On a Functional Contraction Method

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 12
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Henning Sulzbach
aus Frankfurt am Main

Frankfurt am Main 2012
(D30)

Vom Fachbereich 12 der Johann Wolfgang Goethe-Universität angenommen.

Dekan:	Prof. Dr. Tobias Weth
Gutachter:	Prof. Dr. Ralph Neininger Prof. Dr. Svante Janson Prof. Dr. Götz Kersting
Datum der Disputation:	16. 05. 2012

Contents

Acknowledgements

First and foremost, I am deeply grateful to my advisor Ralph Neininger for his extraordinary support, the large amount of fruitful discussions and his encouragement in frustrating moments. His knowledge on the contraction method and its applications has been essential for the outcome of this work. Besides, I owe him my thanks for enabling me to attend several inspiring conferences and workshops.

Secondly, I would like to thank Nicolas Broutin from INRIA, Paris-Rocquencourt, for the fruitful and enduring cooperation which led to some of the major results of this thesis. In particular, I thank him for the opportunity to spend time at INRIA for joint research twice in the last years.

Next, I am much obliged to Svante Janson from Uppsala for bringing valuable references to my attention which have been at the core of this work as well as for further discussions which gave rise to improvements of some results. Moreover I would like to thank him for various comments on the first version of this thesis and in particular for his help in filling a gap in the proof of Donsker's Theorem.

I would also like to express my thanks to Alfredas Račkauskas for pointing out details to me which led to an improved understanding of essential components in my thesis.

There are various people from the university of Frankfurt to whom I owe my thanks. First, I am indebted to all academic teachers whose courses and seminars provided me with the necessary knowledge and tools for my research. My special thanks are due to Götz Kersting and Anton Wakolbinger, who always showed interest in the state of my research as well as both have been available for any question in the context of probability theory during my studies. I am greatly thankful to Jochen Geiger whose courses on "Stochastic processes", "Advanced stochastics" and "Randomized algorithms" awakened my interest for probability theory.
A special word of thanks is owed to Tobias Weth for not hesitating to offer me his help on functional analytic topics as well as to Gaby Schneider for her considerable help with the statistic software R.
I am very grateful to our secretary Nicole Götting for her organizational support and the librarian Cornelia Gilb for her help in locating papers and books which I was not able to get my hands on.
Among my PhD colleagues, my first thanks go to Christian Böinghoff, Arnulf Jentzen and Max Stroh whose blackboards always provided space on which we could discuss some of my problems. Moreover, special thanks go to Benjamin Mühlbauer, Hardy Bosse, Reinhard Steffens, Cordian Riener, Christian Trabandt and Timo de Wolff who provided an enjoyable working atmosphere throughout the past years as well as to Theresa Szczepanski for proofreading parts of this work.
I am also very thankful to my brother Thomas for his help in drawing figures for several talks and the rest of my family for their ongoing and enduring support, especially in the recent years.

Finally, I appreciate the financial support of the German Academic Exchange Service DAAD in the course of the PROCOPE program; my research stays in Paris have proven to be essential for this work.

1. Introduction

Since the middle of the past century, the theory of weak convergence in function spaces has become an important concept in probability theory and its applications. Fundamental contributions from various origins have been made by Kolmogorov [Kol31], Erdős and Kac [EK46], Doob [Doo49] and Donsker [Don51], [Don52]. The theory has been established systematically as it is known today by Prokhorov [Pro53], [Pro56] and Skorokhod [Sko56]. Relying on the latter works, Billingsley offered an accessible account to the area in his book *Convergence of Probability Measures* in 1968. Undoubtedly, it is still the main reference for weak convergence on function or more general metric spaces. Relying on the so-called contraction method, the present thesis is mainly concerned with results that allow to deduce weak convergence in function spaces and applications thereof.

1.1. The idea of contraction

The contraction method is an approach to distributional convergence for sequences of random variables obeying certain recurrence relations on the level of distributions. It has become a powerful tool in the probabilistic analysis of algorithms since its invention in the seminal paper on the running time analysis of the well-known sorting algorithm Quicksort by Rösler [Rös91]. Often, the analysis of divide and conquer algorithms leads to the following recursion for a sequence of random variables (Y_n)

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} A_r(n) Y^{(r)}_{I_r^{(n)}} + b(n), \quad n \geq n_0, \tag{1.1}$$

where $\stackrel{d}{=}$ denotes that left and right hand side are identically distributed, and $(Y_j^{(r)})_{j\geq 0}$ have the same distribution as $(Y_n)_{n\geq 0}$ for all $r = 1, \ldots, K$, where $K \geq 1$ and $n_0 \geq 0$ are fixed integers. Moreover $I^{(n)} = (I_1^{(n)}, \ldots, I_K^{(n)})$ is a vector of random integers in $\{0, \ldots, n\}$. The basic independence assumption is that $(Y_j^{(1)})_{j\geq 0}, \ldots, (Y_j^{(K)})_{j\geq 0}$ and $(A_1(n), \ldots, A_K(n), b(n), I^{(n)})$ are independent. This assumption also determines the law of the right hand side of (1.1). Dependencies between the coefficients $A_r(n)$, $b(n)$ and the integers $I_r^{(n)}$ appear in various applications. Apart from the probabilistic analysis of recursive algorithms, recurrences of the form (1.1) appear in several fields, e.g., in the study of random trees, in branching processes, in the context of random fractals, in models from stochastic geometry and in information and coding theory. For surveys of such occurrences see [NR04b] and [Nei04]. Mostly, (Y_n) is a sequence of real-valued random variables and $A_r(n), b(n)$ are random coefficients also with values in $\mathbb{R}$. However, there is no harm in considering random variables with values in arbitrary vector spaces provided $A_r(n), r = 1, \ldots, K$ denote random linear operators and the right hand side of (1.1) remains a well-defined random variable. Recurrences for $\mathbb{R}^d$ or Hilbert space valued random variables have been treated in the literature and will be reviewed later. In the thesis we develop the contraction method in separable Banach spaces, where we mainly focus on the case $\mathcal{C}[0,1]$, the space of continuous functions on the unit interval endowed with uniform topology. We also give analogous results for $\mathcal{D}[0,1]$, the space of càdlàg functions, that is right-continuous functions with left limits, equipped with Skorokhod topology.

In applications the quantities Y_n grow large as n tends to infinity, an appropriate scaling is typically obtained by centering Y_n and normalizing by the order of the standard deviation. The scaling leads to a recursion similar to (1.1) for the rescaled sequence (X_n):

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \quad n \geq n_0, \tag{1.2}$$

with conditions on identical distributions and independence similar to recurrence (1.1). The coefficients $A_r^{(n)}$ and $b^{(n)}$ in the modified recurrence (1.2) are typically directly computable from the original coefficients $A_r(n)$, $b(n)$ and the scaling, see e.g., for the case of random vectors in $\mathbb{R}^d$, [NR04b, equations (4)].

The main idea: The rough idea of the contraction method is the following: First, it usually follows directly from the coefficients $A_r^{(n)}$ and the sequence $b^{(n)}$ that there exists random operators A_r and a random variable b such that

$$A_r^{(n)} \to A_r, \quad b^{(n)} \to b, \tag{1.3}$$

as $n \to \infty$ in a suitable sense. If also $I_r^{(n)}$ grows large as $n \to \infty$ for all $r = 1, \dots, K$ and it is plausible that the quantities X_n converge, say to a random variable X, then, by letting formally $n \to \infty$, equation (1.2) turns into

$$X \stackrel{d}{=} \sum_{r=1}^{K} A_r X^{(r)} + b, \tag{1.4}$$

with $X^{(1)}, \dots, X^{(K)}$ distributed as X and $X^{(1)}, \dots, X^{(K)}$, $(A_1, \dots, A_k, b)$ independent. The distributional fixed point equation (1.4) will then serve as a characterization of the limiting distribution $\mathcal{L}(X)$. Here $\mathcal{L}(X)$ denotes the distribution of a random variable X. Solutions of (1.4) are usually considered as fixed-points of the following map T which is at the heart of the contraction method:

$$\begin{aligned} T &: \mathcal{M}(B) \to \mathcal{M}(B) \\ T(\mu) &= \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b \right), \end{aligned} \tag{1.5}$$

where $(A_1, \dots, A_K, b), Z^{(1)}, \dots, Z^{(K)}$ are independent and $Z^{(1)}, \dots, Z^{(K)}$ have distribution μ. Here $\mathcal{M}(B)$ denotes the set of probability distributions on the state space B, in which the sequence of random variables (Y_n) [hence also (X_n)] attains their values. Rösler's approach to turn these ideas into rigorous statements consists of the following two steps:

Step 1. Show existence and uniqueness of the solution of the fixed-point equation (1.4) in an appropriate subspace $\mathcal{M}'(B)$ of probability measures on $\mathcal{M}(B)$. To this end endow $\mathcal{M}'(B)$

with a complete metric d that turns T into a contractive self-map on $\mathcal{M}'(B)$. In light of Banach's fixed point theorem one then obtains a unique fixed-point of T, hence a unique solution of (1.4) in $\mathcal{M}'(B)$.

Step 2. Show convergence of X_n to this unique fixed-point in the metric d and infer weak convergence (and possibly more). The proof of the convergence of X_n in the metric d usually runs along similar lines as **Step 1** by contraction arguments. It relies on the quality of convergence of the coefficients in (1.3). The transition to weak convergence depends mainly on properties of d and will turn out to be a hard task in our applications.

We shortly mention the classical case of Quicksort which will be used as a prototype example of the real-valued case throughout the thesis.

Quicksort: Introduced in 1961 by Hoare [Hoa61, Hoa62], Quicksort has become one of the most important sorting algorithms. Its median of three version serves as standard sorting routine in Unix. It is well-known and easily seen that, given a list of n distinct elements from an ordered set, the number of key comparisons Z_n of the standard randomized Quicksort algorithm satisfies a recursion of type (1.1). More precisely, it holds

$$Z_n \stackrel{d}{=} Z^{(1)}_{I_n-1} + Z^{(2)}_{n-I_n} + n - 1, \quad n \geq 1, \tag{1.6}$$

where I_n is uniformly distributed on $\{1, \ldots, n\}, Z_0 = 1$ and conditions of independence as for (1.1). The mean number of comparisons is known explicitly, it holds $\mathbb{E}[Z_n] = 2n \log n + cn + o(n)$ for some real constant c. From (1.6) and the scaling of the form $n^{-1}(Z_n - \mathbb{E}[Z_n])$, it is plausible that a possible limit random variable Z satisfies the following fixed-point equation, that is nowadays known as the Quicksort equation:

$$Z = UZ^{(1)} + (1-U)Z^{(2)} + 2U \log U + 2(1-U)\log(1-U) + 1. \tag{1.7}$$

Here U denotes a random variable on $[0, 1]$ with uniform distribution and conditions as in (1.4) are satisfied. Rösler [Rös91] was able to carry out both **Step 1** and **Step 2** working on the subspace of probability distributions with zero mean and finite variance. Subsequently, we denote this space by $\mathcal{M}_{2,0}(\mathbb{R})$. Endowing $\mathcal{M}_{2,0}(\mathbb{R})$ with the minimal ℓ_2 metric, see (2.8), he proved that the map T in (1.5) has the Lipschitz property with Lipschitz constant bounded by $\sqrt{\mathbb{E}[U^2] + \mathbb{E}[(1-U)^2]} = \sqrt{2/3}$. By the Theorem of Riesz-Fisher completeness of ℓ_2 is easily checked and convergence of the rescaled sequence to the unique solution of (1.7) in $\mathcal{M}_{2,0}(\mathbb{R})$ is shown. One should not forget that the sequence $(n+1)^{-1}(Z_n - \mathbb{E}[Z_n])$ was identified as an L^2-bounded martingale by Régnier [Rég89]. Hence, the convergence was already known at that time.

The ℓ_2 approach: The approach has been established further and applied to a couple of examples in Rösler [Rös92] and Rachev and Rüschendorf [RR95]. Later on general convergence theorems have been derived stating conditions under which convergence of the coefficients of the form (1.3) together with a contraction property of the map (1.5) implies convergence in distribution $X_n \to X$. For random variables in $\mathbb{R}$ with the minimal ℓ_2 metric see Rösler [Rös01], and Neininger [Nei01] for $\mathbb{R}^d$ with the same metric. In $\mathbb{R}^d$ note that the linear operators $A_r(n)$ and $A_r^{(n)}$ coincide with random $d \times d$ matrices. As a prototype result, we cite (a slightly modified and simplified version of) Theorem 3 in [Rös01].

Theorem 1.1. *Let* (X_n) *be a sequence of real-valued random variables satisfying* (1.2) *with* $n_0 = 1$ *such that for all* $r = 1, \ldots, K$ *and* $n \geq 1$

$$0 \leq I_r^{(n)} \leq n-1, \qquad \mathbb{E}\,[b_n] = 0 \qquad \textit{and} \qquad \mathcal{L}(X_0) \in \mathcal{M}_{2,0}(\mathbb{R}).$$

Then $\mathcal{L}(X_n) \in \mathcal{M}_{2,0}(\mathbb{R})$ *for all* n. *Assume further that*

$$L := \sum_{r=1}^{K} \mathbb{E}\left[|A_r|^2\right] < 1 \tag{1.8}$$

and $\mathbb{E}\,[b] = 0, \mathbb{E}\left[b^2\right] < \infty$. *Then* T *is a contraction on* $\mathcal{M}_{2,0}(\mathbb{R})$ *with Lipschitz constant at most* $\sqrt{L}$ *with respect to* ℓ_2. *Hence, it has a unique fixed-point* μ *in* $\mathcal{M}_{2,0}(\mathbb{R})$. *Moreover, if*

$$\mathbb{E}\left[|A_r^{(n)}|^2 \mathbf{1}_{\{I_r^{(n)} \leq m\}}\right] \to 0,$$

for all $r = 1, \ldots, K$ *and* $m \in \mathbb{N}$ *and*

$$(A_1^{(n)}, \ldots, A_K^{(n)}, b^{(n)}) \xrightarrow{\ell_2} (A_1, \ldots, A_K, b),$$

then $\ell_2(X_n, X) \to 0$, *where* X *has distribution* μ.

When applying this result, the scaling of Y_n requires precise asymptotic information on the mean $\mathbb{E}\,[Y_n]$ in advance whereas the order of the standard deviation σ_n may be guessed. Typically, an expansion of the form

$$\mathbb{E}\,[Y_n] = f(n) + o(\sigma_n)$$

turns out to be sufficient. The assertion of Theorem 1.1 can be stated similarly using ℓ_p metrics with $0 < p \leq 1$ assuming only a finite absolute of order p by simply replacing 2 by p in every occurrence. In this case, no expansion of moments of Y_n has to be available in advance. A survey on the contraction method mostly in the context of ℓ_2 metrics including various applications mainly from the area of random trees is given in [RR01].

Limitations of ℓ_p: The ℓ_p approach is restricted in two ways: First, for $0 < p \leq 1$ or $p = 2$, the Lipschitz constant of T (restricted to suitable subspaces) is bounded by

$$\left[\sum_{i=1}^{K} \mathbb{E}\,[|A_i|^p]\right]^{\min(1/p,1)},$$

an analogous result can not be obtained for $1 < p \leq 2$ or $p > 2$ along the same lines. In general, only a bound of the form $\mathbb{E}\left[\left(\sum_{i=1}^{K} |A_i|\right)^p\right]^{1/p}$ can be given easily. This term is clearly increasing in p which is illustrated by the following example.

Example: Consider the fixed-point equation

$$X \stackrel{d}{=} \frac{X + X'}{\sqrt{2}}, \tag{1.9}$$

where X' is an independent copy of X. Assuming a finite second moment of X, it immediately follows from the central limit theorem that X has normal distribution with zero mean. Historically, Pólya [Pól23] was the first to observe this in a remarkable paper from 1923. Later, based on characteristic functions, Laha and Lukacs [LL65] removed the assumption on the finite second moment, see also [KPS96] for a purely probabilistic proof. However, no metric of ℓ_p type seems to provide the contraction property of T.

The second problem is concerned with the generalization of the approach to infinite dimensional spaces. In separable Hilbert spaces, the bound (1.8) remains valid if we restrict to the zero mean case. In [RR95] and [Rüs06], similar ideas are discussed also for $1 < p \leq 2$ in general Banach spaces. Arguments based on Woyczynski's inequality, see [Woy80, Proposition 2.1], imply a bound on the Lipschitz constant of the form $c_p \sum_{i=1}^{K} \mathbb{E}\left[|A_i|^p\right]$ with a positive constant c_p only depending on p which turns out to be too large for applications. Rüschendorf [Rüs06] showed uniqueness of the fixed-point equation (under mean zero and finite p-th moment condition) if only the more natural condition $\sum_{i=1}^{K} \mathbb{E}\left[|A_i|^p\right] < 1$ is satisfied. However, the result is given in the real-valued case and the ideas only extend directly to Banach spaces of type $1 < p \leq 2$. In our main application $\mathcal{C}[0,1]$ endowed with the uniform topology, which is a Banach space only of trivial type 1, we do not know of any useful generalizations of the ℓ_p approach for $1 < p \leq 2$.

The Zolotarev distances: In the context of the contraction method, various other probability metrics, among them the class of ζ_s metrics that is also used in the present work, have been mentioned first in [RR95], see in particular Proposition 1. ζ_s metrics are also used in [RR01, Theorem 5]. The approach has been worked out to its full extent for random variables in $\mathbb{R}^d$ by Neininger and Rüschendorf [NR04b]. The family of metrics of Zolotarev type which we study intensively in Chapter 2 has proved to be more flexible than the classical ℓ_p metrics, the main improvement being a reduction of the Lipschitz constant of T for arbitrary $s > 1$ to

$$\sum_{i=1}^{K} \mathbb{E}\left[|A_i|^s\right].$$

Note that Theorem 4.1 in [NR04b] naturally extends Theorem 1.1. Thus, various problems with normal limit laws could be solved in [NR04b]. As it will turn out later, for $s > 2$, this approach requires an exact scaling of Y_n, in particular a first order expansion of the standard deviation has to be known a priori. An important case is when $A_1(n) = \ldots = A_K(n) = 1$ for all n in (1.1), see [NR04b, Section 5] for many examples. Here, the ζ_s approach for $s > 2$ gives normal limit laws for the rescaled quantity (X_n) if

$$\mathbb{E}\left[Y_n\right] = f(n) + o(g^{1/2}(n)), \qquad \mathbf{Var}\left[Y_n\right] = g(n) + o(g(n)).$$

This results confirms a heuristic principle by Pittel [Pit99] where he proposes that the first two moments accompanied by a linear recurrence might be sufficient to obtain normal limit laws. Note that the contraction method relying on ζ_s with $s > 2$ yields normal limits in all applications known so far. Yet another indication for the flexibility of Zolotarev's metrics is given in [NR04a] where normal limit laws for sequences (X_n) satisfying recurrences leading to degenerate fixed-point equations of type $X \stackrel{d}{=} X$ under certain properties of the moments of Y_n are given.

Recent results: An extension of the method and results in [NR04b] to continuous time, i.e., to processes $(X_t)_{t\geq 0}$ satisfying recurrences similar to (1.2) was given in Janson and Neininger [JN08]. For the case of random variables in a separable Hilbert space leading to functional limit laws general limit theorems for recurrences (1.1) have been developed in Drmota, Janson and Neininger [DJN08]. The main application there was a functional limit law for the profile of random trees whose analysis was carried out by an encoding in the so-called profile-polynomial, a generating function of finite degree. This approach led to random variables in the Bergman space of square-integrable analytic functions on a domain in the complex plane. It is remarkable that this approach combines two different methods of analyzing distributional recurrences, namely manipulations with generating functions and the contraction method. In Eickmeyer and Rüschendorf [ER07] general limit theorems for recurrences in $\mathcal{D}[0,1]$ under the L_p-topology were developed. Note, that the uniform topology for $\mathcal{C}[0,1]$ and the Skorokhod topology for $\mathcal{D}[0,1]$ considered in the present work are considerably stronger than the L_p-topology. In $\mathcal{C}[0,1]$, the uniform topology provides more continuous functionals such as the supremum $f \mapsto \sup_{t\in[0,1]} f(t)$ or projections $f \mapsto f(s_1,\ldots,s_k)$, for fixed $s_1,\ldots,s_k \in [0,1]$, to which the continuous mapping theorem can be applied. In $\mathcal{D}[0,1]$ those functionals may also be applied once the limit random variable is known to have continuous sample paths. Note that our approach in $\mathcal{D}[0,1]$ is limited to sequences of random functions with càdlàg paths whose limits have continuous sample paths.

All results based on ζ_s metrics in the context of the contraction method in $\mathbb{R}^d$ or in separable Hilbert spaces rely on the fact that convergence in ζ_s implies weak convergence. However, for general Banach spaces this is not true. Counterexamples have been reported in Bentkus and Rachkauskas [BR85], we give explicit examples in Section 2.4. Furthermore, completeness of the ζ_s metrics on appropriate subspaces of $\mathcal{M}(B)$ is known only for separable Hilbert spaces, see [DJN08, Theorem 5.1]. Summarizing, we face the following major difficulties in the framework of the contraction method using metrics of Zolotarev type in the case of continuous or càdlàg functions on the unit interval.

P1. Distributional convergence can only be inferred by ζ_s convergence under further conditions.

P2. For $s > 1$, by the lack of completeness of ζ_s, a fixed-point of (1.4) has to be found by different means.

P3a. For $1 < s < 2$, the scaling and the convergence of coefficients (1.3) typically require the convergence of $\mathbb{E}[X_n(t)]$ to $\mathbb{E}[X(t)]$ uniformly in t. Moreover, a rate of convergence is needed to solve **P1** which will later be clear.

P3b. When applying the contraction method with $s > 2$, exact scaling is required, i.e. the covariance function of the sequence (X_n) has to be independent of the time parameter n.

Outline - Chapters 2 and 3: We investigate the family of ζ_s metrics in Chapter 2. In Sections 2.2 and 2.3 we give upper and lower bounds on the metrics in general Banach spaces in particular in terms of the family of ℓ_p metrics that is also introduced here. We discuss counterexamples where convergence in ζ_s does not imply weak convergence in Section 2.4. An appropriate formulation of the Zolotarev distance in non-separable spaces, in particular for $\mathcal{D}[0,1]$, is given in Section 2.5. From Section 2.6 on, we concentrate on the cases of $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ where we provide a solution to **P1** by additional assumptions on the rate of convergence and on the regularity of sample paths of the sequence (X_n). Here, the key ingredient is a result in Barbour [Bar90] in the context of an extension of Stein's method to càdlàg processes. We also prove that under the same conditions on the rate of convergence and sample paths regularities, the sequence $(\|X_n\|^s)_{n\geq 0}$ is uniformly integrable. This gives rise to moment convergence of the supremum, a very useful result in applications. Chapter 3 is devoted to the framework of the contraction method in separable Banach spaces, our main result Theorem 3.6 in the case of $\mathcal{C}[0,1]$ or $\mathcal{D}[0,1]$ is to be compared with Theorem 4.1 in [NR04b]. The rate of convergence in the Zolotarev metric needed to infer weak convergence is guaranteed by a transfer theorem from the rate of convergence of the coefficients (1.3). Here, convergence of the sequence of linear operators $A_r^{(n)}$ is with respect to the operator norm. Finally, we point out that the use of ζ_s metrics in the càdlàg space requires the limit to have continuous paths since we have no arguments to deduce distributional convergence otherwise. Moreover note that we deal with the Skorokhod instead of the uniform topology on $\mathcal{D}[0,1]$ solely for reasons of measurability, see the beginning of Section 2.5.

The remaining problems **P2**, **P3a** and **P3b** will be discussed in the second part of the thesis, we only mention **P2** here. The obvious method for finding a solution of (1.4) is by considering the infinite iteration of T. This approach is taken in Chapter 5, where the main difficulties are the verification of continuity and integrability of the supremum of the limit. For $s > 2$, as in the real-valued case, one may guess a solution. This is what we do in Chapter 4.

1.2. Applications

We present applications of the ideas and results of Chapter 3 for recurrences of type (1.1) both in the case $1 < s \leq 2$ and $s > 2$ in the second part of this thesis. As a start, we provide a new and considerably short proof of the classical invariance principle due to Donsker based on recursive time-decompositions of Brownian motions and sums of independent random variables in Chapter 4. Here, the ζ_s approach is worked out in the case of $s > 2$, a way to surmount the difficulties in **P3b** is given by using a piecewise linear interpolation of the Brownian motion.

In the other case, our main result is concerned with partial match queries in random quadtrees and K-d trees. These tree models, introduced by Finkel and Bentley in [FB74], resp. Bentley in [Ben75], serve as comparison-based data structures for multidimensional databases and may be considered as multidimensional generalizations of binary search trees. Higher-dimensional databases arise in various contexts such as computer graphics, computational geometry, geographical information systems and statistical analysis. Using bit representations for the data, digital structures such as tries, digital search trees or patricia tries provide alternatives allowing efficient solutions for retrieval problems. For a general account on multidimensional data structures and their applications, we refer to the series of monographs by Samet [Sam90a, Sam90b, Sam06].

The Quadtree: The quadtree extends the comparison based construction of binary search trees to higher dimensions. For d-dimensional data, the corresponding quadtree has branching number at most 2^d where each dimension causes a factor of 2. For convenience, we assume the data field to be the unit cube $[0,1]^d$. For points $p_1, \dots, p_n \in [0,1]^d$ the tree is constructed as follows: The first point p_1 is inserted at the root, it splits the unit cube into 2^d subregions according to its coordinates. Each subregion corresponds to one of its children. The construction then goes on recursively: Having inserted i points in the tree, the unit cube is covered by $1 + i(2^d - 1)$ subregions each corresponding to an external node in the tree. Point p_{i+1} is then stored at the node u that corresponds to the subregion in which it falls. Insertion divides this region into 2^d new subregions assigned to the children of u.

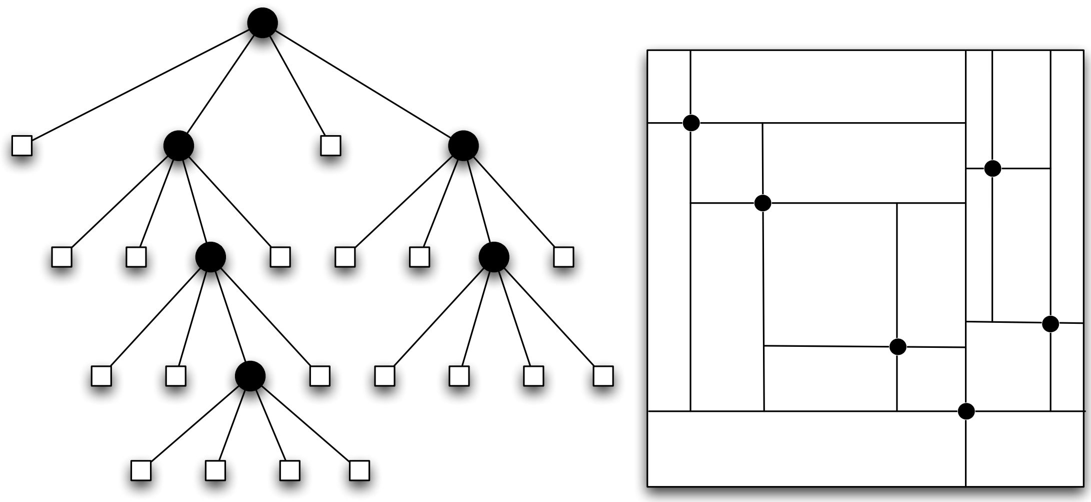


Figure 1.1.: A quadtree of size $n = 6$ with $d = 2$. External nodes are indicated by boxes which correspond to regions in the partition on the right.

Searching: By far the most important property of binary search trees (and variants thereof) as data structure is the fact that these trees are typically well-balanced under reasonable assumptions on the data. Insertion, deletion, searching or retrieving specific data usually requires only logarithmic time. The same is true in quadtrees if one aims at finding completely specified patterns. Under the uniform model, that we will always assume throughout the thesis, the quadtree is generated by a sequence of independent random variables uniformly distributed on the unit cube. For the insertion depth D_n of the n-th node that measures the time for an unsuccessful search in the quadtree, we have

$$\mathbb{E}\left[D_n\right] \sim \frac{2}{d} \log n. \tag{1.10}$$

This has been proved independently in [FGPR93] by means of generating functions and singularity analysis and in [DL90] by probabilistic arguments. In [FL94], the order of the variance of D_n is identified for all $d \geq 2$ and asymptotic normality of D_n after normalization is shown. For the height H_n of the tree that corresponds to worst case search times, it holds

$$\mathbb{E}\left[H_n\right] \sim \frac{c}{d} \log n, \tag{1.11}$$

where $c = 4.31107\ldots$ is the constant known from the height of random binary search trees satisfying $ce^{1/c} = 2e$. This result is due to Devroye [Dev87]. The behaviour of K-d trees, being introduced in Section 5.6, is included in these results choosing $d = 1$ in (1.10) and (1.11) since they coincide with random binary search trees in distribution. It is worthwhile noting that an analogous result is not true for quadtrees: Here, insertion is not performed uniformly at random among the available external nodes but nodes on lower levels are favoured.

Partial match retrieval: In partial match queries, one is interested in finding all data matching a fixed pattern that only imposes some constraints on the data field. Such a query visits considerably more nodes in the tree. The first investigations in partial match retrieval is due to Rivest [Riv76] based on digital structures. For partial match queries with $1 < s < d$ specified components in a quadtree of size n, Bentley and Stanat proposed $n^{1-s/d}$ as the order of magnitude for the cost of the retrieval algorithm, i.e. the number of traversed nodes, see [BS75] (and [Ben75] for a similar statement in the model of K-d trees). This claim was disproved by Flajolet et al. [FGPR93] in the quadtree model (and in [FP86] for K-d trees). The main result in this paper states that

$$\mathbb{E}\left[C_n^{s,d}(\xi^{(s)})\right] \sim \gamma_{s,d} n^{\alpha-1} \tag{1.12}$$

for some constant $\gamma_{s,d} > 0$ and $\alpha \in [1, 2]$ solving

$$\alpha^{d-s}(\alpha+1)^s = 2^d. \tag{1.13}$$

Here $C_n^{s,d}(x)$, $x = (x_1, \ldots, x_s) \in [0,1]^s$, denotes the cost, i.e. the number of visited nodes, for a partial match query with s specified components equalling $x_1, \ldots, x_s$ and $\xi^{(s)}$ is assumed to be uniform on $[0,1]^s$, independent of the quadtree. Note that $\alpha > 2 - s/d$. We give a simple heuristic argument for the occurrence of the constant α at the beginning of Chapter 5. The result (1.12) has been strengthened by Chern and Hwang in [CH03] where the value of the leading constant $\gamma_{s,d}$ and rates of convergence are provided in all dimensions. Distributional limit laws and asymptotics of the variance at a uniform query line have been studied by several authors; however, neglecting subtle dependencies between the contributions of subtrees, the order of the variance, higher moments and a limit law have remained open and will be solved in this thesis for the case $d = 2$. It is worthwhile noting that, concerning partial match retrieval, comparison-based structures are outperformed by digital structures. In multidimensional tries, called K-d tries in [FP86] and also quadtries in [DZC04], the order of the average cost, i.e. the number of bit comparisons, of partial match retrieval is indeed $n^{1-s/d}$, see [FP86].

The behaviour changes dramatically on the boundary where we enforce at least one coordinate to be zero. A search for those lowest points in the tree visits considerably less nodes, we denote by $R_n^{s,d}$ the cost of the retrieval algorithm if s components are chosen to be zero and $d-s$ components are left unspecified. Then Flajolet et al. [FGPR93] proved

$$\mathbb{E}\left[R_n^{d-1,d}\right] \sim \eta_{d-1,d} n^{\alpha-1}, \qquad \alpha = 2^{1/d},$$

for some constant $\eta_{d-1,d} > 0$, where $\eta_{1,2}$ is explicitly known to be $\Gamma(2\alpha)/(\alpha\Gamma^3(\alpha))$. Here $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ denotes the Gamma function for $x > 0$. We do not know of any result in the case of general s. Our heuristic approach at the beginning of 5 which can be worked out analogously here, suggests that $R_n^{s,d}$ is of order n^α with $\alpha = 2^{1-s/d}$. For $d = 2$, a limit law for $R_n^{1,2}$ based on fragmentation processes is given in [CJ11], a proof thereof could also be obtained by means of the contraction method based on the results in [FGPR93].

The two-dimensional case: In the thesis, we only treat two-dimensional quadtrees and 2-d trees. The following observation is crucial for our approach: In a partial match query with fixed first component $s \in [0,1]$, any node in the tree is visited if and only if the region it is inserted in is intersected by the horizontal line $x = s$. Moreover this is the case if and only if the horizontal line $x = s$ crosses the vertical coming from the node. We abbreviate $C_n(s) := C_n^{1,2}(s)$ for the

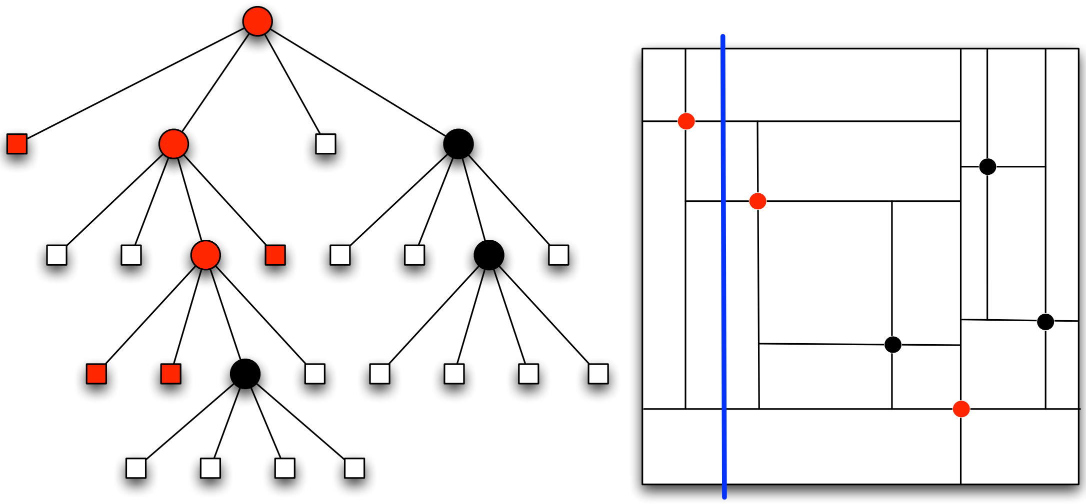


Figure 1.2.: A partial match retrieval with fixed query line $x = 0.2$. Visited internal nodes are marked red and so are external nodes that correspond to the four regions intersected by the line.

number of vertical lines in the partition of the unit square governing the quadtree that intersect the horizontal line at $x = s$. The constant $\gamma_{1,2}$ was already known in [FGPR93], we will subsequently denote it by κ. In the two-dimensional case Chern and Hwang [CH03] provide an expansion for the mean of arbitrary order; for our purposes, it will be sufficient to use

$$\mathbb{E}\left[C_n(\xi)\right] = \kappa n^{\beta} - 1 + O(n^{\beta-1}), \tag{1.14}$$

where

$$\kappa = \frac{\Gamma(2\beta+2)}{2\Gamma^3(\beta+1)}, \qquad \beta = \frac{\sqrt{17}-3}{2}, \tag{1.15}$$

$\xi = \xi^{(1)}$ and $\beta = \alpha - 1$. Recently, Curien and Joseph [CJ11] were the first to give a result on the mean of $C_n(s)$ for a fixed query line $x = s$. Based on fragmentation processes, using (1.12) and an ingenious coupling argument for Markov chains, they proved convergence of $C_n(s)/n^{\beta}$ in mean for fixed s. Based on fixed-point arguments for the limit, their main result states that

$$n^{-\beta}\mathbb{E}\left[C_n(s)\right] \sim K_1 h(s), \tag{1.16}$$

where

$$K_1 = \frac{\Gamma(2\beta+2)\Gamma(\beta+2)}{2\Gamma^3(\beta+1)\Gamma^2(\beta/2+1)}, \qquad h(s) = (s(1-s))^{\beta/2}. \tag{1.17}$$

Note that we use a refinement of their method and result in our work.

Recursive decomposition and main results: Given the number of points in each of the four regions (equivalently, nodes in the subtrees), those points are again independent and uniformly distributed. Moreover, the behaviour within each subregion is independent of each other. A partial match query with fixed first component $x = s$ then searches in two of these subtrees, or equivalently, certain horizontal lines are counted either in the first and second or in the third and fourth subregion. Here we abbreviate the first subtree to correspond to the lower-left region, the second to the upper-left, the third to the lower-right and the fourth to the upper-right. Note, that the query line in the regions under investigation has to be considered relative to the x-component of the first inserted node. Thus, a decomposition at the root gives the following fundamental distributional recurrence

$$\begin{aligned} C_n(s) \quad &\stackrel{d}{=} \quad 1 + \mathbf{1}_{\{s<U\}} \left[C^{(1)}_{I_1^{(n)}}\left(\frac{s}{U}\right) + C^{(2)}_{I_2^{(n)}}\left(\frac{s}{U}\right) \right] \\ &\quad\quad + \mathbf{1}_{\{s\geq U\}} \left[C^{(3)}_{I_3^{(n)}}\left(\frac{1-s}{1-U}\right) + C^{(4)}_{I_4^{(n)}}\left(\frac{1-s}{1-U}\right) \right]. \end{aligned} \tag{1.18}$$

Here, U, V denote the components of the first inserted point, $I_1^{(n)}, \ldots, I_4^{(n)}$ denote the number of points in the subregions and $(C_n^{(1)}), \ldots, (C_n^{(4)})$ are independent copies of (C_n), independent of $\left(U, V, I_1^{(n)}, \ldots, I_4^{(n)}\right)$. Obviously, (1.18) can not be seen as a one-dimensional recurrence of type (1.1) for fixed s. The crucial observation is that the recursive decomposition holds simultaneously for an arbitrary finite number of coordinates $s_1, \ldots, s_k$. Thus, partial match queries for different values of s are coupled and considered in one and the same quadtree! Naturally, there exist càdlàg versions of $(C_n(s))_{s\in[0,1]}$, hence (1.18) can be viewed as a recursion in the space of càdlàg functions. Given (U, V), we have

$$\mathcal{L}\left(I_1^{(n)}, \ldots, I_4^{(n)}\right) = \mathrm{Mult}(n-1; UV, U(1-V), (1-U)V, (1-U)(1-V)). \tag{1.19}$$

Hence scaling by n^β suggests that any limit process Z of the scaled quantity $n^{-\beta}C_n(s)$ satisfies

$$\begin{aligned} Z(s) \quad &\stackrel{d}{=} \quad \mathbf{1}_{\{s<U\}} \left[(UV)^\beta Z^{(1)}\left(\frac{s}{U}\right) + (U(1-V))^\beta Z^{(2)}\left(\frac{s}{U}\right) \right] \\ &\quad + \mathbf{1}_{\{s\geq U\}} \left[((1-U)V)^\beta Z^{(3)}\left(\frac{1-s}{1-U}\right) + ((1-U)(1-V))^\beta Z^{(4)}\left(\frac{1-s}{1-U}\right) \right]. \end{aligned} \tag{1.20}$$

Based on the contraction method, our main result of Chapter 5 states that the process $n^{-\beta}C_n(s)$ indeed converges in distribution in the space of càdlàg functions endowed with Skorokhod topology to a continuous solution of (1.20) which is unique under the constraints that its mean at ξ equals κ, see (1.15), and its supremum is square-integrable. For a simulation of the limit, see Figure 1.3. A direct consequence of the result is a limit law for $C_n(\xi)$ for uniform ξ, independent of the quadtree and a first order expansion of the variance where we also identify the leading constant. Our result also implies distributional convergence of the rescaled supremum of $C_n(s)$ which can be strengthened to convergence of all moments. This provides asymptotic information on the worst-case behaviour of the algorithm and solves several long-open problems. Finally note that the costs of partial match queries in quadtrees are not concentrated, typical fluctuations are of the order of their mean. The behaviour in K-d tries is different. Here, under the symmetric Bernoulli model, costs are extremely stable, see [DZC04]. For results on the variance and a normal limit law (after normalization) for the costs in K-d tries see [Sch95, Sch00].

Outline - Chapters 4 and 5: In Section 4.1, we start with the proof of Donsker's theorem relying on the contraction method. It is based on a fixed-point characterization of the Wiener measure that we strengthen in Section 4.2. We also provide convergence of moments of the supremum of the rescaled random walk when assuming corresponding moments for the increments.
In Section 5.1 we collect all results for the retrieval problem on quadtrees, that is process convergence of C_n after rescaling, convergence of the supremum in distribution and with all moments and a characterization of all one-dimensional marginals of Z in terms of a single probability distribution on the real line. The proof of our main result is given in Section 5.2. Solutions for the problems **P2** and **P3a** are provided in Sections 5.3 resp. 5.4. In the latter section we also give further illustrations for the occurrence of size-biasing effects that play a major role in this context. In Section 5.6 we introduce 2-d trees and give analogous results for this class of two-dimensional trees. Section 5.7 is devoted to further open problems in the partial match retrieval problem. Finally, in Section 5.8, we present a problem left open in [CLG11] from the theory of random recursive triangulations that exhibits similar behaviour as the problem of partial match queries. Based on the methods we develop in this thesis, a proof thereof seems to be within reach.

Several parts of the Chapter 5 have recently been published as an extended abstract in [BNS12].

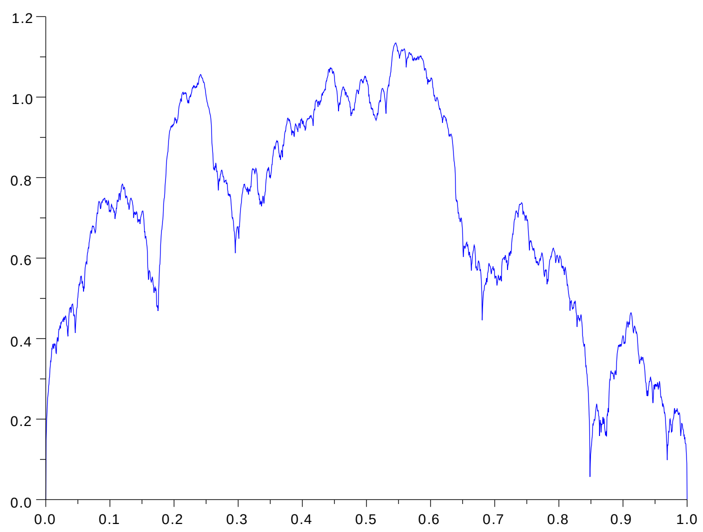


Figure 1.3.: Simulation of the limit process Z established by Nicolas Broutin.

2. The Zolotarev metric

In his seminal work on probability metrics [Zol78], Zolotarev observes that "In probability theory and its applications where there are especially many approximation problems, the use of metrics [...] is also a common occurrence, although we must note that the arsenal employed in this connection is not very large." To this end, he justifies the use of metric distances by the following three rules which we cite from [Zol76].

i) In approximation problems, a metric formulation of a problem is preferable to others equivalent to it because it enables one to consider the question of obtaining quantitative estimates for the approximations.

ii) In a metric formulation of a problem, a cardinal question is that of making a proper choice of the metric used since the naturalness and completeness of a solution will depend on this in many respects.

iii) If we have at our disposal a solution to an approximation problem in terms of some metric, then going over to other metrics in the same problem can be accomplished by means of a comparison of the metrics in the form of estimates for the metric with the help of others.

Based on these guidelines, Zolotarev introduces plenty of different metrics in his works in the late seventies. Additionally, he finds relations between them and applications to justify each of his rules. This *method of metric distances* was later investigated by various other researchers. Zolotarev' rules also play an important role throughout this thesis.
We start by introducing the class of *Zolotarev metrics* ζ_s which we use for the purpose of the contraction method. It was invented by Zolotarev [Zol78, Zol76] [1]. In [Zol77], he summarizes lots of its properties and gives further results. From his remaining works in the context of probability metrics we use results from [Zol79a, Zol79b] and the survey article [Zol84]. In the real-valued case, one should also take into account the comprehensive book [Zol97]. Note that this class of metrics is only one possible choice and it would be very enlightening to find other distances that can serve in this area.

2.1. Definition and main properties

We aim at considering random variables taking their values in a real vector space B that can be endowed with a complete norm $\|\cdot\|$ which turns $(B, \|\cdot\|)$ in a Banach space. There are several reasons why the concept of non-separable Banach spaces equipped with the usual Borel-σ-algebra as state spaces for random variables is inappropriate, for more details see the beginning of Section 2.5. There we will endow B with a considerably smaller σ-algebra. Henceforth, for

[1] Indeed, [Zol78] was his first work on probability metrics and had already appeared 1976 in Russian language in *Mat. Sb. (N.S.)*,101(143)(3):416–454, 1976.

the current section and Sections 2.2 and 2.3, we will always assume $(B, \|\cdot\|)$ to be a separable Banach space equipped with Borel-σ-algebra $\mathcal{B}$. Moreover, we will, here and in the case of non-separable B, always endow $\mathbb{R}^d$ for $d \geq 1$ with the usual Euclidean norm and Borel-σ-algebra. We denote by $\mathcal{M}(B)$ the set of probability measures on B. For functions $f : B \to \mathbb{R}$ which are Fréchet differentiable we denote the derivative of f at a point x by $Df(x)$. Observe that $Df(x)$ is an element of the topological dual $L(B, \mathbb{R})$ of continuous linear forms on B. We also consider higher order derivatives, where $D^m f(x)$ denotes the mth derivative of f at point x and $D^m f$ is a continuous multilinear form on B. Note that the space of multilinear functions $g : B^m \to \mathbb{R}$ is equipped with the norm

$$\|g\| = \sup_{\|h_1\|\leq 1,\ldots,\|h_m\|\leq 1} \|g(h_1, \ldots, h_m)\|.$$

For a comprehensive account on differentiability in Banach spaces we refer to book of Cartan [Car71]. Subsequently, $s > 0$ is fixed and for $m = \lceil s \rceil - 1, \alpha = s - m$ we define

$$\mathcal{F}_s := \{f : B \to \mathbb{R} : \|D^m f(x) - D^m f(y)\| \leq \|x - y\|^\alpha \, \forall \, x, y \in B\} \tag{2.1}$$

Definition 2.1. *Let $s > 0$ and $m = \lceil s \rceil - 1, \alpha = s - m$. For $\mu, \nu \in \mathcal{M}(B)$ the* Zolotarev distance *between μ and ν is defined by*

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}\,[f(X) - f(Y)]\,|, \tag{2.2}$$

where X and Y are B-valued random variables with $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$. X.

A priori the expression (2.2) is not necessarily finite or well-defined. A simple application of Taylor's formula, see Lemma 2.9, shows that (2.2) is well-defined, if

$$B_s(\mu, \nu) := \int \|x\|^s d\mu + \int \|x\|^s d\nu < \infty \tag{2.3}$$

and moreover finite if additionally

$$\int f(x, \ldots, x) d\mu = \int f(x, \ldots, x) d\nu, \tag{2.4}$$

for any continuous k-linear form on B and $1 \leq k \leq m$.

Let μ, ν be probability measures on B satisfying condition (2.3) and f be a k-linear form on B with $k \leq m$ such that

$$\int f(x, \ldots, x) d\mu \neq \int f(x, \ldots, x) d\nu.$$

Then, as the mth derivative of f is constant, the function Cf belongs to $\mathcal{F}_s$ for any $C > 0$. Thus, we have

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}\,[f(X) - f(Y)]\,| = \infty.$$

As a consequence we can say that, for any probability measures μ, ν satisfying (2.3), finiteness of $\zeta_s(\mu, \nu)$ is equivalent to condition (2.4). Moreover, for various Banach spaces, such as Hilbert spaces or $\mathcal{C}[0, 1]$, we have the following: If μ, ν satisfy (2.3) then condition (2.4) can be replaced by

$$\int f_1(x) \cdots f_k(x) d\mu(x) = \int f_1(x) \cdots f_k(x) d\nu(x), \tag{2.5}$$

for all $1 \leq k \leq m$ and continuous linear forms $f_1, \ldots, f_k$. We discuss this property in more detail in Section 3.2 and only give a proof in the Hilbert space case here. Note that this equivalence does not hold for arbitrary Banach spaces, see [JK].

Lemma 2.2. *Let B be a separable Hilbert space with scalar product $< \cdot, \cdot >$ and μ and ν be two probability measures on B satisfying* (2.3)*. Then, for any $k \in \mathbb{N}$, the conditions* (2.4) *and* (2.5) *are equivalent.*

Proof. Let $(e_i)_{i\geq 1}$ be an orthonormal basis of B and f be a continuous bilinear form on B. Then, using the Riesz representation theorem, we have

$$f(x, y) =< Ax, y >= \sum_{i\geq 1} < Ax, e_i >< y, e_i >= \sum_{i\geq 1} < x, v_i >< y, e_i >$$

for some continuous linear operator A and sequence $(v_i)_{i\geq 1}$ in B. The theorem of dominated convergence in connection with Parseval's identity and the Cauchy-Schwarz inequality implies

$$\int f(x, x)d\mu(x) = \sum_{i\geq 1} \int < x, v_i >< x, e_i > d\mu(x)$$

since $\int \|x\|^2 d\mu(x) < \infty$. This shows the assertion for $k = 2$. The remaining cases follow analogously. □

For μ, ν satisfying (2.3), the definition of $\zeta_s(\mu, \nu)$ does not involve the common distribution of μ and ν, hence we will use the abbreviation

$$\zeta_s(X, Y) := \zeta_s(\mathcal{L}(X), \mathcal{L}(Y))$$

for random variables X, Y in B with finite absolute moments of order s. Let $\mathcal{M}_s(B)$ be the subset of $\mathcal{M}(B)$ of distributions μ such that $\int \|x\|^s d\mu(x) < \infty$. We fix a probability measure $\nu \in \mathcal{M}_s(B)$ and denote by $\mathcal{M}_s(\nu)$ the set of all $\mu \in \mathcal{M}_s(B)$ such that (2.3) and (2.4) are satisfied. The first Lemma follows directly from the Definition.

Lemma 2.3. *ζ_s is a pseudometric on $\mathcal{M}_s(\nu)$.*

The next Lemma exhibits a very useful property of ζ_s for the purpose of recursive decompositions of stochastic processes. It is Theorem 3 in [Zol77].

Lemma 2.4. *Let B' be a Banach space and $g : B \to B'$ be a linear and continuous operator. Then we have*

$$\zeta_s(g(X), g(Y)) \leq \|g\|^s \zeta_s(X, Y).$$

for $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ where $\|g\|$ denotes the operator norm of g, i.e. $\|g\| = \sup_{\|x\|\leq 1} \|g(x)\|$.

Proof. Note that g is bounded. It suffices to show that

$$\{\|g\|^{-s} f \circ g : f \in \mathcal{F}'_s\} \subseteq \mathcal{F}_s,$$

where $\mathcal{F}'_s$ is defined analogously to $\mathcal{F}_s$ in B'. Let $f \in \mathcal{F}_s$ and $\eta := \|g\|^{-s} f \circ g$. Then η is m-times continuously differentiable and we have $D^m\eta(x) = \|g\|^{-s} (D^m(f(g(x))) \circ g^{\otimes m}$ for $x \in B$. Here, $g^{\otimes m} : B^m \to (B')^m$ denotes the tensor product $g^{\otimes m}(h_1, \ldots, h_m) = (g(h_1), \ldots, g(h_n))$. This implies

$$\begin{aligned}
\|D^m\eta(x) - D^m\eta(y)\| &= \|g\|^{-s} \|(D^m f(g(x))) \circ g^{\otimes m} - (D^m f(g(y))) \circ g^{\otimes m}\| \\
&\leq \|g\|^{-\alpha} \|g(x) - g(y)\|^\alpha \\
&= \|g\|^{-\alpha} \|g(x - y)\|^\alpha \leq \|x - y\|^\alpha.
\end{aligned}$$

The assertion follows. □

Another basic property is that ζ_s is $(s,+)$ ideal. This is Lemma 3 in [Zol76].

Lemma 2.5. *ζ_s is ideal of order s in $\mathcal{M}_s(\nu)$ for any $\nu \in \mathcal{M}_s(B)$, that is*

$$\begin{aligned} \zeta_s(cX, cY) &= |c|^s \zeta_s(X,Y), \\ \zeta_s(X+Z, Y+Z) &\leq \zeta_s(X,Y) \end{aligned}$$

for $c \in \mathbb{R}\backslash\{0\}$, $(X,Y), Z$ independent, $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ and $\mathcal{L}(Z) \in \mathcal{M}_s(B)$.

Proof. The map $h_c : B \to B, h_c(x) = cx$ is a linear and continuous operator for any $c \in \mathbb{R}$. Hence, by Lemma 2.4

$$\zeta_s(cX, cY) \leq |c|^s \zeta_s(X,Y).$$

Applying the Lemma with $h_{1/c}$ gives the inequality in the other direction. For any $z \in B, f \in \mathcal{F}_s$ the map $g_z : B \to B, g_z(x) = f(x+z)$ is also element of $\mathcal{F}_s$. Conditioning on the value of Z yields $|\mathbb{E}[f(X+Z) - f(Y+Z)]| \leq \zeta_s(X,Y)$ for all $f \in \mathcal{F}_s$ which implies the second assertion of the lemma. □

The Lemma directly implies the following corollary by an adaption of the triangle inequality.

Corollary 2.6. *Let $\mathcal{L}(X_1), \mathcal{L}(Y_1) \in \mathcal{M}_s(\nu_1)$ and $\mathcal{L}(X_2), \mathcal{L}(Y_2) \in \mathcal{M}_s(\nu_2)$ with arbitrary $\nu_1, \nu_2 \in \mathcal{M}_s(B)$ such that (X_1, Y_1) and (X_2, Y_2) are independent. Then*

$$\zeta_s(X_1 + X_2, Y_1 + Y_2) \leq \zeta_s(X_1, Y_1) + \zeta_s(X_2, Y_2).$$

We want to give a result similar to Lemma 2.4 where the linear operator may also be random itself. We focus on the case that B' either equals B or $\mathbb{R}$ where an extension to $\mathbb{R}^d$ for $d > 1$ is straightforward. Let $B^* = L(B, \mathbb{R})$ be the topological dual of B and $\widehat{B}$ be the space of all continuous linear functions from B to B (continuous endomorphisms). Endowed with the norms

$$\|a\| = \sup_{x \in B, \|x\| \leq 1} |a(x)|, \qquad \|b\| = \sup_{x \in B, \|x\| \leq 1} \|b(x)\|,$$

for $a \in B^*, b \in \widehat{B}$ both spaces are Banach spaces. However, these spaces are typically non-separable, hence not suitable for purposes of measurability. Therefore, we will equip them with considerably smaller σ-algebras. We start with the dual space: Similarly to the weak-* topology, we let $\mathcal{B}^*$ be the σ-algebra on B^* that is generated by all norm-continuous linear forms φ on B^* [that is elements of the bidual B^{**}] of the form $\varphi(a) = a(x)$ for some $x \in B$. Note that the set of these continuous linear forms coincides with the bidual B^{**} if and only if B is reflexive, a property that is not satisfied in our applications. We move on to $\widehat{B}$ and define $\widehat{\mathcal{B}}$ to be generated by all norm-continuous linear functions ψ from $\widehat{B}$ to B of the form $\psi(a) = a(x)$ for some $x \in B$. By Pettis' Theorem, we have $\mathcal{B} = \sigma(\ell \in B^*)$. Hence, if $S \subseteq B^*$ with $\mathcal{B} = \sigma(\ell \in S)$, then $\widehat{\mathcal{B}}$ is also generated by the continuous linear forms ϱ on $\widehat{B}$ that can be written as $\varrho(a) = \ell(a(x))$ for $\ell \in S$ and $x \in B$. Using the separability of B it is easy to see that the map $a \mapsto \|a\|$ is $\mathcal{B}^* - \mathcal{B}(\mathbb{R})$ measurable for $a \in B^*$. In the same way, one shows that $b \mapsto \|b\|$ is measurable with respect to $\widehat{\mathcal{B}} - \mathcal{B}(\mathbb{R})$.

Definition 2.7. *By* random continuous linear form on B *we denote any random variable with values in $(B^*, \mathcal{B}^*)$. Analogously,* random continuous linear operators on B *are random variables with values in $(\widehat{B}, \widehat{\mathcal{B}})$.*

To settle issues of measurability, note the following: For any $a \in B^*, x \in B$, random continuous linear form A and random variable X in B, we have that $a(X), A(x)$ and $A(X)$ are again random variables. The latter follows from measurability of the map $(a,x) \mapsto a(x)$ with respect to $\mathcal{B}^* \otimes \mathcal{B} - \mathcal{B}$. This is due to the separability of B, we have

$$\begin{aligned}&\{(a,x) \in B^* \times B : a(x) < r\} \\ &\quad = \bigcup_{k\geq 1} \bigcup_{m \geq 1} \bigcap_{n \geq m} \bigcup_{i \geq 1} \{a \in B^* : a(e_i) < r - 1/k\} \times \{x \in B : \|x - e_i\| < 1/n\},\end{aligned}$$

where $\{e_i : i \geq 1\}$ denotes a dense subset of B. Exactly the same is true for $b(x), A(x), A(X)$ when $b \in \widehat{B}, x \in B$, A denotes a random continuous linear operator and X a random variable in B. The following Lemma is immediate from Lemma 2.4 by conditioning.

Lemma 2.8. *Let $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$. Then, for any random linear continuous form or operator A with $\mathbb{E}[\|A\|^s] < \infty$ independent of X and Y, we have*

$$\zeta_s(A(X), A(Y)) \leq \mathbb{E}[\|A\|^s]\, \zeta_s(X,Y).$$

Note that the assumptions in Lemma 2.4, Lemma 2.5, Corollary 2.6 and Lemma 2.8 are sufficient to guarantee finiteness of all ζ_s-distances in the statements.

We close this section with the simple observation that any relevant property of ζ_s is invariant under isomorphisms. Indeed, if B' denotes a Banach space and $\varphi : B \to B'$ is isomorphic, then

$$\zeta_s(\varphi(X), \varphi(Y)) \leq \|\varphi\|^s \zeta_s(X,Y) \tag{2.6}$$

for $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ by Lemma 2.4.

2.2. Upper bounds for ζ_s

In this section we give upper bounds for ζ_s, mainly for two reasons: First, we address the question of finiteness of the distance and second to infer convergence in ζ_s from other types of convergence. Zolotarev gave many upper (and lower) bounds for ζ_s, some of them being valid only if more structure of B is assumed. The only upper bound we will use subsequently comes from Theorem 2.17 and therefore we include the short proof for the reader's convenience. For any m times continuously differentiable function $f : B \to \mathbb{R}$, we have by Taylor's formula

$$f(x) = \sum_{i=0}^{m-1} \frac{f^{(i)}(0)(x,\dots,x)}{i!} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} f^{(m)}(tx)(x,\dots,x)dt.$$

Hence we let

$$g_f(x) = f(x) - \sum_{i=0}^{m} \frac{f^{(i)}(0)(x,\dots,x)}{i!} = \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} Q(tx)(x,\dots,x)dt \tag{2.7}$$

be the remainder term in the Taylor expansion of order m of f at point 0, where $Q(x) = f^{(m)}(x) - f^{(m)}(0)$. For $f \in \mathcal{F}_s$ we have

$$\|Q(tx)(x,\dots,x)\| \leq t^\alpha \|x\|^s$$

which gives $\mathbb{E}\left[|f(X)|\right] < \infty$ for $f \in \mathcal{F}_s, \mathcal{L}(X) \in \mathcal{M}_s(B)$ and

$$|\mathbb{E}\left[f(X) - f(Y)\right]| \leq \frac{\Gamma(1+\alpha)}{\Gamma(1+s)} B_s(X,Y),$$

for $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ where $B_s(X,Y) := B_s(\mathcal{L}(X), \mathcal{L}(Y))$ and $B_s(\mu,\nu)$ was defined in (2.3). Since the right hand side does no longer depend on f this immediately implies the following result which is Lemma 2 in [Zol76].

Lemma 2.9. *If $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ then*

$$\zeta_s(X,Y) \leq \frac{\Gamma(1+\alpha)}{\Gamma(1+s)} B_s(X,Y).$$

For improved upper bounds we introduce other metrics on the space of probability measures. For $s > 0$ and probability measures μ and ν, let

$$\ell_s(\mu,\nu) := \inf(\mathbb{E}\left[\|X - Y\|^s\right])^{\min(1/s,1)}, \tag{2.8}$$

where the infimum is taken over all random variables (X,Y) on the product space such that $\mathcal{L}(X) = \mu$ and $\mathcal{L}(Y) = \nu$. For random variables X, Y with values in B, we set $\ell_s(X,Y) = \ell_s(\mathcal{L}(X), \mathcal{L}(Y))$. Note that $\ell_s(X,Y)$ is finite if both $\|X\|$ and $\|Y\|$ have finite s-th moment. By the separability of B, it is not hard to see that the infimum in (2.8) is attained if both X and Y have finite absolute moment of order s. The short proof is given in Lemma A.2 in the appendix. The historical background of the metric is diverse, it seems that it first appeared in Gini's work on the Gini coefficient in 1914. There are several other terms for ℓ_s, among them *Wasserstein* and *Kantorovich* metric. We list several of its properties here and include proofs in the appendix. For more detailed information on ℓ_s we refer to [BF81] or the book of Rachev [Rac91]. The following characterization of ℓ_s convergence can be found in both references [Lemma 8.3 in [BF81] and Theorem 8.2.1 in [Rac91]]. A short proof is given in Lemma A.3 in the appendix.

Lemma 2.10. *Let $s > 0$ and $\mathbb{E}\left[\|X_n\|^s\right], \mathbb{E}\left[\|X\|^s\right] < \infty$ for all n. Then $\ell_s(X_n, X) \to 0$ implies $X_n \to X$ in distribution and $\mathbb{E}\left[\|X_n\|^s\right] \to \mathbb{E}\left[\|X\|^s\right]$. The converse is true as well.*

A further quantity that serves in the context of bounding ζ_s from above is defined by

$$\kappa_s(X,Y) = \ell_1(X\|X\|^{s-1}, Y\|Y\|^{s-1})$$

for B valued random variables X, Y. $\kappa_s(X,Y)$ is also called the *difference pseudomoment of order s*. It is easy to see that ℓ_s and κ_s are not equivalent. However, there are close in the sense of the following Lemma, whose proof is contained in Lemma A.4 in the appendix. Note that both ℓ_s and κ_s are indeed metrics on $\mathcal{M}_s(B)$, we refer to [BF81, Lemma 8.3] for a proof.

Lemma 2.11. *For $s > 0$, the topologies induced by ℓ_s, κ_s on $\mathcal{M}_s(B)$ are equal.*

The quantities ℓ_1 and κ_s are constructed as minimal versions of $\mathbb{E}\left[\|g(X) - g(Y)\|\right]$ for a given function $g : B \to B$ whereas ζ_s is defined as the supremum of $|\mathbb{E}\left[f(X) - f(Y)\right]|$ with certain constraints on the function f. Both techniques are related by a classical Theorem of Kantorovich and Rubinstein, see [KR58] (for B compact) and [Dud76] for the general case.

Theorem 2.12. *Let (S,d) be a separable metric space and X,Y be S-valued random variables such that $\mathbb{E}[d(x,X)], \mathbb{E}[d(y,Y)] < \infty$ for some (hence all) $x,y \in S$. Then*

$$\inf \mathbb{E}[d(X,Y)] = \sup\{|\mathbb{E}[f(X)-f(Y)]| : f : S \to \mathbb{R}, |f(x)-f(y)| \le d(x,y)\},$$

where the infimum is taken over all possible realizations of (X,Y) in S^2 with fixed marginals. The supremum on the right hand side is attained for some 1-Lipschitz function $f : S \to \mathbb{R}$.

This leads to an alternative representation of κ_s. For B valued random variables X,Y with $\mathbb{E}[\|X\|^s], \mathbb{E}[\|Y\|^s] < \infty$, we have

$$\kappa_s(X,Y) = \sup\{|\mathbb{E}[f(X)-f(Y)]| : f : B \to \mathbb{R}, |f(x)-f(y)| \le \|x\|x\|^{s-1} - y\|y\|^{s-1}\|\}.$$

For $s = 1$, this immediately follows from Theorem 2.12. For general s, use the observation that $d(x,y) := \|x\|x\|^{s-1} - y\|y\|^{s-1}\|$ is a metric on B that defines the same topology as $\|\cdot\|$. Note that the direction $\ge$ of the latter expression is immediate from the definition of κ_s.

Example 2.13. Let us consider the case $B = \mathbb{R}$. Here, both metrics ℓ_s and κ_s have nice representations. Let U be uniform on the unit interval, X,Y real-valued random variables with $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$ and $s \ge 1$. Then

$$\ell_s(X,Y) = \|F_X^{-1}(U) - F_Y^{-1}(U)\|_s = \left(\int_0^1 |F_X^{-1}(u) - F_Y^{-1}(u)|^s du\right)^{1/s}, \tag{2.9}$$

where $F_X^{-1}(u) = \sup\{x : F_X(x) \le t\}$ with $F_X(t) = \mathbf{P}(X \le t)$ denotes the generalized inverse of the distribution function of X [If $\ell_s(X,Y)$ is infinite, then the same applies to the right hand side of (2.9)]. This was first proved by Dall'Aglio [Dal56], see [Maj78] for a proof in English language. For $s = 1$, a simple geometric argument shows that this representation is equivalent to

$$\ell_1(X,Y) = \int_{-\infty}^{\infty} |F_X(u) - F_Y(u)| du. \tag{2.10}$$

Moreover, if $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$ then $\zeta_1(X,Y)$ coincides with $\ell_1(X,Y)$ and the right hand sides of both (2.10) and (2.9) (with $s = 1$). We provide a short self-contained proof thereof not relying on the Kantorovich-Rubinstein Theorem in Lemma A.5 in the appendix. A simple parameter transformation leads from (2.10) to a similar representation of κ_s. We have

$$\kappa_s(X,Y) = s\left(\int_{-\infty}^{\infty} |u|^{s-1}|F_X(u) - F_Y(u)| du\right). \tag{2.11}$$

for all $s > 0$.

We move on comparing the metrics ℓ_s, κ_s with ζ_s. For $s \le 1$, the definition of $\mathcal{F}_s$ directly gives

$$\zeta_s(X,Y) \le \ell_s(X,Y) \tag{2.12}$$

for $\mathbb{E}[\|X\|^s], \mathbb{E}[\|Y\|^s] < \infty$. Again, by Theorem 2.12, we have equality in (2.12) by the separability of B.

Corollary 2.14. *Let $s \le 1$ and $\mathbb{E}[\|X\|^s], \mathbb{E}[\|Y\|^s] < \infty$. Then $\zeta_s(X,Y) = \ell_s(X,Y)$.*

For applications with $s > 1$ we will use an upper bound for ζ_s in terms of ℓ_s which we state as Theorem 2.17 at the end of this section. Zolotarev himself gave upper bounds in terms of κ_s in [Zol76, Section 4]. Like Lemma 2.9, they are based on a Taylor expansion of functions from class $\mathcal{F}_s$.

Theorem 2.15. *Let $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ and $b_s = \min(\mathbb{E}[\|X\|^s], \mathbb{E}[\|Y\|^s])$. Then, for $s > 0$,*

$$\zeta_s(X,Y) \leq \frac{\Gamma(1+\alpha)}{\Gamma(1+s)}(2m\kappa_s + (2\kappa_s)^\alpha b_s^{1-\alpha}). \tag{2.13}$$

Under the same conditions and for integer values of s, it holds

$$\zeta_s(X,Y) \leq \frac{2\kappa_s}{\Gamma(s)}. \tag{2.14}$$

If B is a Hilbert space both inequalities remain valid if we remove the 2's on the right hand sides. On $\mathcal{M}_s(\nu)$, convergence in κ_s implies convergence in ζ_s.

Remark 2.16. Zolotarev emphasizes that, in general, the condition (2.3) would not be necessary for finiteness of $\zeta_s(\mu,\nu)$. For $s \geq 1$ integer, $\zeta_s(\mu,\nu)$ is well-defined and finite if only (2.4) and $\kappa_s(X,Y) < \infty$ where $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$ hold true. In this case (2.14) remains valid. By Theorem 4 in [Zol77] a similar statement is true for any $s > 0$ if $\nu_s(X,Y) < \infty$ where $\nu_s(X,Y)$ denotes the *absolute pseudomoment* of order s of X and Y. It is defined by

$$\nu_s(X,Y) = \int |u|^s d|\mathbb{P}_X - \mathbb{P}_Y|(u),$$

where $|\mathbb{P}_X - \mathbb{P}_Y|$ is the variation measure of the signed measure $\mathbb{P}_X - \mathbb{P}_Y$. However, $\mathbb{E}[\|X\|^s] = \infty, \mathbb{E}[\|Y\|^s] = \infty$ may imply non-integrability of $f(X)$ for certain functions of $\mathcal{F}_s$. Thus, an appropriate definition of $\zeta_s(\mu,\nu)$ involves a minimization over all possible random variables (X,Y) on the product spaces with marginals μ and ν just as in the definition of ℓ_s. We do not want to deal at all with this case.

We now state and prove the ℓ_s bound for $s > 1$. It is Lemma 5.7 in [DJN08].

Theorem 2.17. *Let $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ with $s > 1$. Then*

$$\zeta_s(X,Y) \leq \left[(\mathbb{E}[\|X\|^s])^{1-1/s} + (\mathbb{E}[\|Y\|^s])^{1-1/s}\right] \ell_s(X,Y).$$

In particular, ℓ_s convergence implies ζ_s convergence in $\mathcal{M}_s(\nu)$.

Proof. Recall the function g_f from (2.7) for $f \in \mathcal{F}_s$

$$g_f(x) = f(x) - f(0) - \sum_{i=1}^m \frac{f^{(i)}(0)(x,\ldots,x)}{i!}.$$

The i-th summand is i-linear in its argument hence it vanishes for $x = 0$ and its i-th derivative is constant. This gives

$$g_f(0) = Dg_f(0) = \cdots = D^m g_f(0) = 0$$

and

$$\|D^m g_f(x) - D^m g_f(y)\| = \|D^m f(x) - D^m f(y)\| \leq \|x-y\|^\alpha.$$

By the mean value theorem, for any $x \in B$ it exists $0 \le \theta \le 1$ with

$$\|D^{m-1} g_f(x)\| = \|D^m g_f(\theta x)(x)\| \le \|x\|^{\alpha+1}$$

and by backward induction

$$\|D^k g_f(x)\| \le \|x\|^{s-k},$$

for all $0 \le k \le m$. Now, for $x, y \in B$, again using the mean value theorem, there exists $0 \le \theta \le 1$ such that

$$\begin{aligned} \|g_f(y) - g_f(x)\| &= \|Dg(x + \theta(y-x))(y-x)\| \le \|(1-\theta)x + \theta y\|^{s-1}\|y-x\| \\ &\le (\|x\|^{s-1} + \|y\|^{s-1})\|y-x\| \end{aligned}$$

using the triangle inequality and $(1-\theta)\|x\| + \theta\|y\| \le \max(\|x\|, \|y\|)$ in the final step. Hölder's inequality now implies

$$\begin{aligned} |\mathbb{E}[f(X) - f(Y)]| &\le \mathbb{E}[\|g(X) - g(Y)\|] \le \mathbb{E}\left[(\|X\|^{s-1} + \|Y\|^{s-1})\|Y - X\|\right] \\ &\le \left[\mathbb{E}[\|X\|^s]^{1-1/s} + \mathbb{E}[\|Y\|^s]^{1-1/s}\right] \mathbb{E}[\|Y - X\|^s]^{1/s}. \end{aligned}$$

Taking the supremum over $f \in \mathcal{F}_s$ and the infimum over realizations of $\mathcal{L}(X)$ and $\mathcal{L}(Y)$ the assertion follows. □

2.3. Lower bounds on ζ_s

Upper bounds are of interest to prove convergence of a sequence of random variables (X_n) to a random variable X in the ζ_s distance. Lower bounds however are of great importance aiming to infer other modes of convergence, in particular weak convergence, from convergence in the Zolotarev distance. In principle, lower bounds can be easily obtained by choosing arbitrary functions f from $\mathcal{F}_s$. Therefore, the richness of $\mathcal{F}_s$ plays a major role. Our main focus in the section lies on the following two problems.

- Does $\zeta_s(\mu, \nu) = 0$ imply $\mu = \nu$?
- Does $\zeta_s(\mu_n, \mu) \to 0$ imply $\mu_n \to \mu$ weakly for $n \to \infty$?

In general, only the first problem has a positive answer. However, we will give considerably weak additional assumptions on (μ_n) and μ to obtain weak convergence for the cases of continuous or càdlàg functions on the unit interval in Section 2.6. A simple, yet useful bound in the case of real-valued random variables comes from by the observation that for all $s > 0$ integer, we have $C_s^{-1} f \in \mathcal{F}_s$ with

$$f(x) = x^s, \qquad C_s = s!$$

This gives

$$|\mathbb{E}[X^s] - \mathbb{E}[Y^s]| \le C_s \zeta_s(X, Y) \tag{2.15}$$

for $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$ and $B = \mathbb{R}$. This result is rather simple, however, based on it as a lower bound and on the upper bound given in Theorem 2.15, Neininger and Rüschendorf [NR02] were able to prove that

$$\zeta_3(Z_n^*, \sigma^{-1} Z) = \Theta\left(\frac{\log n}{n}\right),$$

where, as mentioned in the introduction, $Z_n^* = (\mathbf{Var}\,[Z_n])^{-1/2}(Z_n - \mathbb{E}\,[Z_n])$ is the precise rescaled version of Z_n, the number of key comparisons in the randomized Quicksort algorithm, Z its weak limit given by the solution of (1.7) and $\sigma^2 = \mathbf{Var}\,[Z]$. It is still an open problem to determine the order of convergence in this problem for more common metrics, e.g. the Kolmogorov metric

$$\varrho(X, Y) = \sup_{x \in \mathbb{R}} |\mathbf{P}\,(X \le x) - \mathbf{P}\,(Y \le x)\,|, \tag{2.16}$$

or the Lévy-Prokhorov metric

$$\pi(X, Y) \;=\; \inf\{\varepsilon > 0 : \mathbf{P}\,(X \in C) \le \mathbf{P}\,(Y \in C^\varepsilon) + \varepsilon\, \forall\, C \text{ closed}\}, \tag{2.17}$$

where $C^\varepsilon = \{x \in \mathbb{R} : \|x - y\| < \varepsilon \text{ for some } y \in C\}$. Bounds on the convergence rates in the Quicksort problem in the Kolmogorov metric and the ℓ_p metrics for $p > 1$ have been obtained in [FJ02]. Note that the Lévy-Prokhorov metric has an obvious extension to the case of arbitrary metric spaces.

(2.15) naturally poses the question whether there exists constants $\bar{C}_s$ such that $\bar{C}_s\|x\|^s \in \mathcal{F}_s$. In general, this cannot be the case, since the norm function may not be differentiable, $B = \mathcal{C}[0, 1]$ is one of these cases [see Lemma A.6 in the appendix for a proof]. Nevertheless, it is true in Hilbert spaces. This is easy to show for integer values of s and more involved in the general case. Zolotarev gives a proof containing a slight mistake in the last step in the case of $1 < s \le 2$ [Zol76, Theorem 6]. Based on his arguments, we give a proof covering all cases here.

Lemma 2.18. *Let B be a Hilbert space, $s > 0$. Then, there exists a constant $\bar{C}_s > 0$ such that $\bar{C}_s\|x\|^s \in \mathcal{F}_s$. Thus, for any $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$, we have*

$$|\mathbb{E}\,[\|X\|^s] - \mathbb{E}\,[\|Y\|^s]\,| \le \bar{C}_s^{-1}\zeta_s(X, Y).$$

Proof. Let $\nu(x) = \|x\|$ and $\nu_s(x) = \|x\|^s$. Obviously, $\nu_s(x) \in \mathcal{F}_s$ for $s \le 1$. For $s \in 2\mathbb{N}$, it is easy to see that $\|D^s\nu_s(x)\|$ is constant and equals $s!$. Hence $\bar{C}_s\nu_s(x) \in \mathcal{F}_s$ where $\bar{C}_s = (s!)^{-1}$. For the remaining of the proof fix $s \in \mathbb{R}^+ \backslash 2\mathbb{N}$. Let $x \ne 0$. Then, it is not difficult to see that for every $k > 0$

$$D^k\nu_s(x)[h_1, \ldots, h_k] = \sum_{0 \le \ell \le k, \ell = k \bmod 2} \; \sum_{*} c_\ell \|x\|^{s-k-\ell} \prod_{r=1}^{\ell} < x, h_{i_r} > \prod_{r=1}^{(k-\ell)/2} < h_{j_r}, h_{p_r} >$$

where, for fixed ℓ, the second sum ranges over all disjoint sets $\{i_1, \ldots, i_\ell\}, \{j_1, \ldots, j_{(k-\ell)/2}\}$ and $\{p_1, \ldots, p_{(k-\ell)/2}\}$ with $1 \le i_1 < \ldots < i_\ell \le k$, $1 \le j_1 < \ldots < j_{(k-\ell)/2} \le k$, $j_i < p_i$ for $i = 1, \ldots, (k - \ell)/2$ and $c_\ell = s(s-2)\cdots(s-(k+\ell)+2)$. This representation is used also in [Jam77, Lemma 1]. Thus, for $y \ne 0$, the term $D^m\nu_s(x) - D^m\nu_s(y)$ consists of summands of the form

$$c_\ell \left(\|x\|^{\alpha-\ell} \prod_{r=1}^{\ell} < x, h_{i_r} > - \|y\|^{\alpha-\ell} \prod_{r=1}^{\ell} < y, h_{i_r} > \right) \prod_{r=1}^{(m-\ell)/2} < h_{j_r}, h_{p_r} > \tag{2.18}$$

For $\ell = 0$, the norm of this term is bounded by $c_\ell | \|x\|^\alpha - \|y\|^\alpha | \leq c_\ell \|x-y\|^\alpha$. For $1 \leq \ell \leq m$, we can express (2.18) as

$$\sum_{0\leq t\leq \ell-1} c_\ell \|x\|^{\alpha-\ell} \prod_{r=1}^{t} < x, h_{i_r} >< x-y, h_{i_{t+1}} > \prod_{r=t+2}^{\ell} < y, h_{i_r} > \prod_{r=1}^{(m-\ell)/2} < h_{j_r}, h_{p_r} > \tag{2.19}$$

$$+ c_\ell (\|x\|^{\alpha-\ell} - \|y\|^{\alpha-\ell}) \prod_{r=1}^{\ell} < y, h_{i_r} > \prod_{r=1}^{(m-\ell)/2} < h_{j_r}, h_{p_r} > \tag{2.20}$$

Let $\|y\| \leq \|x\|$. Then, the norm of each of the summands in (2.19) is bounded by

$$c_\ell \|x\|^{\alpha-1} \|x-y\| \leq c_\ell \left(\frac{\|x-y\|}{\|x\|} \right)^{1-\alpha} \|x-y\|^\alpha \leq c_\ell 2^{1-\alpha} \|x-y\|^\alpha.$$

Additionally, the norm of the term (2.20) is bounded by $c_\ell \beta(\ell, \|x\|, \|y\|)$ where $\beta(\ell, x, y) = y^\alpha - y^\ell x^{\alpha-\ell}$ for $x, y \in \mathbb{R}^+$ and $\ell \in \mathbb{N}$. It is easy to see that $\beta(\ell, x, y) \leq (\ell/\alpha - 1)(x^\alpha - y^\alpha)$ for $x \geq y > 0$. Thus,

$$\beta(\ell, \|x\|, \|y\|) \leq (\ell/\alpha - 1)(\|x\|^\alpha - \|y\|^\alpha) \leq (\ell/\alpha - 1)\|x-y\|^\alpha.$$

Overall, for $\|x\| \geq \|y\| > 0$, the term $\|D^m \nu_s(x) - D^m \nu_s(y)\| / \|x-y\|^\alpha$ is bounded uniformly in x, y. Finally, note that $\nu_s(0) = D\nu_s(0) = \cdots = D^{(m)}\nu_s(0) = 0$, so the case that $y = 0$ or / and $x = 0$ can be handled with ease. $\square$

A useful concept in the issues of uniqueness and convergence of distributions is that of *characteristic functions* (or *Fourier transforms*).

Definition 2.19. *Let X be a B valued random variable. The function $\phi_X : L(B, \mathbb{R}) \to \mathbb{C}$ defined by*

$$\phi_X(h) = \mathbb{E}\left[e^{ih(X)} \right]$$

is called characteristic function of X.

Obviously ϕ_X only depends on the distribution of X so the characteristic function is naturally defined for measures, we omit this reformulation here. The theory of characteristic functions in infinite dimensional spaces is considerably more involved than in $\mathbb{R}^d$; nevertheless, our first result resembles the corresponding statement in the multivariate real case and can actually be easily proved relying on the latter. It was first noted by Kolmogorov [Kol35].

Theorem 2.20. *The characteristic function uniquely determines the distribution, i.e. $\phi_X(h) = \phi_Y(h)$ for all $h \in L(B, \mathbb{R})$ implies $\mathcal{L}(X) = \mathcal{L}(Y)$.*

The next Lemma which essentially coincides with Theorem 2 in [Zol77] allows the transition from characteristic functions to functions from class $\mathcal{F}_s$.

Lemma 2.21. *Let $h \in L(B, \mathbb{R})$ and $e(x) = e^{ih(x)}$ for $x \in B$. Then e is smooth. Remembering $s = m + \alpha$, we have $2^{\alpha-1}\|h\|^{-s} \sin(h(x)) \in \mathcal{F}_s$ and $2^{\alpha-1}\|h\|^{-s} \cos(h(x)) \in \mathcal{F}_s$.*

Proof. The smoothness property is trivial since continuous linear forms are smooth and also $x \mapsto e^{ix}$ is. Observe that

$$e^{(m)}(x)(u_1, \ldots, u_m) = i^m e(x) h(u_1) \cdots h(u_m).$$

Using Lipschitz continuity of e^{ix} and $\min(2, x) \leq 2^{1-\alpha} x^\alpha$ for $x > 0$ we obtain

$$\begin{aligned} \|e^{(m)}(x) - e^{(m)}(y)\| &\leq |e(x) - e(y)| \|h\|^m \\ &\leq \|h\|^m \min(2, |h(x) - h(y)|) \leq 2^{1-\alpha} \|h\|^s \|x - y\|^\alpha. \end{aligned}$$

Since $e(x) = \cos(h(x)) + i \sin(h(x))$ the result now follows by linearity of the derivative. □

The Lemma immediately implies

$$|\phi_X(h) - \phi_Y(h)| \leq 2 \cdot 2^{1-\alpha} \|h\|^s \zeta_s(X, Y) \tag{2.21}$$

for all h. In particular we have $\phi_X = \phi_Y$ if $\zeta_s(X, Y) = 0$. Together with Theorem 2.20 this gives the following theorem.

Theorem 2.22. *Let $\mathcal{L}(X), \mathcal{L}(Y) \in \mathcal{M}_s(\nu)$. If $\zeta_s(X, Y) = 0$ then $\mathcal{L}(X) = \mathcal{L}(Y)$, in particular ζ_s is a metric on $\mathcal{M}_s(\nu)$.*

In his papers, Zolotarev defines the ζ_s distances using functions from B to the space of complex numbers. This would allow to work directly with $e(x)$ and save a factor of 2 in (2.21).
We now move on to the question whether convergence in the Zolotarev distance implies weak convergence. As in Lemma 2.18, the smoothness of the norm function $x \mapsto \|x\|$ plays a crucial role. That is one main reason why Hilbert spaces are much easier to handle than general Banach spaces. There, one can only hope for good approximations of the norm by smoother functions. By Theorem 2.12 and Lemma 2.10 ζ_s convergence implies weak convergence for $0 \leq s \leq 1$. A direct proof not relying on the Kantorovich-Rubinstein Theorem is contained in the proof of the classical Portementeau Lemma [Bil68, Theorem 2.1, ii)⇒ iii)] together with Theorem 1.2 there. We will make use of the latter theorem several times, so we state it as Lemma A.1 in the appendix.

Corollary 2.23. *Let $0 \leq s \leq 1$. Then ζ_s convergence implies weak convergence.*

We now move on to the general case. Let $\mathcal{L}(X_n), \mathcal{L}(X) \in \mathcal{M}_s(\nu)$ for all n with $\zeta_s(X_n, X) \to 0$. According to (2.21), we have $\phi_{X_n}(h) \to \phi_X(h)$ for every $h \in L(B, \mathbb{R})$. This immediately connects our question to Lévy's continuity theorem and motivates the following definition.

Definition 2.24. *A Banach space B has property* PL *if the following statements are equivalent:*

i) $X_n \to X$ for $n \to \infty$ in distribution,

ii) $\phi_{X_n}(h) \to \phi_X(h)$ for every $h \in L(B, \mathbb{R})$ as $n \to \infty$.

By definition (ii) follows from (i) so the interesting direction is $ii) \to i)$. It is well-known that any finite dimensional Banach space has property PL, thus ζ_s convergence implies weak convergence in finite dimensional Banach spaces. However, it is easy to see that not every Banach space B has property PL. The following example is taken from [Mou53].

Example 2.25. Let B be an infinite dimensional separable Hilbert space with a basis $(e_i)_{i\in\mathbb{N}}$ of orthonormal vectors. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of B-valued random variables with $\mathcal{L}(X_i) = \delta_{e_i}$. Fix $h \in L(B, \mathbb{R})$. By the Riesz representation theorem there exists $y_h \in B$ such that $h(x) = \langle x, y_h \rangle$ for all $x \in B$. Furthermore if $y_h = \sum_{i\in\mathbb{N}} \alpha_i e_i$ then $\alpha_i \to 0$ since $\|y_h\| = \|h\| < \infty$. Therefore $h(e_n) = \alpha_n \to 0$ which implies $\phi_{X_n}(h) \to 1$. This means $\phi_{X_n}(h) \to \phi_X(h)$ if X has distribution δ_0. But obviously X_n does not converge to X in distribution.

In fact, by results in [Bou73], no normed vector space of infinite dimension has property PL. In their works on the central limit theorem in Hilbert spaces, Giné and León [GL80] proved the following theorem, see also [DJN08, Theorem 5.1].

Theorem 2.26. *Let B be a separable Hilbert space. Then ζ_s convergence implies weak convergence.*

In the general case, note that ζ_s convergence uniquely determines the possible limit in the weak topology. Solely the question of tightness of (X_n) remains open. We summarize the relations between ℓ_s, κ_s and ζ_s by combining the statements of Lemma 2.10, Theorem 2.17 and Lemma 2.18.

Corollary 2.27. *Let $\mathcal{L}(X_n), \mathcal{L}(X) \in \mathcal{M}_s(B)$ for all n. Then, $\kappa_s(X_n, X) \to 0$ if and only if $\ell_s(X_n, X) \to 0$. Furthermore, if $\mathcal{L}(X_n), \mathcal{L}(X) \in \mathcal{M}_s(\nu)$ for all n, then $\ell_s(X_n, X) \to 0$ implies $\zeta_s(X_n, X) \to 0$. $\zeta_s(X_n, X) \to 0$ implies $\ell_s(X_n, X) \to 0$ if and only if $\zeta_s(X_n, X) \to 0$ implies $X_n \to X$ in distribution. If B is a Hilbert space then ζ_s, κ_s and ℓ_s induce the same topology on $\mathcal{M}_s(\nu)$.*

The proofs of Corollary 2.23 and Theorem 2.26 are straightforward in the sense that no other metrics on the space of probability distributions are involved. For the remaining of this section we discuss lower bounds on the Zolotarev metrics in terms of the Lévy-Prokhorov metric as defined in (2.17). Bounds of this type have been proved by Zolotarev [Zol76], Jamukov [Jam77] and Senatov [Sen84] (based on results in [Sen83]). We collect them in the following Theorem. Here, for a subset $\mathfrak{A}$ of the Borel sets in B, let

$$\begin{aligned}\pi(X, Y; \mathfrak{A}) = \inf \{&\varepsilon > 0 : \mathbf{P}(X \in A) \le \mathbf{P}(Y \in A^\varepsilon) + \varepsilon, \\ &\mathbf{P}(Y \in A) \le \mathbf{P}(X \in A^\varepsilon) + \varepsilon \,\forall\, A \in \mathfrak{A}\},\end{aligned}$$

with $A^\varepsilon = \{x \in B : \|x - y\| < \varepsilon \text{ for some } y \in A\}$. Moreover, we abbreviate $\mathfrak{C}$ for the set of convex measurable subsets of B and $\mathfrak{S}$ for the set of closed spheres.

Theorem 2.28. *For all $0 < s \le 1$ we have*

$$\pi^{1+s}(X, Y) \le \zeta_s(X, Y). \tag{2.22}$$

Let $s > 0$ and the norm function $\eta(x) = \|x\|$ be $m + 1$ times differentiable for all $x \neq 0$ such that there exists constants $A_1, \ldots, A_{m+1}$ with

$$\|\eta^{(k)}(x)\| \le A_k \|\eta^{1-k}(x)\| \tag{2.23}$$

for all $k = 1, \ldots, m + 1$. Then there exists a constant $C_s = C_s(\alpha, A_1, \ldots, A_{m+1})$ such that

$$\pi^{1+s}(\|X\|, \|Y\|) \le C_s \zeta_s(X, Y). \tag{2.24}$$

$$\pi^{1+s}(X, Y; \mathfrak{S}) \le C_s \zeta_s(X, Y). \tag{2.25}$$

Assumption (2.23) *is satisfied in any Hilbert space. If B is a Hilbert space then, for any $0 < s \leq 2$,*

$$\pi^{1+s}(X, Y; \mathfrak{C}) \leq C\zeta_s(X, Y) \tag{2.26}$$

for some constant $C > 0$ which can chosen independently of s. Finally, if $B = \mathbb{R}^d$, then

$$\pi^{1+s}(X, Y; \mathfrak{C}) \leq C_s d^{(s-1-\alpha)/2}\zeta_s(X, Y), \qquad s > 1, \tag{2.27}$$

$$\pi^{1+s}(X, Y) \leq C_s d^{(s-\alpha)/2}\zeta_s(X, Y), \qquad s > 0. \tag{2.28}$$

Here C_s denotes a constant which depends on s (but not on d).

The simplest bound (2.22) is mentioned in [Zol84] and [Sen83]. (2.24) is proved in [Zol76], (2.25) can be deduced easily from (2.24) as indicated in [Zol79b]. The final three bounds (2.26), (2.27) and (2.28) are obtained in [Sen84]; weaker versions of (2.28) have been proved in [Jam77]. We only outline the proofs here: Let $\varepsilon < \pi(X, Y)$ and C be a closed set with

$$\varepsilon \leq \mathbf{P}(X \in C) - \mathbf{P}(Y \in C^\varepsilon).$$

It is sufficient to find a real-valued function f (depending on s, ε and C on B) and a constant $c > 0$ only depending on s such that $\|f\| \leq 1$, $f(x) = 1$ for $x \in C$, $f(x) = -1$ for $x \notin C^\varepsilon$ and $c\,\varepsilon^s f(x) \in \mathcal{F}_s$. Then, $\pi^{1+s}(X, Y) \leq 1/(2c)\zeta_s(X, Y)$ follows from the observation that

$$\mathbb{E}[f(X) - f(Y)] \geq 2(\mathbf{P}(X \in B) - \mathbf{P}(Y \in B^\varepsilon)) \geq 2\varepsilon.$$

For $0 < s \leq 1$, the existence of such a function f and constant c which can be chosen to be two is guaranteed by Lemma A.1 in the appendix. This gives the bound (2.22).

2.4. ζ_s in type p Banach spaces

Bentkus and Rachkauskas [BR85] were the first to mention that ζ_s convergence does not always imply weak convergence. Additionally, they claimed that any bound of the form $\pi^\beta \leq c\,\zeta_s$ with $\beta, c > 0, s > 1$ can not be valid for arbitrary Banach spaces (and not even for separable Hilbert spaces). We discuss this by drawing attention to the central limit theorem and start with the real-valued case. Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed real-valued random variables with zero mean and finite variance σ^2. The central limit theorem, short CLT, states that $S_n^* = n^{-1/2}\sum_{i=1}^n X_i$ converges in distribution to a normally distributed random variable N with zero mean and variance σ^2. Additionally, if $\mathbb{E}\left[|X|^3\right]$ is finite, the classical Berry-Essen Theorem gives

$$\varrho(S_n^*, N) = O(n^{-1/2}),$$

where ϱ denotes the Kolmogorov distance defined in (2.16) and the rate is known to be of this order. The bound is actually uniform for all distributions of X_1 with fixed variance and bounded absolute third moment but we do not emphasize this here. In the Lévy-Prokhorov distance, an upper bound of the form

$$\pi(S_n^*, N) = O(n^{-1/2})$$

is known as well [Yur75]. Let us now consider the Zolotarev distance. Assuming $\mathbb{E}[|X_1|^s] < \infty$ for some $s > 2$, the ideal property of ζ_s and the convolution property of the normal distribution directly imply

$$\zeta_s(S_n^*, N) = \zeta_s(X_1, N)n^{1-s/2}. \tag{2.29}$$

Hence, for $s = 3$, we have the same order of convergence as in the Kolmogorov distance. Furthermore, assuming $\mathbb{E}\left[X_1^3\right] = 0$ and $\mathbb{E}\left[|X_1|^s\right] < \infty$ for some $3 < s \leq 4$, the ζ_s rate improves whereas assuming higher moments does not improve the rate in the Kolmogorov distance in general. The crucial observation is: The proof of (2.29) works analogously in *any* Banach space assuming that X_1, N both have mean zero [that is $\mathbb{E}\left[f(X)\right] = \mathbb{E}\left[f(N)\right] = 0$ for all $f \in B^*$], $\mathbb{E}\left[\|X_1\|^s\right] < \infty$ for some $2 < s \leq 3$, N is normally distributed [that is $f(N)$ is normally distributed for all $f \in B^*$] and (2.4) is satisfied for $k = 2$. In fact, the behaviour of the distance between S_n and N in the weak topology may change dramatically when the structure of B is decreased.

Hilbert space case: By the classical result of Varadhan [Var62] the CLT remains valid in Hilbert spaces; that is X_1 satisfies the CLT if X_1 has mean zero [that is $\mathbb{E}\left[< X, v >\right] = 0$ for all $v \in B$] and $\mathbb{E}\left[\|X\|^2\right] < \infty$. However, according to a result by Senatov [Sen81], there exists a symmetric probability distribution μ concentrated on the unit sphere of ℓ^2 (such that $\int \|x\|^k d\mu(x) < \infty$ for all $k > 0$) with the following property: If X_1 has distribution μ then, for any $\varepsilon > 0$,

$$\pi(S_n^*, N) = \Omega(n^{-\varepsilon}),$$

where N is normal with zero mean and the same covariance operator as μ [that is

$$\mathbb{E}\left[< X, v >< X, w >\right] = \mathbb{E}\left[< N, v >< N, w >\right]$$

for all $v, w \in B$]. Note that equality of the covariance operator of X_1 and N implies (2.4) for $k = 2$ by Lemma 2.2. This shows that any bound of the form $\pi^\beta \leq c\,\zeta_s$ for $\beta > 0, c > 0, 2 < s \leq 3$ is false in ℓ^2.

Banach space case: The central limit theorem is considerably more involved in general Banach spaces. As it turns out, square-integrability of the norm of the distribution is neither sufficient nor necessary to imply the CLT. In the context of the CLT, the following properties of a Banach space are of great interest.

Definition 2.29. *A separable Banach space B is of* type $1 \leq p$ *if for a sequence (Y_n) of independent and identically distributed random variables with $\mathbf{P}(Y_1 = 1) = \mathbf{P}(Y_1 = -1) = 1/2$ and every finite sequence $x_1, \ldots, x_k$ from B, there exists a constant $C > 0$ such that*

$$\mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i Y_i\right\|^p\right] \leq C \sum_{i=1}^{k} \|x_i\|^p.$$

Similarly, B is of cotype $1 \leq q \leq \infty$ *if*

$$\sum_{i=1}^{k} \|x_i\|^q \leq C\,\mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i Y_i\right\|^q\right]$$

under the same conditions. For $q = \infty$, this reduces to

$$\sup_{i=1,\ldots,k} \|x_i\| \leq C\,\mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i Y_i\right\|\right].$$

It is easily seen that the definition of type only makes sense for $p \leq 2$ and that B has type $p' < p$ if it has type p. In the same way, the cotype definition is meaningful only for $q \geq 2$ and cotype q implies cotype q' for $q' > q$. However, note that being of type $p' < p$ for all $p' < p$ does not imply B to be of type p. The analogous result holds for the cotype. Every Banach space is of trivial type 1 and of trivial cotype ∞. Type and cotype properties are related to the geometry of the space B, we refer to [LT91] for a comprehensive discussion of this topic. Every Hilbert space is of type and cotype 2 and a fundamental result by Kwapień [Kwa72] states the converse: Every Banach space of type and cotype 2 is isomorphic to a Hilbert space. This generalizes Theorem 2.26 to Banach spaces with this property. The following two Theorems are fundamental with the first dating back to Hoffmann-Jørgensen and Pisier [HJP76] and the second to Jain [Jai77].

Theorem 2.30. *Let X be a zero mean random variable with $\mathbb{E}\left[\|X\|^2\right] < \infty$ and B be of type 2. Then X satisfies the central limit theorem. Conversely, if every zero mean random variable X with $\mathbb{E}\left[\|X\|^2\right] < \infty$ satisfies the central limit theorem, then B is of type 2.*

In general, a zero mean random variable X with $\mathbb{E}\left[\|X\|^2\right] < \infty$ (or at least well defined covariance operator $\mathbb{E}\left[f(X)g(X)\right]$ for all $f, g \in B^*$) may not satisfy the central limit theorem due to two reasons: First, the corresponding normal distribution with the covariance operator given by X does not exist on B. Second, and more important for us, X might be *pregaussian*, that is of mean zero such that the corresponding normal distribution exists [i.e. N has zero mean and $\mathbb{E}\left[f(N)g(N)\right] = \mathbb{E}\left[f(X)g(X)\right]$ for all $f, g \in B^*$], but the sequence (S_n^*) fails to be tight. Note that the sequence $\mathcal{L}(S_n^*)_{n\geq 0}$ can have at most one accumulation point in the weak topology which is necessarily normal. The following result is the analogue to Theorem 2.30 in the case where the first problem is ruled out.

Theorem 2.31. *Let X be pregaussian and B be of cotype 2. Then X satisfies the central limit theorem. Conversely, if every pregaussian random variable satisfies the central limit theorem, then B is of cotype 2. Additionally, $\mathbb{E}\left[\|X\|^2\right] < \infty$ for any pregaussian random variable X in a space of cotype 2.*

In general, nothing more can be said: First, for any $q > 2$ there exists a Banach space of type 2 and cotype q where one finds pregaussian random variables not satisfying the CLT. Second, for any $p < 2$ there exists a Banach space of type p and cotype 2 in which bounded random variables are not necessarily pregaussian. Furthermore Ledoux [Led84] shows that there exists a Banach space of type $2 - \varepsilon$ and cotype $2 + \varepsilon$ for any $\varepsilon > 0$ in which one finds bounded pregaussian random variables not satisfying the CLT. Relating these results to the Zolotarev metrics, we can state the following corollary.

Corollary 2.32. *Let B be of cotype $q > 2$ such that there exists a zero mean pregaussian random variable X with*

- $\mathbb{E}\left[\|X\|^s\right] < \infty$ *for some* $2 < s \leq 3$,
- X, N *satisfy* (2.4) *for* $k = 2$ *where* N *is the corresponding normal distribution,*
- X *does not satisfy the CLT.*

Then $\zeta_s(S_n^*, N) \to 0$ but S_n^* does not converge to N in distribution.

We present the following example in $\mathcal{C}[0,1]$ which goes back to Strassen and Dudley [SD69]. Note that $\mathcal{C}[0,1]$ is of trivial type 1 and cotype ∞.

Example 2.33. For $n \geq 1, j = 0, 1, \ldots, n^2 - 1$, let

$$g_{nj}(x) = \begin{cases} 0 & \text{if } x \in \{0,1\}, \\ 1 & \text{if } 6i+1 \leq 6^n(n!)^2 x \leq 6i+2, \\ -1 & \text{if } 6i+4 \leq 6^n(n!)^2 x \leq 6i+5, \end{cases}$$

where $i = j + rn^2, r = 0, 1 \ldots, 6^{n-1}((n-1)!)^2 - 1$ and linear on all remaining intervals on $[0,1]$ where it is not defined otherwise. Let (p_n) be a distribution on the natural numbers with $p_n \sim n^{-5/4}$. Now, define the distribution of random variable X_1 by $\mathbf{P}(X_1 = g_{nj}) = \mathbf{P}(X_1 = -g_{nj}) = p_n/(2n^2)$ for all $n \geq 1$ and $j = 0, 1, \ldots, n^2 - 1$. Dudley and Strassen show that X_1 is pregaussian but (S_n^*) is not tight. This, together with the obvious boundedness of X_1 and the fact that (2.4) is automatically satisfied in $\mathcal{C}[0,1]$, see Section 2.6, shows that $\zeta_s(S_n, N) \to 0$ whereas S_n^* does not converge in distribution.

Note that, if ζ_s convergence implies weak convergence in B, the same is true for any closed subspace of B and any space B' isomorphic to B as observed in (2.6). Hence, by the Theorem of Banach-Mazur, if a sequence of probability measures (μ_n) converging in ζ_s but not converging weakly can be found in some Banach space B, then it can also be found in $\mathcal{C}[0,1]$.

2.5. The non-separable case and $\mathcal{D}[0,1]$

In this section we only treat cases of Banach spaces which are non-separable with respect to their norm. The application we have in mind is the space of càdlàg functions. The concept of Borel measurability is unsuitable for functions mapping from an underlying probability space $(\Omega, \mathcal{A}, \mathbb{P})$ to $(B, \|\cdot\|)$. We can give several reasons here: First, a classical result (using the continuum hypothesis) by Marczewiski and Sikorski [MS48] shows that any random variable with values in a normed space equipped with Borel-σ-algebra has to be concentrated on a separable subset. Second, it is a non-trivial result involving the axiom of choice that the sum of two Borel-measurable functions in metric spaces is again Borel-measurable [For a counterexample in a non-metrizable space see [Pac05]]. Finally, the random function X with values in $\mathcal{D}[0,1]$ defined by

$$X_t = 1_{\{t \geq U\}}, \qquad t \in [0,1],$$

where U has uniform distribution on the unit interval is not Borel-measurable in $(\mathcal{D}[0,1], \|\cdot\|)$, see [Bil68, Section 18]. One can find a certain number of alternative approaches to this problem in the literature. Concerning $\mathcal{D}[0,1]$, the uniform topology is still appropriate if we assume the points of discontinuities to attain values only in a countable set. In this case, it is sufficient to consider a separable subset of $(\mathcal{D}[0,1], \|\cdot\|)$. A more general approach is given by Dudley [Dud66, Dud67] also with applications to the case $\mathcal{D}[0,1]$. He works with the σ-algebra $\mathcal{B}_0$ generated by the set of open spheres in $(B, \|\cdot\|)$. Obviously $\mathcal{B}_0 \subseteq \mathcal{B}$, the inclusion can be strict as we will see in the case of $\mathcal{D}[0,1]$ later. However, $\mathcal{B}_0$ may also coincide with $\mathcal{B}$ in non-separable case as noted in [Tal78]. For a further approach to weak convergence in non-separable metric spaces see also [Pol79].

The general setting: Subsequently, we will equip B with a σ-algebra $\mathcal{B}_*$ that is smaller than the Borel-σ-algebra generated by the norm which we will always denote by $\mathcal{B}$. In other words $\mathcal{B}_* \subseteq \mathcal{B}$. We have to impose the following restrictions on the richness of $\mathcal{B}_*$:

A1. For any $c \in \mathbb{R}$ and $y \in B$ the functions $x \mapsto cx$ and $x \mapsto x + y$ from B to B are $\mathcal{B}_* - \mathcal{B}_*$ measurable. Moreover, the function $(x, y) \mapsto x + y$ from $B \times B$ to B is $\mathcal{B}_* \otimes \mathcal{B}_* - \mathcal{B}_*$ measurable.

A2. The norm function $x \mapsto \|x\|$ is measurable with respect to $\mathcal{B}_*$.

We make the following general abbreviation: In any definition, lemma, theorem and corollary in the Sections 2.1, 2.2 and 2.3, assume the following:

R1. For any norm-continuous linear operator $f : B \to B'$, additionally suppose that it is $\mathcal{B}_* - \mathcal{B}_*$ measurable if B' equals B and $\mathcal{B}_* - \mathcal{B}(\mathbb{R}^d)$ measurable if $B' = \mathbb{R}^d$ for some $d \geq 1$.

R2. For any norm-continuous multilinear from on B, additionally suppose that this function is measurable with respect to the product-σ-algebra $\mathcal{B}_*^{\otimes k}$.

R3. The class of functions $\mathcal{F}_s$ used to define ζ_s in (2.2) is replaced by the subset of functions from $\mathcal{F}_s$ that are additionally measurable with respect to $\mathcal{B}_*$. We denote this set by $\mathcal{F}_s^*$.

To illustrate our approach, note that

- $\mathcal{M}(B)$ is the set of all probability distributions on $(B, \mathcal{B}_*)$,
- $\mathcal{M}_s(B)$ is the subset of all distributions μ from $\mathcal{M}(B)$ with $\int \|x\|^s d\mu(x) < \infty$ (remember that $x \mapsto \|x\|$ is measurable by **A2**),
- for $\nu \in \mathcal{M}_s(B)$, $\mathcal{M}_s(\nu)$ is the set of all $\mu \in \mathcal{M}_s(B)$ such that (2.4) is satisfied for norm-continuous multilinear functions that are measurable with respect to $\mathcal{B}_*^{\otimes k}$.

The definition of Zolotarev distance remains as in the separable case under consideration of **R3**. The quantities ℓ_s and κ_s are defined as in the separable case.

Using these abbreviations, the following results remain valid in the present case: Lemma 2.3, Lemma 2.4 if $B' = B$ or $B' = \mathbb{R}^d$ for some $d \geq 1$, Lemma 2.5, Corollary 2.6, Lemma 2.9, the direction $\Rightarrow$ in Lemma 2.10, inequality (2.12), Theorems 2.15 and 2.17, Lemma 2.18 in the Hilbert space case and Lemma 2.21 together with inequality (2.21) under the constraint that h is additionally $\mathcal{B}_* - \mathcal{B}(\mathbb{R})$ measurable.
The converse direction in Lemma 2.10 relies on separability; however, it remains valid if we only assume the limit X to be concentrated on a separable subset of B.

Assuming that $\mathcal{B}_*$ is generated by a separable topology that is induced by a metric d, we restrict the set of norm-continuous linear forms or operators in Definition 2.7 to those linear forms that are additionally continuous with respect to d and those operators that are continuous as endomorphisms on (B, d). On these smaller state spaces the corresponding σ-algebras are generated by the same set of functions as in the separable case and additionally also by the norm function. Then Lemma 2.8 remains valid.

In general, the characteristic function may not uniquely determine the law of a random variable, whence we do not know whether $\zeta_s(\mu, \nu) = 0$ implies $\mu = \nu$. There is one important exception: The proof of Theorem 2.20 only relies on the fact that $\mathcal{B}$ is generated by the continuous linear forms; whence, if $\mathcal{B}_*$ is chosen to be generated by this class of functions, Theorem 2.20 remains valid. In this case ζ_s is a metric on $\mathcal{M}_s(\nu)$. A discussion of weak convergence requires that $\mathcal{B}_*$ is generated by a topology. Then, Corollary 2.23 remains valid, if this topology is induced by a metric d such that $d(x, y) \leq C\|x - y\|$ for some $C > 0$ and all $x, y \in B$.

The case $\mathcal{D}[0,1]$**:** The application we have in mind is the space of càdlàg functions. On $\mathcal{D}[0,1]$, there is a well-known topology T_{sk} introduced by Skorokhod [Sko56] which is induced by the so-called Skorokhod metric d_{sk}. All relevant properties of the metric [also its precise definition that is not of interest here] can be found in [Bil68, Section 3]; $(\mathcal{D}[0,1], d_{sk})$ is a complete, separable, i.e. Polish space. Convergence $d_{sk}(x_n, x) \to 0$ is equivalent to the existence of a sequence of monotonically increasing bijections (λ_n) on the unit interval such that $\lambda_n(t) \to t$ and $x_n(\lambda_n(t)) \to x(t)$ both uniformly on $[0,1]$. Thus, we choose $\mathcal{B}_*$ to be the σ-algebra generated by d_{sk} and denote it by $\mathcal{B}_{sk}$. It is well known that $\mathcal{B}_{sk}$ is generated by the finite dimensional projections which shows that the norm function is measurable with respect to $\mathcal{B}_{sk}$. Note that these properties also imply that $\mathcal{B}_{sk}$ coincides with $\mathcal{B}_0$, where $\mathcal{B}_0$ was introduced as the σ-algebra generated by the open spheres in the uniform metric. This proves the conditions **A1** and **A2** to be satisfied. According to Theorem 2 in [Pes95], any norm-continuous linear form on $\mathcal{D}[0,1]$ is measurable with respect to the Skorokhod topology. Moreover, these observations immediately imply that any norm-continuous linear function from $\mathcal{D}[0,1]$ to $\mathcal{D}[0,1]$ is $\mathcal{B}_{sk} - \mathcal{B}_{sk}$ measurable. Hence, the restriction **R1** is negligible. Moreover, according to the results in [JK], any norm-continuous k-linear form is measurable with respect to $\mathcal{B}_{sk}^{\otimes k}$. Thus, restriction **R2** is negligible as well. We do not know whether the classes $\mathcal{F}_s^*$ and $\mathcal{F}_s$ coincide, i.e. if every function from $\mathcal{F}_s$ is measurable with respect to $\mathcal{B}_{sk}$. However, it will turn out that this issue is not problematic. Lemma 2.8 is valid in the càdlàg case, where one should keep in mind that we only allow random norm-continuous linear forms A that are continuous with respect to d_{sk} (or norm-continuous linear operators which are continuous as endomorphisms on $(\mathcal{D}[0,1], d_{sk})$ to $(\mathcal{D}[0,1], d_{sk})$) such that $\|A\|$ is a real-valued random variable. By our remarks on characteristic functions in the previous section, Theorems 2.20 and 2.22 remain valid on $(\mathcal{D}[0,1], d_{sk})$. Moreover, ζ_s convergence implies weak convergence for $0 < s \leq 1$ since $d_{sk}(x, y) \leq \|x - y\|$.

2.6. The Zolotarev distance on $(\mathcal{C}[0,1], \|\cdot\|)$ and $(\mathcal{D}[0,1], d_{sk})$

In the following we consider the separable Banach space $B = (\mathcal{C}[0,1], \|\cdot\|)$ and the Polish space $B = (\mathcal{D}[0,1], d_{sk})$ with the supremum norm $\|\cdot\|$ resp. the Skorokhod metric d_{sk}. First note, for random variables X, Y in $(\mathcal{C}[0,1], \|\cdot\|)$ with $\zeta_s(X, Y) < \infty$ we have

$$\zeta_s((X(t_1), \dots, X(t_k)), (Y(t_1), \dots, Y(t_k)) \leq k^{s/2} \zeta_s(X, Y) \tag{2.30}$$

for all $0 \leq t_1 \leq \ldots \leq t_k \leq 1$. This follows from Lemma 2.4 using the continuous and linear function $g : \mathcal{C}[0,1] \to \mathbb{R}^k, g(f) = (f(t_1), \dots, f(t_k))$ and observing that $\|g\| = \sqrt{k}$ [The more natural bound $\zeta_s((X(t_1), \dots, X(t_k)), (Y(t_1), \dots, Y(t_k)) \leq \zeta_s(X, Y)$ is obtained if $\mathbb{R}^k$ is endowed with the $\max$-norm. For the purpose of this thesis, this improvement is not essential] . Thus, we obtain for random variables X_n, X in $(\mathcal{C}[0,1], \|\cdot\|)$, $n \geq 1$, the implication

$$\zeta_s(X_n, X) \to 0 \quad \Rightarrow \quad X_n \xrightarrow{\text{fdd}} X. \tag{2.31}$$

Here, $\xrightarrow{\text{fdd}}$ denotes weak convergence of all finite dimensional marginals of the processes. Additionally, if Z is a random variable in $[0,1]$, independent of (X_n) and X then Lemma 2.8 implies

$$\zeta_s(X_n(Z), X(Z)) \leq \mathbb{E}\left[Z^s\right] \zeta_s(X_n, X).$$

Finally, if $(X_n), X$ are real-valued random variables, then $\zeta_s(X_n, X) \to 0$ also implies convergence of absolute moments of order up to s by Lemma 2.18. In the càdlàg case, only inequality (2.30) and the implication (2.31) remain valid, the additional statements are based on the continuous linear form $A(f) = f(t)$ for $t \in [0,1]$ which is not continuous with respect to the Skorokhod topology. However, ζ_s convergence implies convergence of the characteristic function of $X_n(t)$ uniformly in t, hence we also have distributional convergence of $X_n(Z)$. The same argument works for the moments of $X_n(Z)$. We summarize these properties in the following proposition.

Proposition 2.34. *For random variables X_n, X in $(\mathcal{C}[0,1], \|\cdot\|)$ or $(\mathcal{D}[0,1], d_{sk})$, $n \geq 1$, with $\zeta_s(X_n, X) \to 0$ for $n \to \infty$ we have*

$$X_n \xrightarrow{\text{fdd}} X.$$

In particular, $\mathcal{L}(X)$ is the only accumulation point of $(\mathcal{L}(X_n))_{n\geq 1}$ in the weak topology. For all $t \in [0,1]$ we have

$$X_n(t) \xrightarrow{d} X(t), \quad \mathbb{E}\left[|X_n(t)|^s\right] \to \mathbb{E}\left[|X(t)|^s\right].$$

For any random variable Z in $[0,1]$ independent of (X_n) and X we have

$$X_n(Z) \xrightarrow{d} X(Z), \quad \mathbb{E}\left[|X_n(Z)|^s\right] \to \mathbb{E}\left[|X(Z)|^s\right].$$

To conclude from convergence in ζ_s to weak convergence on $B = (\mathcal{C}[0,1], \|\cdot\|)$ further assumptions are needed. In view of Proposition 2.34 a tightness criterion is required. Let, for $r > 0$,

$$\begin{aligned}\mathcal{C}_r[0,1] := \{f \in \mathcal{C}[0,1] \mid \exists\, 0 = t_1 < t_2 < \cdots < t_\ell = 1 \;\forall\, i = 1, \ldots, \ell :\\ |t_i - t_{i-1}| \geq r, f|_{[t_{i-1}, t_i]} \text{ is linear}\}\end{aligned} \tag{2.32}$$

denote the set of all continuous functions for which there is a decomposition of $[0,1]$ into intervals of length at least r_n such that the function is piecewise linear on those intervals. Analogously, we define

$$\begin{aligned}\mathcal{D}_r[0,1] := \{f \in \mathcal{D}[0,1] \mid \exists\, 0 = t_1 < t_2 < \cdots < t_\ell = 1 \;\forall\, i = 1, \ldots, \ell :\\ |t_i - t_{i-1}| \geq r, f|_{[t_{i-1}, t_i)} \text{ is constant, continuous in } 1\}.\end{aligned} \tag{2.33}$$

Note that for $r > 0$, the set $\mathcal{C}_r[0,1]$ is Borel-measurable in $\mathcal{C}[0,1]$ and $\mathcal{D}_r[0,1]$ is measurable in B_{sk}.

Theorem 2.35. *Let $(X_n)_{n\geq 0}, X$ be $\mathcal{C}[0,1]$ valued random variables and $0 < s \leq 3$. Suppose $X_n \in \mathcal{C}_{r_n}[0,1]$ for all n and*

$$\zeta_s(X_n, X) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right). \tag{2.34}$$

Then $X_n \to X$ in distribution. The assertion remains valid if $\mathcal{C}[0,1], \mathcal{C}_{r_n}[0,1]$ are replaced by $\mathcal{D}[0,1]$, $\mathcal{D}_{r_n}[0,1]$ and X has continuous sample paths.

As discussed in Section 2.4, ζ_s convergence does not imply weak convergence in the spaces $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ without any further assumption such as (2.34). In the counterexample 2.33 presented at the end of Section 2.4, the sequence S_n^* converges to a Gaussian limit with respect to ζ_s for $2 < s \leq 3$ and is piecewise constant but the sequence r_n can only be chosen of the order (roughly) $(cn)^{-2n}$ for some $c > 0$. Thus, (2.34) is not satisfied.
In applications such as the proof of Donsker's theorem or the application to the partial match retrieval problem presented in Chapters 5 and 4 the rate of convergence will typically be of polynomial order which is fairly sufficient. We postpone the proof of the theorem to the end of the section and state two variants, where the first one, Corollary 2.36, contains a slight relaxation of the assumptions that is useful in applications. The second one will be needed in the case $s > 2$, see Subsection 4.

Corollary 2.36. *Let $(X_n)_{n\geq 0}, X$ be $\mathcal{C}[0,1]$ valued random variables and $0 < s \leq 3$. Suppose $X_n = Y_n + h_n$ with (Y_n) being a sequence $\mathcal{C}[0,1]$ valued random variables and $h_n \in \mathcal{C}[0,1]$ for all n, such that $\|h_n - h\| \to 0$ for a continuous function h and*

$$\mathbf{P}\left(Y_n \notin \mathcal{C}_{r_n}[0,1]\right) \to 0.$$

If

$$\zeta_s(X_n, X) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right),$$

then we have

$$X_n \xrightarrow{d} X.$$

The statement remains true if $\mathcal{C}[0,1]$ and $\mathcal{C}_{r_n}[0,1]$ are replaced by $\mathcal{D}[0,1]$ and $\mathcal{D}_{r_n}[0,1]$ endowed with Skorokhod topology respectively, X has continuous sample paths and h remains continuous.

Corollary 2.37. *Let $(X_n)_{n\geq 0}, (Y_n)_{n\geq 0}, X$ be $\mathcal{C}[0,1]$ valued random variables and $0 < s \leq 3$. Suppose $X_n \in \mathcal{C}_{r_n}[0,1]$ for all n and $Y_n \to X$ in distribution. If*

$$\zeta_s(X_n, Y_n) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right),$$

then

$$X_n \to X$$

in distribution. The statement remains true if $\mathcal{C}[0,1]$ and $\mathcal{C}_{r_n}[0,1]$ are replaced by $\mathcal{D}[0,1]$ and $\mathcal{D}_{r_n}[0,1]$ endowed with Skorokhod topology respectively and X has continuous sample paths.

In $\mathcal{C}[0,1]$ (or $\mathcal{D}[0,1]$ if the limit X has continuous paths), convergence in distribution implies distributional convergence of the supremum norm $\|X_n\|$ by the continuous mapping theorem. In applications, one is also interested in convergence of moments of the supremum. For random variables X in $\mathcal{C}[0,1]$ or $\mathcal{D}[0,1]$, we denote by

$$\|X\|_s := \left(\mathbb{E}\left[\|X\|^s\right]\right)^{(1/s)\wedge 1}$$

the L_s-norm of the supremum norm. For technical reasons, we have to restrict ourselves to integer $s \in \{1,2,3\}$ in the following theorem. Note that we then have $m = s - 1$ and $\alpha = 1$.

Theorem 2.38. *Let $(X_n)_{n\geq 0}, X$ be $\mathcal{C}[0,1]$ valued random variables and $s \in \{1,2,3\}$ with $\|X_n\|_s, \|X\|_s < \infty$ for all n. Suppose one of the following assumptions is satisfied:*

i) $X_n \in \mathcal{C}_{r_n}[0,1]$ *for all* n *and*

$$\zeta_s(X_n, X) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right).$$

ii) $X_n = Y_n + h_n$ *with* $Y_n, h_n \in \mathcal{C}[0,1]$ *for all* n*,* $\|h_n - h\| \to 0$ *for a continuous function* h*,*

$$\mathbb{E}\left[\|X_n\|^s \mathbf{1}_{\{Y_n \notin \mathcal{C}_{r_n}[0,1]\}}\right] \to 0$$

and

$$\zeta_s(X_n, X) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right).$$

iii) $(Y_n)_{n\geq 0}$ *is a sequence of* $\mathcal{C}[0,1]$ *valued random variables with* $|Y_n| \leq Z$ *almost surely for a* $\mathcal{C}[0,1]$ *valued random variable* Z *with* $\|Z\|_s < \infty$*,* $X_n \in \mathcal{C}_{r_n}[0,1]$ *for all* n *and*

$$\zeta_s(X_n, Y_n) = o\left(\log^{-m}\left(\frac{1}{r_n}\right)\right).$$

Then $(\|X_n\|^s)$ *is uniformly integrable. All statements remain true if* $\mathcal{C}[0,1], \mathcal{C}_{r_n}[0,1]$ *are replaced by* $\mathcal{D}[0,1], \mathcal{D}_{r_n}[0,1]$ *endowed with Skorokhod topology and* h *in item* $ii)$ *is continuous.*

In applications of the contraction method one shows distributional convergence and convergence of the s-th absolute moments with the help of the previous results. Convergence of higher moments is a direct consequence of these considerations under mild additional assumptions.

Proposition 2.39. *Let* (X_n) *be a sequence of* B*-valued random variables satisfying recursion* (1.2) *where* B *is a separable Banach space or* $\mathcal{D}[0,1]$ *endowed with the Skorokhod topology. Suppose that, with* $k \geq 1$ *integer,*

i) $\mathbb{E}\left[\|X_n\|^k\right] < \infty$ *for all* n *and* $\sup_n \mathbb{E}\left[\|b^{(n)}\|^k\right] < \infty$*,*

ii) $\mathbb{E}\left[\|A_r^{(n)}\|^k \mathbf{1}_{\{I_r^{(n)}=n\}}\right] \to 0$ *for* $r = 1, \ldots, K$*,*

iii) we have

$$\limsup_{n\to\infty} \sum_{i=1}^{K} \mathbb{E}\left[\|A_r^{(n)}\|^k\right] < 1,$$

iv) for any $k' < k$ *(or only* $k' = k-1$*), we have* $\sup_n \mathbb{E}\left[\|X_n\|^{k'}\right] < \infty$*.*

Then $\mathbb{E}\left[\|X_n\|^k\right]$ *is bounded in* n*. In particular, if* $X_n \to X$ *in distribution, items* $i) - iii)$ *are satisfied for all* $k > p$ *and* $\sup_n \mathbb{E}\left[\|X_n\|^p\right] < \infty$*, then* $\|X\|$ *has finite moments or arbitrary order and* $\mathbb{E}\left[\|X_n\|^\kappa\right] \to \mathbb{E}\left[\|X\|^\kappa\right]$ *for all* $\kappa > 0$*.*

Proof. The recursion implies that, stochastically,

$$\|X_n\| \leq \sum_{r=1}^{K} \left\|A_r^{(n)} X_{I_r^{(n)}}^{(r)}\right\| + \|b^{(n)}\|.$$

For simplicity, we assume $I_r^{(n)} < n$ for all r. Then, by the assumptions, conditioning on the coefficients and Hölder's inequality, there exists a constant $C > 0$ (independent of n) with

$$\mathbb{E}\left[\|X_n\|^k\right] \leq \sum_{r=1}^{K} \mathbb{E}\left[\left\|A_r^{(n)} X_{I_r^{(n)}}^{(r)}\right\|^k\right] + C.$$

Choose $0 < q < 1, n_1 \in \mathbb{N}$ and $\bar{C} > 0$ such that $\mathbb{E}\left[\|X_i\|^k\right] \leq \bar{C}$ for all $i < n_1$ and $\sum_{i=1}^{K} \mathbb{E}\left[\|A_r^{(n)}\|^k\right] \leq 1 - q$ for all $n \geq n_1$. This implies

$$\mathbb{E}\left[\|X_{n_1}\|^k\right] \leq (1-q)\bar{C} + C$$

Further increasing $\bar{C}$ yields $\mathbb{E}\left[\|X_{n_1}\|^k\right] \leq \bar{C}$, hence the desired uniform boundedness by induction. □

Completeness: It is of interest whether the metric space $(\mathcal{M}_s(\nu), \zeta_s)$ is complete. This is true for $0 < s \leq 1$. Also, in the case that B is a separable Hilbert space, this holds true, see Theorem 5.1 in [DJN08]. Nevertheless, the problem remains open in the general case, in particular in the cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ with $s > 1$. We can only state the following proposition.

Proposition 2.40. *Let $(\mu_n)_{n\geq 0}$ be a sequence of probability measures on $\mathcal{C}[0,1]$ or $\mathcal{D}[0,1]$ that is a Cauchy sequence with respect to the ζ_s metric for some $s > 0$. Then there exists a probability measure μ on $\mathbb{R}^{[0,1]}$ such that*

$$\mu_n \xrightarrow{\text{fdd}} \mu. \tag{2.35}$$

Proof. According to (2.30), $(X_n(t_1), \dots, X_n(t_k))_{n\geq 0}$ is a Cauchy sequence and hence it exists a random variable $Y_{t_1,\dots,t_k}$ in $\mathbb{R}^k$ with

$$(X_n(t_1), \dots, X_n(t_k)) \xrightarrow{d} Y_{t_1,\dots,t_k}.$$

The set of distributions of $Y_{t_1,\dots,t_k}$ for $0 \leq t_1 < \dots < t_k \leq 1$ and $k \in \mathbb{N}$ is consistent so there exists a process X on the product space $\mathbb{R}^{[0,1]}$ satisfying (2.35). Note that condition (2.4) would be satisfied for μ_n and a version of μ with continuous paths and finite absolute moment of order s. □

2.7. Proof of the main results of Section 2.6

We now come to the proofs of Theorem 2.35, its two corollaries and Theorem 2.38. The first essentially coincides with Theorem 2 in [Bar90], see also [BJ09], and we present a version of the proof given there so that we can deduce the variants and implications given in our other statements. The basic tool is Corollary 2 in Section 2 of [Bil68]. We state it here as a Lemma.

Lemma 2.41. *Let $(\mu_n)_{n\geq 0}, \mu$ be probability measures on a separable metric space (S, d). For $r > 0, x \in S$ let $B_r(x) = \{y \in S : d(x,y) < r\}$. If for any $x_1, \dots, x_k \in S, \gamma_1, \dots, \gamma_k > 0$ with $\mu(\partial B_{\gamma_i}(x_i)) = 0$ for $i = 1, \dots, k$ it holds*

$$\mu_n\left(\bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \to \mu\left(\bigcap_{i\in I} B_{\gamma_i}(x_i)\right),$$

where $I = \{1, \dots, k\}$, then $\mu_n \to \mu$ weakly.

A main difficulty in deducing weak convergence from convergence in ζ_s compared to the Hilbert space case is the non-differentiability of the norm function $x \mapsto \|x\|$. We will instead use the smoother L_p-norm which approximates the supremum norm in the sense that

$$L_p(x) \to \|x\|, \tag{2.36}$$

for any fixed $x \in \mathcal{C}[0,1]$ as $p \to \infty$. For the remaining part of this section, p, for fixed values or tending to infinity, is always understood as an even integer with $p \geq 4$.

Lemma 2.42. *For $x, y \in \mathcal{C}[0,1]$ let*

$$L_p(x) = \left(\int_0^1 [x(t)]^p dt\right)^{1/p}, \qquad \psi_{p,y}(x) = L_p\left((1+[x-y]^2)^{1/2}\right).$$

Then L_p is smooth on $\mathcal{C}[0,1]\backslash\{\mathbf{0}\}$ where $\mathbf{0}$ is the zero-function and $\psi_{p,y}$ is smooth on $\mathcal{C}[0,1]$ for all $y \in \mathcal{C}[0,1]$. Furthermore for $k \in \{1,2,3\}$, we have

$$\|D^k L_p(x)\| = O(p^{k-1} L_p^{1-k}(x)),$$

uniformly for p and $x \in \mathcal{C}[0,1]\backslash\{\mathbf{0}\}$. Moreover, again for $k \in \{1,2,3\}$,

$$\|D^k \psi_{p,y}(x)\| = O\left(p^{k-1}\right) \tag{2.37}$$

uniformly for p and $x, y \in \mathcal{C}[0,1]$. All assertions remain valid when $\mathcal{C}[0,1]$ is replaced by $\mathcal{D}[0,1]$, moreover both functions L_p and $\psi_{p,y}$ are continuous with respect to the Skorokhod metric for all p and $y \in \mathcal{D}[0,1]$.

Proof. The smoothness properties are obvious. Differentiating L_p by the chain rule yields

$$DL_p(x)[h] = \left(\int_0^1 [x(t)]^p dt\right)^{1/p-1} \int_0^1 [x(t)]^{p-1} h(t) dt.$$

For $h \in \mathcal{C}[0,1]$ with $\|h\| \leq 1$ by Jensen's inequality and $L_p(h) \leq \|h\|$ we obtain that the right hand side of the latter display is uniformly bounded by 1. The bounds on the norms of the higher order derivatives follow along the same lines. Using the same ideas, it is easy to see that

$$\|D^k \psi_{p,y}(x)\| = O\left(\sum_{j=1}^{k} p^{j-1} L_p^{1-j}(\omega_y(x))\right),$$

uniformly in p and $x, y \in \mathcal{C}[0,1]$ where $\omega_y(x) = (1+|x-y|^2)^{1/2}$. This gives (2.37). □

We stress that the convergence in (2.36) only holds pointwise, it is easy to construct a sequence of continuous functions $(x_p)_{p\geq 0}$ such that $L_p(x_p) \to 0$ and $\|x_p\| \to \infty$ for $p \to \infty$. Aside from the obvious bound $L_p(x) \leq \|x\|$ we will need the following simple Lemma which contains sort of a converse of this inequality.

Lemma 2.43. *Let $f \in \mathcal{D}_r[0,1]$ and denote by $\lambda(\cdot)$ the Lebesgue measure on the unit interval. Then for any $\gamma > 0$ and $0 < \theta < 1$,*

$$\|f\| \geq \gamma \Rightarrow \lambda\left(\{t : |f(t) - g(t)| \geq (1-\theta)\gamma\}\right) \geq \frac{1}{2} r.$$

Moreover, for any $g \in \mathcal{C}[0,1]$ *it exists* $\delta = \delta(g, \gamma, \theta) > 0$ *such that*

$$\|f - g\| \geq \gamma \Rightarrow \lambda\left(\{t : |f(t) - g(t)| \geq (1-\theta)\gamma\}\right) \geq \frac{1}{2} \min\left(r, \delta\right).$$

Let $f \in \mathcal{C}_r[0,1]$ *and* γ, θ *as above. Then,*

$$\|f\| \geq \gamma \Rightarrow \lambda\left(\{t : |f(t) - g(t)| \geq (1-\theta)\gamma\}\right) \geq \frac{\theta}{8} r.$$

Moreover, for $g \in \mathcal{C}[0,1]$ *there exists* $\delta = \delta(g, \gamma, \theta) > 0$ *with*

$$\|f - g\| \geq \gamma \Rightarrow \lambda\left(\{t : |f(t) - g(t)| \geq (1-\theta)\gamma\}\right) \geq \frac{\theta}{8} \min\left(r, \delta\right).$$

We first give the proofs in the continuous case.

Proof. (Theorem 2.35) For $r > 0, x \in \mathcal{C}[0,1]$ let $B_r(x) = \{y \in \mathcal{C}[0,1] : \|y - x\| < r\}$. According to Lemma 2.41, we need to verify that

$$\mathbf{P}\left(X_n \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \to \mathbf{P}\left(X \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right), \tag{2.38}$$

for $I = \{1, \ldots, k\}$ and $x_1, \ldots, x_k \in S, \gamma_1, \ldots, \gamma_k > 0$ such that $\mathbf{P}\left(X \in (\partial B_{\gamma_i}(x_i))\right) = 0$. The lack of uniformity in (2.36) leads us to find lower and upper bounds for the desired quantity. We will establish

$$\limsup_{n \to \infty} \mathbf{P}\left(X_n \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \leq \mathbf{P}\left(X \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \tag{2.39}$$

and

$$\liminf_{n \to \infty} \mathbf{P}\left(X_n \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \geq \mathbf{P}\left(X \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \tag{2.40}$$

separated from each other. To this end we construct functions $g_{i,n}, \bar{g}_{i,n} : \mathcal{C}[0,1] \to [0,1]$ satisfying

$$\bar{g}_{i,n}(x) \leq \mathbf{1}_{\{B_{\gamma_i}(x_i)\}}(x) \leq g_{i,n}(x), \quad \text{for all } x \in \mathcal{C}_{r_n}[0,1], \tag{2.41}$$

$$g_{i,n}(x), \bar{g}_{i,n}(x) \to \mathbf{1}_{\{B_{\gamma_i}(x_i)\}}(x), \quad \text{for all } x \in \mathcal{C}[0,1] \backslash \partial B_{\gamma_i}(x_i), \tag{2.42}$$

and such that $a_n \prod_{i \in I} g_{i,n}, \bar{a}_n \prod_{i \in I} \bar{g}_{i,n} \in \mathcal{F}_s$ for appropriate constants $a_n, \bar{a}_n > 0$. Then we can conclude

$$\mathbf{P}\left(X_n \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \leq \mathbb{E}\left[\prod_{i \in I} g_{i,n}(X_n)\right] \leq \mathbb{E}\left[\prod_{i \in I} g_{i,n}(X)\right] + a_n^{-1} \zeta_s(X_n, X) \tag{2.43}$$

and

$$\mathbf{P}\left(X_n \in \bigcap_{i \in I} B_{\gamma_i}(x_i)\right) \geq \mathbb{E}\left[\prod_{i \in I} \bar{g}_{i,n}(X_n)\right] \geq \mathbb{E}\left[\prod_{i \in I} \bar{g}_{i,n}(X)\right] - \bar{a}_n^{-1} \zeta_s(X_n, X). \tag{2.44}$$

Now, if $a_n^{-1} \zeta_s(X_n, X)$ for $n \to \infty$ then (2.43) implies (2.39) and similarly (2.40) follows from (2.44)) if $\bar{a}_n^{-1} \zeta_s(X_n, X)$ as $n \to \infty$.

Let us a give motivation of how to construct the functions $g_{i,n}$. According to (2.42), asymptotically, the functions $g_{i,n}$ have to separate points $x \in \mathcal{C}[0,1]$ which are in $B_{\gamma_i}(x_i)$ from those which are not. This is why we use the L_p norm. Consider ψ_{p,x_i} as introduced in Lemma 2.42. If $x \in \overline{B_{\gamma_i}(x_i)}$ then $\psi_{p,x_i}(x) \le (1+\gamma_i^2)^{1/2}$ whereas if $x \notin \overline{B_{\gamma_i}(x_i)}$ then $\liminf_{p\to\infty} \psi_{p,x_i}(x) > (1+\gamma_i^2)^{1/2}$.
Let $\varphi : \mathbb{R} \to [0,1]$ be a three times continuously differentiable function with $\varphi(u) = 1$ for $u \le 0$ and $\varphi(u) = 0$ for $u \ge 1$. For $\varrho \in \mathbb{R}$ and $\eta > 0$ we denote $\varphi_{\varrho,\eta} : \mathbb{R}^+ \to [0,1]$ by $\varphi_{\varrho,\eta}(u) = \varphi((u-\varrho)/\eta)$.
Let $g_i(x) = \varphi_{(1+\gamma_i^2)^{1/2},\eta}(\psi_{p,x_i}(x))$. Let $g_{i,n} = g_i$ with $\eta = \eta_n \downarrow 0$ and $p = p_n \uparrow \infty$. Then $g_{i,n}$ has the properties in (2.41) and (2.42).
Now we construct $\bar g_{i,n}$. Let $0 < \theta < 1$ and $x \in \mathcal{C}_{r_n}[0,1]$. Since the family $(x_i)_{i\in I}$ is uniformly equicontinuous, by Lemma 2.43 we can find $\delta = \delta(\theta)$ (also depending on $x_1, \dots, x_k, \gamma_1, \dots, \gamma_k$ which are kept fixed) with

$$
\begin{aligned}
\{\|x - x_i\| \ge \gamma_i\} &\subseteq \left\{\lambda(\{t : |x(t) - x_i(t)| \ge \gamma_i(1-\theta)\}) \ge \frac{\theta}{8}\min(r_n, \delta)\right\} \\
&\subseteq \left\{\psi_{p,x_i}(x) \ge (1+\gamma_i^2(1-\theta)^2)^{1/2}\left(\frac{\theta}{8}\min(r_n,\delta)\right)^{1/p}\right\} \\
&\subseteq \{\bar g_{i,n}(x) = 0\},
\end{aligned}
$$

with $\bar g_{i,n}(x) = \varphi_{(1+\gamma_i^2(1-\theta)^2)^{1/2}(\theta\min(r_n,\delta)/8)^{1/p}-\eta,\eta}(\psi_{p,x_i}(x))$. This gives (2.41). $\bar g_{i,n}$ does not fulfill (2.42), but we have

$$
\bar g_{i,n}(x) \to \mathbf{1}_{\{B_{\gamma_i(1-\theta)}(x_i)\}}(x)
$$

for $x \in \mathcal{C}[0,1] \backslash \partial B_{\gamma_i(1-\theta)}(x_i)$ and $p = p_n \uparrow \infty, \eta = \eta_n \downarrow 0$ such that $r_n^{1/p_n} \to 1$. This gives for every $0 < \theta < 1$ with $\mathbf{P}\left(X \in \partial B_{\gamma_i(1-\theta)}(x_i)\right) = 0$ for all $i \in I$

$$
\lim_{n\to\infty} \mathbb{E}\left[\prod_{i\in I} \bar g_{i,n}(X)\right] = \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i(1-\theta)}(x_i)\right).
$$

Assuming that $\bar a_n \prod_{i\in I} \bar g_{i,n} \in \mathcal{F}_s$ and letting n tend to infinity (2.44) rewrites as

$$
\liminf_{n\to\infty} \mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \ge \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i(1-\theta)}(x_i)\right) - \limsup_{n\to\infty} \bar a_n^{-1}\zeta_s(X_n, X), \quad (2.45)
$$

where $\bar a_n$ may depend on θ and δ. Below, we will see that the error term on the right hand side of (2.45) vanishes as $n \to \infty$ uniformly in θ, δ. So choosing $\theta \downarrow 0$ such that $\mathbf{P}\left(X \in \partial B_{\gamma_i(1-\theta)}(x_i)\right) = 0$ for all $i \in I$ the assertion

$$
\liminf_{n\to\infty} \mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \ge \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right)
$$

follows.
It remains to show that the error terms vanish in the limit. By Lemma 2.42 $g(x) = \phi_{\varrho,\eta}(\psi_{p,y}(x))$ and using the Mean Value theorem we achieve for $m = 0,1,2$

$$
\|g^{(m)}(x+h) - g^{(m)}(x)\| \le C_m p^m \eta^{-(m+1)}\|h\|^\alpha
$$

for $p \ge 4, \eta < 1$ and some constants $C_m > 0$. It is easy to check that the same is valid for products of functions of form g with different constants, independent of the parameters. It follows that both error terms in (2.43) and (2.45) are bounded by $C'_m p_n^m \eta_n^{-(m+1)} \zeta_s(X_n, X)$ for all n, uniformly in θ, δ, where C'_m denotes a fixed constant for each $m \in \{0,1,2\}$. By (2.34) we can choose $p_n \uparrow \infty$ and $\eta_n \downarrow 0$ such that both $r_n^{1/p_n} \to 1$ and the error terms vanish in the limit.

□

Proof. (Corollary 2.36) Again, according to Lemma 2.41 we only have to verify (2.38), for which we modify the proof of Theorem 2.35: First note that the assumption of piecewise linearity of X_n and the convergence rate for $\zeta_s(X_n, X)$ are not necessary for the upper bound

$$\limsup_{n\to\infty} \mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \le \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right).$$

For the lower bound let $\varepsilon > 0$ and note that

$$\mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \quad \ge \quad \mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i) \cap \{Y_n \in \mathcal{C}_{r_n}[0,1]\}\right)$$

We modify the functions $\bar{g}_{i,n}(x)$. Let $0 < \gamma_{K_i} < \gamma_i$ such that

$$\mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_{K_i}}(x_i)\right) \ge \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) - \varepsilon.$$

and $\mathbf{P}\left(X \in \partial B_{\gamma_{K_i}}(x_i)\right) = 0$ for all i. Let $0 < \theta < 1$ and n_0 be large enough such that $\varrho_n = \|h_n - h\| < \min_i(\gamma_{K_i}(1-\theta) \wedge \gamma - \gamma_{K_i})$ and $\mathbf{P}\left(Y_n \notin \mathcal{C}_{r_n}[0,1]\right) < \varepsilon)$ for all $n \ge n_0$. Then, since the functions $(x_i - h)_{i\in I}$ are uniformly equicontinuous, by Lemma 2.43 there exists $\delta = \delta(\theta)$ such that for $y \in \mathcal{C}_{r_n}[0,1]$ with $x = y + h_n$ and $n \ge n_0$

$$\begin{aligned}
& \{\|x - x_i\| \ge \gamma_i\} \subseteq \{\|y + h - x_i\| \ge \gamma_{K_i}\} \\
\subseteq \quad & \left\{\lambda(\{t : |y(t) + h(t) - x_i(t)| \ge \gamma_{K_i}(1-\theta)\}) \ge \frac{\theta}{8}\min(r_n, \delta)\right\} \\
\subseteq \quad & \left\{\lambda(\{t : |x(t) - x_i(t)| \ge \gamma_{K_i}(1-\theta) - \varrho_n\}) \ge \frac{\theta}{8}\min(r_n, \delta)\right\} \\
\subseteq \quad & \left\{\psi_{p,x_i}(x) \ge (1 + (\gamma_{K_i}(1-\theta) - \varrho_n)^2)^{1/2}\left(\frac{\theta}{8}\min(r_n, \delta)\right)^{1/p}\right\} \\
\subseteq \quad & \{\bar{g}_{i,n}(x) = 0\},
\end{aligned}$$

with $\bar{g}_{i,n}(x) = \phi_{(1+(\gamma_{K_i}(1-\theta)-\varrho_n)^2)^{1/2}(\theta\min(r_n,\delta)/8)^{1/p}-\eta,\eta}(\psi_{p,x_i}(x))$. Hence,

$$\begin{aligned}
\mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \quad & \ge \quad \mathbb{E}\left[\prod_{i\in I} \bar{g}_{i,n}(X_n)\mathbf{1}_{\{Y_n \in \mathcal{C}_{r_n}[0,1]\}}\right] \\
& \ge \quad \mathbb{E}\left[\prod_{i\in I} \bar{g}_{i,n}(X_n)\right] - \varepsilon
\end{aligned}$$

for $n \ge n_0$. The upper bound of the error term $\bar{a}_n^{-1}\zeta_s(X_n, X)$ is a function of p and η so it is uniform in $\varrho_n, \theta, \delta$. Following the same lines as in the proof of Theorem 2.35 gives

$$\begin{aligned}\liminf_{n\to\infty} \mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) &\geq \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_{K_i}}(x_i)\right) - \varepsilon \\ &\geq \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) - 2\varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the result follows. $\square$

Proof. (Corollary 2.37) In the setting of the proof of Theorem 2.35, (2.43) rewrites as

$$\begin{aligned}\mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) &\leq \mathbb{E}\left[\prod_{i\in I} g_{i,n}(X_n)\right] \leq \mathbb{E}\left[\prod_{i\in I} g_{i,n}(Y_n)\right] + a_n^{-1}\zeta_s(X_n, Y_n) \\ &= \mathbb{E}\left[\prod_{i\in I} g_{i,n}(Y_n)\right] - \mathbb{E}\left[\prod_{i\in I} g_{i,n}(X)\right] \\ &\quad +\mathbb{E}\left[\prod_{i\in I} g_{i,n}(X)\right] + a_n^{-1}\zeta_s(X_n, Y_n)\end{aligned}$$

We may choose $Y_n \to X$ almost surely. On the event $\{X \in B_{\gamma_i}(x_i)\}$ we have $\lim_n g_{i,n}(Y_n) = \lim_n g_{i,n}(X) = 1$ and on $\{X \notin \overline{B_{\gamma_i}(x_i)}\}$ we have $\lim_n g_{i,n}(Y_n) = \lim_n g_{i,n}(X) = 0$. Since $\mathbf{P}\left(X \in \partial B_{\gamma_i}(x_i)\right) = 0$ it follows

$$\prod_{i\in I} g_{i,n}(Y_n) - \prod_{i\in I} g_{i,n}(X) \to 0$$

for $n \to \infty$ almost surely and dominated convergence yields

$$\limsup_{n\to\infty} \mathbf{P}\left(X_n \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right) \leq \mathbf{P}\left(X \in \bigcap_{i\in I} B_{\gamma_i}(x_i)\right),$$

just like in the proof of Theorem 2.35. The lower bound follows similarly. $\square$

We now head to the case of càdlàg functions. We only discuss the approach in the proof of Theorem 2.35. Following exactly the same arguments as in the continuous case and using the additional statements of Lemma 2.42 and Lemma 2.43, it is easy to see that we also obtain (2.38) if the balls $B_{\gamma_i}(x_i)$ are defined with the uniform metric in $\mathcal{D}[0,1]$. Remember that we still have $x_i \in \mathcal{C}[0,1]$. Note, that it is at the core of Skorokhod's representation theorem [Bil99, Theorem 6.7] that, if X is continuous and (2.38) is satisfied, we can find versions X_n that converge almost surely to X in the sense that $\|X_n - X\| \to 0$ as $n \to \infty$. Here, we might have to change the underlying probability space which is inessential. This implies $d_{sk}(X_n, X) \to 0$ almost surely, hence the assertion.

The proof of Theorem 2.38 is close to the one of Lemma 5.3 in [DJN08]. The L_p approximation of the supremum norm complicates the argument slightly. We only give the proof in the case of $\mathcal{C}[0,1]$, the modifications in the càdlàg case are obvious.

Proof. (Theorem 2.38) Suppose $s \in \{1, 2, 3\}$ and that the first assumption of Theorem 2.38 is satisfied. Let $\kappa : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ be a smooth, monotonic function with $\kappa(u) = 0$ for $0 \leq u \leq \frac{1}{2}$ and $\kappa(u) = u^s$ for $u \geq 1$. We may as well assume that the interpolation for $\frac{1}{2} \leq u \leq 1$ is done smoothly such that we have $\kappa(u) \leq u^s$ for $\frac{1}{2} \leq u \leq 1$, thus $\kappa(u) \leq u^s$ for all $u \in \mathbb{R}^+$. Let $f, f^{(p)} : \mathcal{C}[0,1] \to \mathbb{R}$ be given by

$$\begin{aligned} f(x) &= \kappa(\|x\|), \\ f^{(p)}(x) &= \kappa(L_p(x)). \end{aligned}$$

By Lemma 2.42, the restrictions of L_p and $f^{(p)}$ to $\mathcal{C}[0,1]\backslash\{\mathbf{0}\}$ are smooth. Furthermore, all derivatives of $f^{(p)}$ vanish for $\|x\| < 1/2$ which implies that $f^{(p)}$ is smooth on $\mathcal{C}[0,1]$. Again, by Lemma 2.42 it is easy to check that for any $k \in \{1, \dots, s\}$,

$$\|D^k f^{(p)}(x)\| = O(p^{k-1}\|x\|^{s-k}),$$

uniformly in p and $x \in \mathcal{C}[0,1]$. Hence, $\|D^s f^{(p)}(z)\| = O(p^m)$ uniformly for all z, in particular for the set $[x, y] := \{\lambda x + (1 - \lambda)y \,|\, \lambda \in [0,1]\}$, and by the mean value theorem

$$\|D^m f^{(p)}(x) - D^m f^{(p)}(y)\| = O(p^m \|x - y\|).$$

Hence, there is a constant $c > 0$ such that $cp^{-m} f^{(p)} \in \mathcal{F}_s$ for all $p \geq 4$. We define, for $r > 0$,

$$\begin{aligned} f_r(x) &:= cr^s f(x/r), \\ f_r^{(p)}(x) &:= cr^s f^{(p)}(x/r). \end{aligned}$$

Then $p^{-m} f_r^{(p)} \in \mathcal{F}_s$. Furthermore, $f_r(x)$ and $f_r^{(p)}(x)$ are bounded by $c\|x\|^s$ for all $x \in \mathcal{C}[0,1]$, uniformly in p. For any fixed x we have $f_r(x) \to 0$ and $\sup_{p \geq 4} f_r^{(p)}(x) \to 0$ as $r \to \infty$. Hence, by $\mathbb{E}[\|X\|^s] < \infty$ and dominated convergence this implies

$$\mathbb{E}\left[\sup_{p \geq 4} f_r^{(p)}(X)\right] \to 0, \quad r \to \infty. \tag{2.46}$$

By definition of ζ_s we have

$$\mathbb{E}\left[f_r^{(p)}(X_n)\right] \leq \mathbb{E}\left[f_r^{(p)}(X)\right] + p^m \zeta_s(X_n, X).$$

By definition of f_r, for $\|x\| > r$ we have $\|x\|^s = c^{-1} f_r(x)$. Hence,

$$\begin{aligned} \mathbb{E}\left[\|X_n\|^s \mathbf{1}_{\{\|X_n\| \geq 2r\}}\right] &= c^{-1}\mathbb{E}\left[f_r(X_n)\mathbf{1}_{\{\|X_n\| \geq 2r\}}\right] \\ &\leq c^{-1}\mathbb{E}\left[f_r^{(p)}(X_n)\right] + c^{-1}\left(\mathbb{E}\left[(f_r(X_n) - f_r^{(p)}(X_n))\mathbf{1}_{\{\|X_n\| \geq 2r\}}\right]\right) \\ &\leq c^{-1}\mathbb{E}\left[f_r^{(p)}(X)\right] + c^{-1}p^m \zeta_s(X_n, X) \\ &\quad + c^{-1}\left(\mathbb{E}\left[(f_r(X_n) - f_r^{(p)}(X_n))\mathbf{1}_{\{\|X_n\| \geq 2r\}}\right]\right). \end{aligned} \tag{2.47}$$

Now, let $\varepsilon > 0$ be arbitrary. By (2.46) fix $r > 0$ such that $\mathbb{E}\left[f_r^{(p)}(X)\right] < \varepsilon$ for all $p \geq 4$. Additionally, by the given assumptions there exists a sequence $p_n \uparrow \infty$ such that

$$\frac{\log r_n}{p_n} \to 0, \qquad p_n^m \zeta_s(X_n, X) \to 0, \qquad (n \to \infty).$$

Therefore, let N_0 be large enough such that $p_n^m \zeta_s(X_n, X) < \varepsilon$ for all $n \geq N_0$. It remains to bound the third summand in (2.47). Using Lemma 2.43, piecewise linearity of X_n implies that for all $0 < \theta < 1$,

$$L_p(X_n) \geq \|X_n\|(1-\theta)\left(\frac{\theta r_n}{8}\right)^{1/p_n}.$$

In particular, we have $L_p(X_n) \geq \frac{\|X_n\|}{2}$ for all n sufficiently large. For those n and $\|X_n\| > 2r$ we also have $f_r^{(p)}(X_n) = cL_p^s(X_n)$. This yields

$$\mathbb{E}\left[(f_r(X_n) - f_r^{(p)}(X_n))\mathbf{1}_{\{\|X_n\| \geq 2r\}}\right] = c\mathbb{E}\left[(\|X_n\|^s - L_p^s(X_n))\mathbf{1}_{\{\|X_n\| \geq 2r\}}\right] \tag{2.48}$$

$$\leq c(1-2^{-s})\mathbb{E}\left[\|X_n\|^s \mathbf{1}_{\{\|X_n\| \geq 2r\}}\right]. \tag{2.49}$$

for all n sufficiently large. Increasing N_0 if necessary, inserting (2.49) into (2.47) and rearranging terms implies

$$\mathbb{E}\left[\|X_n\|^s \mathbf{1}_{\{\|X_n\| \geq 2r\}}\right] \leq 2^{1+s}c^{-1}\varepsilon.$$

for all $n \geq N_0$. Since ε was arbitrary, the assertion follows.
Now, suppose the second assumption to be satisfied. Then, we have to modify the last part of the proof. In (2.48) we can decompose

$$L_p^s(X_n) = L_p^s(X_n)\mathbf{1}_{\{Y_n \in \mathcal{C}_{r_n}[0,1]\}} + L_p^s(X_n)\mathbf{1}_{\{Y_n \notin \mathcal{C}_{r_n}[0,1]\}}.$$

Using $L_p^s(X_n) \leq \|X_n\|^s$, the assumptions guarantee the expectation of the second term to be small in the limit $n \to \infty$. For the first one, using similar arguments as above, given $\{Y_n \in \mathcal{C}_{r_n}[0,1]\}$, we find

$$L_p(X_n) \geq \frac{\|X_n\|}{2} - 2\varrho_n$$

with $\varrho_n = \|h_n - h\|$ for all n sufficiently large. Proceeding as in the first part, we obtain the result. Given the third assumption, it only remains to bound $\mathbb{E}\left[f_r^{(p)}(Y_n)\right]$ which appears instead of $\mathbb{E}\left[f_r^{(p)}(X)\right]$ by $\mathbb{E}\left[f_r^{(p)}(Z)\right]$ in (2.47). □

3. The contraction method

Originally, the contraction method for distributional recurrences (1.2) illustrated in the introduction is based on a metric on the set of probability distributions satisfying the following three properties:

- The metric distance between $\mathcal{L}(X_n)$ and $\mathcal{L}(X_m)$ is finite for all n, m.
- Convergence in the given metric implies weak convergence.
- The metric is complete on appropriate subsets of the entire set of measures.

With respect to these main points the results of the last chapter are rather disappointing. First, the restrictions imposed on distributions to have finite ζ_s distance are considerably strong; the necessary scaling gives rise to the substantial problems **P3a** and **P3b**, which we will solve in our applications in Chapters 5 and 4. Second, weak convergence can only be deduced after establishing a rate of convergence with respect to ζ_s and regularity conditions on the paths of X_n. Again, this causes problems that have to be addressed in detail.
Regarding the last point, note that, looking at Banach's fixed-point theorem for complete metric spaces, one usually puts most emphasis on the existence of a fixed-point for a given contractive map. However, one should not forget that, once the fixed-point has been established by different means, both the statement of uniqueness and the exponential speed of convergence for the distance between the successive iteration started at a valid point and the fixed-point itself remain valid independently of the completeness property. Throughout this chapter, we will be led by this observation while developing the contraction method based on the class of ζ_s metrics.

3.1. The main result: A functional limit theorem

The contraction method is developed first for a general separable Banach space B and the space $(\mathcal{D}[0,1], d_{sk})$. Then the framework is specialized to the cases $(\mathcal{C}[0,1], \|\cdot\|)$ and $(\mathcal{D}[0,1], d_{sk})$. For this section B will always denote a separable Banach space or $(\mathcal{D}[0,1], d_{sk})$. We recall the recursive equation (1.2). We have

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \quad n \geq n_0. \tag{3.1}$$

where $A_1^{(n)}, \ldots, A_K^{(n)}$ are random continuous linear operators, $b^{(n)}$ is a B-valued random variable, $(X_n^{(1)})_{n\geq 0}, \ldots, (X_n^{(K)})_{n\geq 0}$ are distributed like $(X_n)_{n\geq 0}$, and $I^{(n)} = (I_1^{(n)}, \ldots, I_K^{(n)})$ is a vector of random integers in $\{0, \ldots, n\}$. Moreover $(A_1^{(n)}, \ldots, A_K^{(n)}, b^{(n)}, I^{(n)})$, $(X_n^{(1)})_{n\geq 0}, \ldots, (X_n^{(K)})_{n\geq 0}$ are independent and $n_0 \in \mathbb{N}$.
Recall that in order to be a random continuous linear operator, A has to take values in the set of continuous endomorphisms on $\mathcal{C}[0,1]$ respectively the set of norm-continuous endomorphisms that are continuous with respect to d_{sk} on $\mathcal{D}[0,1]$ such that $A(x)(t)$ is a real-valued random variable

for all $x \in \mathcal{C}[0,1]$ respectively $x \in \mathcal{D}[0,1]$ and $t \in [0,1]$. In $\mathcal{D}[0,1]$ we additionally have to guarantee $\|A\|$ to be a real-valued random variable.
Next, we make assumptions about the moments and the asymptotic behavior of the coefficients $A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)}$. For a random continuous linear operator A and for random variable X with values in B we write

$$\|A\|_s := \mathbb{E}\left[\|A\|^s\right]^{1\wedge(1/s)}.$$

We consider the following conditions with an $s > 0$:

C1. We have $\|X_0\|_s, \dots, \|X_{n_0-1}\|_s, \|A_r^{(n)}\|_s, \|b^{(n)}\|_s < \infty$ for all $r = 1, \dots, K$ and $n \geq 0$ and there exist random continuous linear operators $A_1, \dots, A_K$ on B and a B-valued random variable b such that, as $n \to \infty$,

$$\gamma(n) := \|b^{(n)} - b\|_s + \sum_{r=1}^K \left(\|A_r^{(n)} - A_r\|_s + \left\| \mathbf{1}_{\{I_r^{(n)} \leq n_0\}} A_r^{(n)} \right\|_s \right) \to 0.$$

and for all $\ell \in \mathbb{N}$,

$$\mathbb{E}\left[\mathbf{1}_{\{I_r^{(n)} \in \{0,\dots,\ell\} \cup \{n\}\}} \|A_r^{(n)}\|^s \right] \to 0.$$

C2. We have

$$L := \sum_{r=1}^K \mathbb{E}\left[\|A_r\|^s\right] < 1.$$

The limits of the coefficients determine the limiting operator T from (1.5):

$$\begin{aligned} T : \mathcal{M}(B) &\to \mathcal{M}(B) \\ \mu &\mapsto \mathcal{L}\left(\sum_{r=1}^K A_r Z^{(r)} + b \right) \end{aligned} \tag{3.2}$$

where $(A_1, \dots, A_K, b)$, $Z^{(1)}, \dots, Z^{(K)}$ are independent and $Z^{(1)}, \dots, Z^{(K)}$ have distribution μ.

C3. The map T has a fixed-point $\eta \in \mathcal{M}_s(B)$, such that $\mathcal{L}(X_n) \in \mathcal{M}_s(\eta)$ for all $n \geq n_0$.

The existence of a fixed-point is not in general implied by contraction properties of T with respect to a Zolotarev metric due to the lack of completeness of the metric on the space B. However, we can argue that there is at most one fixed-point of T in $\mathcal{M}_s(\eta)$:

Lemma 3.1. *Assume the sequence $(X_n)_{n\geq 0}$ satisfies (3.1). Under conditions* **C1**–**C3** *we have $T(\mathcal{M}_s(\eta)) \subseteq \mathcal{M}_s(\eta)$ and*

$$\zeta_s(T(\mu), T(\lambda)) \leq L\zeta_s(\mu, \lambda) \quad \textit{for all } \mu, \lambda \in \mathcal{M}_s(\eta).$$

In particular, the restriction of T to $\mathcal{M}_s(\eta)$ is a contraction and has the unique fixed-point η.

Proof. Let $\mu \in \mathcal{M}_s(\eta)$. Recall that we have $s = m + \alpha$ with $m \in \mathbb{N}_0$ and $\alpha \in (0,1]$. We introduce an accompanying sequence

$$Q_n := \sum_{r=1}^{K} A_r^{(n)} \left(\mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} X_{I_r^{(n)}}^{(r)} + \mathbf{1}_{\left\{I_r^{(n)} \geq n_0\right\}} Z^{(r)} \right) + b^{(n)}, \quad n \geq n_0, \tag{3.3}$$

where $(A_1^{(n)}, \dots, A_K^{(n)}, b^{(n)})$, $Z^{(1)}, \dots, Z^{(K)}$ are independent and $Z^{(1)}, \dots, Z^{(K)}$ have distribution η.
We first show that $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$ for all $n \geq n_0$. Condition **C1**, conditioning on the coefficients and Minkowski's inequality imply $\mathbb{E}\left[\|Q_n\|^s\right] < \infty$ for all n. For $s \leq 1$ we already obtain $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$.
For $s > 1$ we choose arbitrary $1 \leq k \leq m$ and multilinear and bounded $f : B^k \to \mathbb{R}$. We have

$$\begin{aligned} \mathbb{E}\left[f(Z, \dots, Z)\right] &= \mathbb{E}\left[f(X_n, \dots, X_n)\right] \\ &= \mathbb{E}\left[f\left(\sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)}, \dots, \sum_{r=1}^{K} A_r^{(n)} X_{I_r^{(n)}}^{(r)} + b^{(n)} \right)\right]. \end{aligned}$$

To show $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$ we need to verify that the latter display is equal to $\mathbb{E}\left[f(Q_n, \dots, Q_n)\right]$: Since f is multilinear, both terms can be expanded as a sum and it suffices to show that the corresponding summands are equal:

$$\mathbb{E}\left[f\left(C_{j_1}^{(n)}, \dots, C_{j_k}^{(n)} \right)\right] \quad = \quad \mathbb{E}\left[f\left(D_{j_1}^{(n)}, \dots, D_{j_k}^{(n)} \right)\right], \tag{3.4}$$

where $j_1, \dots, j_k \in \{1, \dots, K\}$ and for each $i \in \{1, \dots, k\}$ we either have

$$C_{j_i}^{(n)} = A_{j_i}^{(n)} X_{I_{j_i}^{(n)}}^{(j_i)} \quad \text{and} \quad D_{j_i}^{(n)} = A_{j_i}^{(n)} \left(\mathbf{1}_{\left\{I_{j_i}^{(n)} < n_0\right\}} X_{I_{j_i}^{(n)}}^{(j_i)} + \mathbf{1}_{\left\{I_{j_i}^{(n)} \geq n_0\right\}} Z^{(j_i)} \right) \tag{3.5}$$

or

$$C_{j_i}^{(n)} = b^{(n)} \quad \text{and} \quad D_{j_i}^{(n)} = b^{(n)}. \tag{3.6}$$

The equality in (3.4) is obvious for the case where we have (3.6) for all $i = 1, \dots, k$. For the other cases we have (3.5) for at least $1 \leq \ell \leq k$ arguments of f, say, for simplicity of presentation, for the first ℓ with $1 \leq \ell_1 < \dots < \ell_d = \ell$ such that $j_s = j_{\ell_i}$ for all $s = \ell_{i-1}+1, \dots, \ell_i, i = 1, \dots, d$ and j_{ℓ_i} pairwise different for $i = 1, \dots, d$ (by convention $\ell_0 := 0$). The claim in (3.4) reduces to

$$\begin{aligned} &\mathbb{E}\left[f(C_{j_{\ell_1}}^{(n)}, \dots, C_{j_{\ell_1}}^{(n)}, C_{j_{\ell_2}}^{(n)}, \dots, C_{j_{\ell_d}}^{(n)}, b^{(n)}, \dots, b^{(n)} \right] \\ &= \mathbb{E}\left[f(D_{j_{\ell_1}}^{(n)}, \dots, D_{j_{\ell_1}}^{(n)}, D_{j_{\ell_2}}^{(n)}, \dots, D_{j_{\ell_d}}^{(n)}, b^{(n)}, \dots, b^{(n)} \right] \end{aligned} \tag{3.7}$$

We will prove that, for each $p \in \{1, \dots, d\}$,

$$\begin{aligned} &\mathbb{E}\left[f(C_{j_{\ell_1}}^{(n)}, \dots, C_{j_{\ell_{p-1}}}^{(n)}, C_{j_{\ell_p}}^{(n)}, \dots, C_{j_{\ell_p}}^{(n)}, D_{j_{\ell_{p+1}}}^{(n)}, \dots, D_{j_{\ell_d}}^{(n)}, b^{(n)}, \dots, b^{(n)}) \right] \\ &= \mathbb{E}\left[f(C_{j_{\ell_1}}^{(n)}, \dots, C_{j_{\ell_{p-1}}}^{(n)}, D_{j_{\ell_p}}^{(n)}, \dots, D_{j_{\ell_p}}^{(n)}, D_{j_{\ell_{p+1}}}^{(n)}, \dots, D_{j_{\ell_d}}^{(n)}, b^{(n)}, \dots, b^{(n)}) \right], \end{aligned} \tag{3.8}$$

which in turn implies (3.7). Abbreviating $Y_i^{(r)} = \left(\mathbf{1}_{\{i<n_0\}}X_i^{(r)} + \mathbf{1}_{\{i\geq n_0\}}Z^{(r)}\right)$ and denoting by Υ the joint distribution of $\left(A_{j_{\ell_1}}^{(n)}, \dots, A_{j_{\ell_d}}^{(n)}, I_{j_{\ell_1}}^{(n)}, \dots, I_{j_{\ell_d}}^{(n)}, b^{(n)}\right)$ we have

$$
\begin{aligned}
&\mathbb{E}\left[f(C_{j_{\ell_1}}^{(n)}, \dots, C_{j_{\ell_{i-1}}}^{(n)}, C_{j_{\ell_i}}^{(n)}, \dots, C_{j_{\ell_i}}^{(n)}, D_{j_{\ell_{i+1}}}^{(n)}, \dots, D_{j_{\ell_d}}^{(n)}, b^{(n)}, \dots, b^{(n)}\right] \\
&= \int f\left(\alpha_1 x_1, \dots, \alpha_{p-1}x_{p-1}, \alpha_p x_p, \dots, \alpha_p x_p, \alpha_{p+1}x_{p+1}, \dots, \alpha_d x_d, b, \dots, b\right) \\
&d\mathbb{P}_{X_{i_1}}(x_1)\cdots d\mathbb{P}_{X_{i_p}}(x_p) d\mathbb{P}_{Y_{i_{p+1}}}(x_{p+1})\cdots d\mathbb{P}_{Y_{i_d}}(x_d) d\Upsilon(\alpha_1, \dots, \alpha_d, i_1, \dots, i_d, b) \\
&= \int \mathbb{E}\left[g(X_{i_p}, \dots, X_{i_p})\right] d\mathbb{P}_{X_{i_1}}\cdots d\mathbb{P}_{X_{i_{p-1}}} d\mathbb{P}_{Y_{i_{p+1}}}\cdots d\mathbb{P}_{Y_{i_d}} d\Upsilon,
\end{aligned}
$$

where, for all fixed $\alpha_1, \dots, \alpha_d, i_1, \dots, i_d, b, x_1, \dots, x_{p-1}, x_{p+1}, \dots, x_d$, we use the continuous multilinear function $g : B^{\ell_p - \ell_{p-1}} \to \mathbb{R}$,

$$
\begin{aligned}
&g(y_1, \dots, y_{\ell_p - \ell_{p-1}}) \\
&:= f(\alpha_1 x_1, \dots, \alpha_{p-1}x_{p-1}, \alpha_p y_1, \dots, \alpha_p y_{\ell_p - \ell_{p-1}}, \alpha_{p+1}x_{p+1}, \dots, \alpha_d x_d, b, \dots, b).
\end{aligned}
$$

Since $\mathcal{L}(X_m), \mathcal{L}(Z) \in \mathcal{M}_s(\eta)$ for all $m \geq n_0$ we can replace X_{i_p} by Y_{i_p}. This shows the equality (3.8), hence (3.4). Altogether, we obtain $\mathcal{L}(Q_n) \in \mathcal{M}_s(\eta)$ for all $n \geq n_0$.
Now, we show $T(\mu) \in \mathcal{M}_s(\eta)$. Let W be a random variable with distribution $T(\mu)$. By **C2**, in particular $\|A_r\|_s < \infty$ for $r = 1, \dots, K$, by **C1** we have $\|b\|_s < \infty$. Thus, as for Q_n, from Minkowski's inequality we obtain $\mathbb{E}\left[\|W\|^s\right] < \infty$, hence $T(\mu) \in \mathcal{M}_s(\eta)$ for $s \leq 1$. For the case $s > 1$ we consider again arbitrary $1 \leq k \leq m$ and multilinear and bounded $f : B^k \to \mathbb{R}$. It suffices to show $\mathbb{E}\left[f(Q_n, \dots, Q_n)\right] = \mathbb{E}\left[f(W, \dots, W)\right]$ for some $n \geq n_0$. In fact, we will show that $\lim_{n\to\infty} \mathbb{E}\left[f(Q_n, \dots, Q_n)\right] = \mathbb{E}\left[f(W, \dots, W)\right]$. For this we expand

$$
\mathbb{E}\left[f(W, \dots, W)\right] = \mathbb{E}\left[f\left(\sum_{r=1}^{K} A_r Z^{(r)} + b, \dots, \sum_{r=1}^{K} A_r Z^{(r)} + b\right)\right]
$$

into summands corresponding to (3.4) and have to show that

$$
\lim_{n\to\infty} \mathbb{E}\left[f\left(D_{j_1}^{(n)}, \dots, D_{j_k}^{(n)}\right)\right] = \mathbb{E}\left[f(E_{j_1}, \dots, E_{j_k})\right],
$$

where $j_1, \dots, j_k \in \{1, \dots, K\}$. For each $i \in \{1, \dots, k\}$ we have in case (3.5) that $E_{j_i} = A_{j_i} Z^{(j_i)}$, in case (3.6) that $E_{j_i} = b$. We obtain, introducing a telescoping sum and using Hölder's inequality,

$$
\begin{aligned}
&\left|\mathbb{E}\left[f\left(D_{j_1}^{(n)}, \dots, D_{j_k}^{(n)}\right)\right] - \mathbb{E}\left[f(E_{j_1}, \dots, E_{j_k})\right]\right| \\
&= \left|\sum_{q=1}^{k} \mathbb{E}\left[f\left(E_{j_1}, \dots, E_{j_{q-1}}, D_{j_q}^{(n)}, \dots, D_{j_k}^{(n)}\right) - f\left(E_{j_1}, \dots, E_{j_q}, D_{j_{q+1}}^{(n)}, \dots, D_{j_k}^{(n)}\right)\right]\right| \\
&\leq \sum_{q=1}^{k} \left|\mathbb{E}\left[f\left(E_{j_1}, \dots, E_{j_{q-1}}, D_{j_q}^{(n)} - E_{j_q}, D_{j_{q+1}}^{(n)}, \dots, D_{j_k}^{(n)}\right)\right]\right| \\
&\leq \sum_{q=1}^{k} \|f\| \|D_{j_q}^{(n)} - E_{j_q}\|_k \prod_{v=1}^{q-1} \|E_{j_v}\|_k \prod_{v=q+1}^{k} \|D_{j_v}^{(n)}\|_k.
\end{aligned}
$$

Note that the $\|E_{j_v}\|_k$ and $\|D_{j_v}^{(n)}\|_k$ are all uniformly bounded by independence, **C1**, and $\|X_0\|_s$, $\dots, \|X_{n_0-1}\|_s$, $\|Z\|_s < \infty$. Hence it suffices to show that $\|D_{j_v}^{(n)} - E_{j_v}\|_k \to 0$ for all j_v. In case (3.6) this is $\|b^{(n)} - b\|_k \to 0$ by condition **C1**. In case (3.6) we have, abbreviating $r = j_i$,

$$\begin{aligned}&\left\| A_r^{(n)} \left(\mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} X_{I_r^{(n)}}^{(r)} + \mathbf{1}_{\left\{I_r^{(n)} \geq n_0\right\}} Z^{(r)} \right) - A_r Z^{(r)} \right\|_k \\ &\leq \left\| (A_r^{(n)} - A_r) Z^{(r)} \right\|_k + \left\| A_r^{(n)} \left(\mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} \left(X_{I_r^{(n)}}^{(r)} - Z^{(r)} \right) \right) \right\|_k\end{aligned}$$

The first summand of the latter display tends to zero by independence, $\|Z\|_s < \infty$ and condition **C1**. The second summand tends to zero applying Hölder's inequality, condition **C1**, which implies that $\|A_r^{(n)}\|_s$ in uniformly bounded, $\|X_0\|_s, \dots, \|X_{n_0-1}\|_s, \|Z\|_s < \infty$ and conditions **C1** and **C3**. Altogether we obtain $T(\mu) \in \mathcal{M}_s(\eta)$.
Let $\mu, \lambda \in \mathcal{M}_s(\eta)$. Using Lemma 2.4 and Corollary 2.6 it follows that

$$\zeta_s(T(\mu), T(\lambda)) \leq \left(\sum_{r=1}^{K} \mathbb{E}\left[\|A_r\|^s\right] \right) \zeta_s(\mu, \lambda).$$

Thus, by condition **C2**, the restriction of T to $\mathcal{M}_s(\eta)$ is a contraction with respect to ζ_s.
Assume, μ was a fixed-point of T as well. Then the contraction property implies

$$\zeta_s(\mu, \eta) = \zeta_s(T(\mu), T(\eta)) \leq L\zeta_s(\mu, \eta),$$

hence $\zeta_s(\mu, \eta) = 0$. Since the ζ_s-distance is a metric on $\mathcal{M}_s(\eta)$ it follows $\mu = \eta$. □

Aiming to prove convergence of X_n to a fixed-point of (3.2), the conditions **C1**, **C2** and $\mathcal{L}(X_n) \in \mathcal{M}_s(\mu)$ for $n \geq n_0$, are natural in the context of contraction method. The existence of a solution of the fixed-point equation in condition **C3** is required since we miss knowledge about completeness of the ζ_s metrics. If we only assume **C1**, **C2** and $\mathcal{L}(X_n) \in \mathcal{M}_s(\mu)$ for $n \geq n_0$, then $(T^n(\mu))_{n\geq 0}$ is a Cauchy sequence with respect to ζ_s, a proof thereof runs along similar lines as for the previous proposition. Thus, by Proposition 2.40, $T^n(\mu)$ converges in $\xrightarrow{\text{fdd}}$ to some measure ν on $\mathbb{R}^{[0,1]}$, the natural candidate for a fixed-point of (3.2). Indeed, if η is such a fixed-point in $\mathcal{M}_s(\mu)$, then $\zeta_s(T^n(\mu), \eta) \to 0$ exponentially fast and therefore η has to be a continuous version of ν.

The following proposition uses the ideas developed so far to infer convergence of X_n to X in the ζ_s distance. The proof extends a similar proof for the case $B = \mathbb{R}^d$, see [NR04b, Theorem 4.1]. We draw further implications from this proof, see Corollary 3.5.

Proposition 3.2. *Let $(X_n)_{n\geq 0}$ satisfy recurrence (3.1) with conditions* **C1 – C3**. *Then for the fixed-point $\eta = \mathcal{L}(X)$ of T in (3.2) we have, as $n \to \infty$,*

$$\zeta_s(X_n, X) \to 0.$$

Proof. We use the accompanying sequence defined in (3.3). Throughout the proof let $n \geq n_0$. Again since the ζ_s-distance is a metric we have

$$\zeta_s(X_n, X) \leq \zeta_s(X_n, Q_n) + \zeta_s(Q_n, X). \tag{3.9}$$

First, we consider the second term. By **C1** and Minkowski's inequality, absolute moments of order s of the sequence $(Q_n)_{n\geq n_0}$ are bounded, hence using Theorem 2.17 it suffices to show

$$\ell_s(Q_n, X) \to 0.$$

Using the same set of independent random variables $X^{(1)}, \dots, X^{(K)}$ for Q_n and in the recurrence of X, we obtain

$$\begin{aligned}
\ell_s(Q_n, X) &\leq \left\| \sum_{r=1}^{K} \left(A_r - \mathbf{1}_{\left\{I_r^{(n)} \geq n_0\right\}} A_r^{(n)} \right) X^{(r)} \right\|_s + \left\| \sum_{r=1}^{K} \mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} A_r^{(n)} X^{(r)}_{I_r^{(n)}} \right\|_s + \|b^{(n)} - b\|_s \\
&\leq \sum_{r=1}^{K} \left(\left\| A_r - A_r^{(n)} \right\|_s + \left\| \mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} \|A_r^{(n)}\| \right\|_s \right) \|X\|_s + \|b^{(n)} - b\|_s \\
&\quad + \left\| \sum_{r=1}^{K} \mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} A_r^{(n)} X^{(r)}_{I_r^{(n)}} \right\|_s
\end{aligned}$$

By **C1** the first two summands tend to zero. Also, the third one converges to zero using **C1** and

$$\left\| \mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} \|A_r^{(n)}\| X^{(r)}_{I_r^{(n)}} \right\|_s \leq \left\| \mathbf{1}_{\left\{I_r^{(n)} < n_0\right\}} \|A_r^{(n)}\| \right\|_s \left\| \sup_{j<n_0} \|X_j\| \right\|_s$$

Furthermore, conditioning on the coefficients and using that ζ_s is $(s,+)$ ideal and Lemma 2.4, it is easy to see that

$$\begin{aligned}
\zeta_s(Q_n, X_n) &\leq p_n \zeta_s(X_n, X) + \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\left\{n_0 \leq I_r^{(n)} \leq n-1\right\}} \|A_r^{(n)}\|^s \zeta_s\left(X_{I_r^{(n)}}, X\right) \right] \qquad (3.10)\\
&\leq p_n \zeta_s(X_n, X) + \left(\sum_{r=1}^{K} \mathbb{E}\left[\|A_r^{(n)}\|^s \right] \right) \sup_{n_0 \leq i \leq n-1} \zeta_s(X_i, X), \qquad (3.11)
\end{aligned}$$

where

$$p_n = \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\left\{I_r^{(n)} = n\right\}} \|A_r^{(n)}\|^s \right] \to 0, \quad n \to \infty.$$

Combining (3.9) and (3.11) implies

$$\zeta_s(X_n, X) \leq \frac{1}{1-p_n} \left[\sum_{r=1}^{K} \mathbb{E}\left[\|A_r^{(n)}\|^s \right] \sup_{n_0 \leq i \leq n-1} \zeta_s(X_i, X) + o(1) \right].$$

From this it follows that $\zeta_s(X_n, X)$ is bounded. Let

$$\bar{\eta} := \sup_{n \geq n_0} \zeta_s(X_n, X), \qquad \eta := \limsup_{n\to\infty} \zeta_s(X_n, X)$$

and $\varepsilon > 0$ arbitrary. Then, there exists $\ell > 0$ with $\zeta_s(X_n, X) \leq \eta + \varepsilon$ for all $n \geq \ell$. Using (3.9), (3.10) and splitting $\{n_0 \leq I_r^{(n)} \leq n-1\}$ into $\{n_0 \leq I_r^{(n)} \leq \ell\}$ and $\{\ell < I_r^{(n)} \leq n-1\}$, we obtain

$$\zeta_s(X_n, X) \leq \frac{\bar{\eta}}{1-p_n} \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\left\{n_0 \leq I_r^{(n)} \leq \ell\right\}} \|A_r^{(n)}\|^s \right] + \frac{\eta+\varepsilon}{1-p_n} \mathbb{E}\left[\sum_{r=1}^{K} \|A_r^{(n)}\|^s \right] + o(1)$$

which, by **C1**, finally implies

$$\eta \leq \mathbb{E}\left[\sum_{r=1}^{K} \|A_r\|^s\right] (\eta + \varepsilon).$$

Since $\varepsilon > 0$ is arbitrary and by condition **C2**, we obtain $\eta = 0$. $\square$

Remark 3.3. As pointed out in [ER07] for a related convergence result, the statement of Proposition 3.2 remains true if condition **C1** is weakened by replacing

$$\sum_{r=1}^{K} \left\| A_r^{(n)} - A_r \right\|_s \to 0$$

by

$$\sum_{r=1}^{K} \left\| (A_r^{(n)} - A_r) f \right\|_s \to 0$$

for all $f \in \mathcal{C}[0,1]$ and uniform boundedness of $\|A_r^{(n)}\|_s$ for all $n \geq 0$ and all $r = 1, \ldots, K$. This follows from the given independence structure and the dominated convergence Theorem.

Remark 3.4. The methodology developed in the present section covers sequences (X_n) with jumps at random times. However, condition **C1** essentially requires these times to be equal for all $n \geq n_0$. In particular sequences of processes with jumps at random times that require a (uniformly small) time scale deformation cannot be treated by our approach.

To be able to deduce weak convergence in the situation of Proposition 3.2 for the special cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$, rates of convergence for ζ_s are required. We impose a further assumption on the convergence rate of the coefficients to establish a rate of convergence for the process that strengthens condition **C2**.

C4. The sequence $(\gamma(n))_{n \geq n_0}$ from condition **C1** satisfies $\gamma(n) = O(R(n))$ as $n \to \infty$ for some positive sequence $R(n) \downarrow 0$ such that

$$L^* = \limsup_{n \to \infty} \mathbb{E}\left[\sum_{r=1}^{K} \|A_r^{(n)}\|^s \frac{R(I_r^{(n)})}{R(n)}\right] < 1.$$

Corollary 3.5. *Let $(X_n)_{n \geq 0}$ satisfy recurrence (3.1) with conditions* **C1**, **C3** *and* **C4**. *Then for the fixed-point $\eta = \mathcal{L}(X)$ of T in (3.2) we have, as $n \to \infty$,*

$$\zeta_s(X_n, X) = O(R(n)).$$

Proof. We consider the quantities introduced in the proof of Proposition 3.2 again. By condition **C4** we have $\zeta_s(Q_n, X) \leq C R(n)$ for some $C > 0$ and all n. Furthermore, we can choose $\gamma > 0$ and $n_1 > 0$ such that

$$\mathbb{E}\left[\sum_{r=1}^{K} \|A_r^{(n)}\|^s \frac{R(I_r^{(n)})}{R(n)}\right] \leq 1 - \gamma, \qquad p_n \leq \frac{\gamma}{2}$$

for $n \geq n_1$. Obviously, for any $n_2 \geq n_1$, we can choose $K \geq 2C/\gamma$ such that $d(n) := \zeta_s(X_n, X) \leq KR(n)$ for all $n < n_2$. Using (3.10), this implies

$$d(n_2) \quad \leq \quad p_{n_2} d(n_2) + \mathbb{E}\left[\sum_{r=1}^{K} \mathbf{1}_{\left\{I_r^{(n_2)} \leq n_2 - 1\right\}} \|A_r^{(n_2)}\|^s d(I_r^{(n_2)})\right] + CR(n_2)$$

hence

$$\begin{aligned} d(n_2) &\leq \frac{1}{1-p_{n_2}}\left(\mathbb{E}\left[\sum_{r=1}^{K} \|A_r^{(n_2)}\|^s KR(I_r^{(n_2)})\right] + CR(n_2)\right) \\ &= \frac{1}{1-p_{n_2}}\left(KR(n_2)\mathbb{E}\left[\sum_{r=1}^{K} \|A_r^{(n_2)}\|^s \frac{R(I_r^{(n_2)})}{R(n_2)}\right] + CR(n_2)\right) \\ &\leq \frac{1}{1-p_{n_2}}((1-\gamma)K + C)R(n_2) \leq KR(n_2). \end{aligned}$$

Inductively, $d(n) \leq KR(n)$ for all n. □

We now consider the special cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$. Related to Corollary 2.36 we consider the following additional assumption, where the notations $\mathcal{C}_r[0,1]$ defined in (2.32) and $\mathcal{D}_r[0,1]$ defined in (2.33) are used:

C5 Case $(\mathcal{C}[0,1], \|\cdot\|)$: We have $X_n = Y_n + h_n$ for all $n \geq 0$, where $\|h_n - h\| \to 0$ with $h_n, h \in \mathcal{C}[0,1]$, and there exists a positive sequence $(r_n)_{n\geq 0}$ such that

$$\mathbf{P}\left(Y_n \notin \mathcal{C}_{r_n}[0,1]\right) \to 0.$$

Case $(\mathcal{D}[0,1], d_{sk})$: We have $X_n = Y_n + h_n$ for all $n \geq 0$, where $\|h_n - h\| \to 0$ with $h_n \in \mathcal{D}[0,1], h \in \mathcal{C}[0,1]$, and there exists a positive sequence $(r_n)_{n\geq 0}$ such that

$$\mathbf{P}\left(Y_n \notin \mathcal{D}_{r_n}[0,1]\right) \to 0.$$

We now state the main theorem of this section. It follows immediately from Proposition 2.34, Corollary 2.36, Proposition 3.2 and Corollary 3.5.

Theorem 3.6. *Let $(X_n)_{n\geq 0}$ be a sequence of random variables in $(\mathcal{C}[0,1], \|\cdot\|)$ or $(\mathcal{D}[0,1], d_{sk})$ satisfying recurrence (3.1) with conditions* **C1**, **C2**, **C3** *being satisfied. Then, for $\mathcal{L}(X) = \eta$ we have for all $t \in [0,1]$*

$$X_n(t) \xrightarrow{d} X(t), \quad \mathbb{E}\left[|X_n(t)|^s\right] \to \mathbb{E}\left[|X(t)|^s\right].$$

If Z is distributed on $[0,1]$ and independent of (X_n) and X then

$$X_n(Z) \xrightarrow{d} X(Z), \qquad \mathbb{E}\left[|X_n(Z)|^s\right] \to \mathbb{E}\left[|X(Z)|^s\right].$$

If moreover conditions **C4** *and* **C5** *are satisfied, where $R(n)$ in* **C4** *and r_n in* **C5** *can be chosen with*

$$R(n) = o\left(\frac{1}{\log^m(1/r_n)}\right), \quad n \to \infty, \tag{3.12}$$

and X has continuous sample paths, then we have convergence in distribution:

$$X_n \xrightarrow{d} X,$$

We comprehend our main result by convergence result for the moments of the norm based on Theorem 2.38.

Corollary 3.7. *Let $(X_n)_{n\geq 0}$ be a sequence of random variables in $(\mathcal{C}[0,1], \|\cdot\|)$ or $(\mathcal{D}[0,1], d_{sk})$ satisfying recurrence (3.1) with conditions* **C1**–**C5** *with $s \in \{1,2,3\}$ and such that also condition* (3.12) *is fulfilled. If, in the continuous case,*

$$\mathbb{E}\left[\|X_n\|^s \mathbf{1}_{\{Y_n \notin \mathcal{C}_{r_n}[0,1]\}}\right] \to 0,$$

where Y_n is defined in **C5***, then*

$$\mathbb{E}\left[\|X_n\|^s\right] \to \mathbb{E}\left[\|X\|^s\right] < \infty.$$

The result remains valid in the càdlàg case, if $\mathcal{C}_{r_n}[0,1]$ is replaced by $\mathcal{D}_{r_n}[0,1]$ and X has continuous sample paths.

Finally, we give sufficient criteria for the cases $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$ to verify condition **C3**. Let $\mathcal{L}(Y) = \nu$ be a probability distribution on $\mathcal{C}[0,1]$ with $\mathbb{E}\left[\|Y\|^s\right] < \infty$. Then for a probability measure $\mathcal{L}(X) = \mu$ on $\mathcal{C}[0,1]$ to be in $\mathcal{M}_s(\nu)$ we have the abstract defining properties in (2.3) and (2.4). Note that the cases $0 < s \leq 3$ are of interest in our main result, Theorem 3.6, and that $\mu \in \mathcal{M}_s(\nu)$ implies $\zeta_s(\mu,\nu) < \infty$.

3.2. The conditions on the moments

In this section, we give a precise characterization of conditions (2.3) and (2.4) in the case of continuous or càdlàg functions on the unit interval. Then we also discuss more general state spaces.

Lemma 3.8. *Let $\mathcal{L}(Y) = \mathcal{L}((Y_t)_{t\in[0,1]}) = \nu$ and $\mathcal{L}(X) = \mathcal{L}((X_t)_{t\in[0,1]}) = \mu$ be probability measures on $\mathcal{C}[0,1]$. For $0 < s \leq 1$ we have $\mu \in \mathcal{M}_s(\nu)$ if*

$$\mathbb{E}\left[\|X\|^s\right], \mathbb{E}\left[\|Y\|^s\right] < \infty. \tag{3.13}$$

For $1 < s \leq 2$ we obtain $\mu \in \mathcal{M}_s(\nu)$ if we have condition (3.13) and

$$\mathbb{E}\left[X_t\right] = \mathbb{E}\left[Y_t\right] \text{ for all } 0 \leq t \leq 1. \tag{3.14}$$

For $2 < s \leq 3$ we obtain $\mu \in \mathcal{M}_s(\nu)$ if we have conditions (3.13), (3.14) and

$$\mathrm{Cov}(X_t, X_u) = \mathrm{Cov}(Y_t, Y_u) \text{ for all } 0 \leq t, u \leq 1. \tag{3.15}$$

For $0 < s \leq 1$ or $1 < s \leq 2$ the assertions remain true if $\mathcal{C}[0,1]$ is replaced by $\mathcal{D}[0,1]$.

Proof. The case $0 < s \leq 1$ follows directly from the definition of the space $\mathcal{M}_s(\nu)$ for both, $\mathcal{C}[0,1]$ and $\mathcal{D}[0,1]$.
We first consider $B = \mathcal{C}[0,1]$ and start with the case $1 < s \leq 2$. By Riesz' representation theorem any linear and continuous function $\varphi : \mathcal{C}[0,1] \to \mathbb{R}$ can be written as

$$\varphi(f) = \int f(t) d\mu(t)$$

where μ is a finite, signed measure on $[0,1]$. Hence, (2.4) is satisfied if $\mathbb{E}[X_t] = \mathbb{E}[Y_t]$ for all $t \in [0,1]$ and (2.3) is condition (3.13).
We move on to the case $2 < s \leq 3$. By the Grothendieck inequality [Gro53], see also [Pis11] for a modern account, for any continuous bilinear form φ, there exist probability measures μ and ν on the unit interval such that

$$|\varphi(f,g)| \leq K\|\varphi\| \left[\int_0^1 f^2(t)d\mu(t) \int_0^1 g^2(s)d\nu(s)\right]^{1/2}.$$

for all $f, g \in \mathcal{C}[0,1]$. Here, K denotes a universal constant whose optimal value, called the Grothendieck constant, is still unknown. Thus, denoting $\eta = (\mu+\nu)/2$ the mixture of the two measures, it follows that φ is continuous on $(\mathcal{C}[0,1])^2$ when the space $\mathcal{C}[0,1]$ is endowed with the $L_2(\eta)$ topology. The set $\mathcal{C}[0,1]$ is dense in $L_2([0,1],\mathcal{B}([0,1]),\eta)$, hence we can extend φ to a continuous bilinear form on $L_2([0,1],\mathcal{B}([0,1]),\eta)$. Being a Hilbert space, the claim follows from Lemma 2.2 together with the Riesz representation theorem.
The description of the dual space of $\mathcal{D}[0,1]$ is slightly more complicated than in the case of $\mathcal{C}[0,1]$, in particular a continuous linear form on $\mathcal{D}[0,1]$ is not uniquely determined by its values on $\mathcal{C}[0,1]$. Pestman [Pes95, Theorem 1] showed that any linear and bounded map $\varphi : \mathcal{D}[0,1] \to \mathbb{R}$ is of the form

$$\varphi(f) = \int f(t)d\mu(t) + \sum_{x\in[0,1]} (f(x) - f(x-))\psi(x), \tag{3.16}$$

where μ is again a finite, signed measure on the unit interval, $f(x-) := \lim_{h\downarrow 0} f(x-h)$, $f(0-) := f(0)$ and $\psi : [0,1] \to \mathbb{R}$ takes values different from zero only on a countable subset F of $[0,1]$ with $\sum_{x\in F} |\psi(x)| < \infty$. Note that the measure μ comes from the restriction of φ to $\mathcal{C}[0,1]$. Furthermore, the representation of φ in terms of μ and ψ is unique. Equation (3.16) implies that $\mu \in \mathcal{M}_s(\nu)$ if $\mathbb{E}[X(t)] = \mathbb{E}[Y(t)]$ for all $t \in [0,1]$ and $\mathbb{E}[\|X\|^s], \mathbb{E}[\|Y\|^s] < \infty$ like in the continuous case. Note that $\mathbb{E}[X(t-)] = \mathbb{E}[Y(t-)]$ for all $t \in [0,1]$ follows from the latter by dominated convergence. □

Remark 3.9. Interpreting $\mathbb{E}[X]$ as Bochner-Integral in the continuous case, it is equivalent to say $\mathbb{E}[X(t)] = \mathbb{E}[Y(t)]$ for all $t \in [0,1]$ and $\mathbb{E}[X] = \mathbb{E}[Y]$. This is simply due to the fact that $\mathbb{E}[X]$ is a continuous function with $\mathbb{E}[X](t) = \mathbb{E}[X(t)]$ and $\varphi(\mathbb{E}[X]) = \mathbb{E}[\varphi(X)]$ for all continuous linear forms φ on $\mathcal{C}[0,1]$. Also the higher moments can be interpreted similarly as expectations of tensor products, cf. [DJN08].

Remark 3.10. Note that condition (3.15) typically cannot be achieved for a sequence $(X_n)_{n\geq 0}$ that arises as in (1.2) by an affine scaling from a sequence $(Y_n)_{n\geq 0}$ as in (1.1). This fundamental problem for developing a functional contraction method on the basis of the Zolotarev metrics ζ_s with $2 < s \leq 3$ was already mentioned in [DJN08, Remark 6.2]. We describe a way to circumvent this problem in our application to Donsker's invariance principle by a perturbation argument, see Section 4.

Remark 3.11. Lemma 3.8 implies that condition (2.4) may be replaced by (2.5) in the case of $\mathcal{C}[0,1]$ for $k = 1, 2$ or $\mathcal{D}[0,1]$ for $k = 1$. In fact, much more can be said. Janson and Kaijser [JK] show that the equivalence of (2.4) and (2.5) holds true for any $k \in \mathbb{N}$ in separable Banach spaces having the approximation property such as $\mathcal{C}[0,1]$ or sequence spaces. In fact, it had been an open problem to find Banach spaces without this property for many years, the first example was given by Enflo in [Enf73]. Based on a Banach space without the approximation property, Janson and Kaijser [JK] also give an example where the equivalence of (2.4) and (2.5) is false already in the case $k = 2$.

4. Donsker's invariance principle

Let $(V_n)_{n\in\mathbb{N}}$ be a sequence of independent, identically distributed real-valued random variables with $\mathbb{E}[V_1] = 0$ and $\mathbf{Var}[V_1] = 1$. In Donsker's theorem one considers the properly scaled and linearly interpolated random walk $S^n = (S^n_t)_{t\in[0,1]}$, $n \geq 1$, defined by

$$S^n_t = \frac{1}{\sqrt{n}}\left(\sum_{k=1}^{\lfloor nt\rfloor} V_k + (nt - \lfloor nt\rfloor)V_{\lfloor nt\rfloor+1}\right), \quad t \in [0,1].$$

With $W = (W_t)_{t\in[0,1]}$ a standard Brownian motion, Donsker's functional limit law states that $S^n \to W$ in distribution on $(\mathcal{C}[0,1], \|\cdot\|)$. Equivalently and more in the spirit of the time when the result was formulated and proved, this means

$$f(S^n) \xrightarrow{d} f(W) \tag{4.1}$$

for any continuous function $f : \mathcal{C}[0,1] \to \mathbb{R}$.

The history started with the idea of Erdős and Kac [Kac46, EK46, EK47] to prove invariance principles for $f(S_n)$ by two steps: First, one provides distributional convergence of $f(S^n)$ and notes its limit to be invariant under the law of V_1. Second, one determines the shape of the limit by focussing on a convenient choice of $\mathcal{L}(V_1)$ that allows one to compute the limit by means of direct calculations. Applying this methodology, Erdős and Kac established (4.1) for certain functions f, e.g. $f(x) = \sup_{t\in[0,1]} x(t)$ and $f(x) = \sup_{t\in[0,1]} |x(t)|$. A much earlier work by Kolmogorov [Kol31] had already been in this spirit. In the works by Mark [Mar49] and Fortet [For49] the idea of Erdős and Kac was extended to various other continuous functionals. The heuristic approach of directly approximating the sequence of processes by its limit goes back to Doob [Doo49], where he uses this idea in the related case of the rescaled empirical distribution function and its limit, the Brownian bridge. As an outcome of his dissertation, Donsker [Don51] gave a rigorous proof of (4.1) for all continuous functions f. The concept of tightness was developed shortly after by Prokhorov [Pro53, Pro56] and the proofs of the invariance principle found in most textbooks involve his arguments based on the theorem that is today named after him.

For the purpose of the contraction method it is necessary to assume an additional moment on V_1. Our aim of the next section is to prove the following theorem.

Theorem 4.1. *Let* $\mathbb{E}\left[|V_1|^{2+\varepsilon}\right] < \infty$ *for some* $\varepsilon > 0$. *Then* $S^n \xrightarrow{d} W$ *as* $n \to \infty$ *in* $(\mathcal{C}[0,1], \|\cdot\|)$.

4.1. A contraction proof

In this section we apply the general methodology of the Chapters 2 and 3 to give a short proof of Theorem 4.1. For a recursive decomposition of S^n and W we define operators for $\beta > 1$,

$$\varphi_\beta : \mathcal{C}[0,1] \to \mathcal{C}[0,1], \quad \varphi_\beta(f)(t) = \mathbf{1}_{\{t \le 1/\beta\}} f(\beta t) + \mathbf{1}_{\{t > 1/\beta\}} f(1),$$
$$\psi_\beta : \mathcal{C}[0,1] \to \mathcal{C}[0,1], \quad \psi_\beta(f)(t) = \mathbf{1}_{\{t \le 1/\beta\}} f(0) + \mathbf{1}_{\{t > 1/\beta\}} f\left(\tfrac{\beta t - 1}{\beta - 1}\right).$$

Note that both φ_β and ψ_β are linear, continuous and $\|\varphi_\beta(f)\| = \|\psi_\beta(f)\| = \|f\|$ for all $f \in \mathcal{C}[0,1]$, hence we have $\|\varphi_\beta\| = \|\psi_\beta\| = 1$. By construction we have

$$S^n \stackrel{d}{=} \sqrt{\frac{\lceil n/2 \rceil}{n}}\varphi_{\frac{n}{\lceil n/2 \rceil}}\left(S^{\lceil n/2 \rceil}\right) + \sqrt{\frac{\lfloor n/2 \rfloor}{n}}\psi_{\frac{n}{\lceil n/2 \rceil}}\left(\widehat{S}^{\lfloor n/2 \rfloor}\right), \quad n \ge 2, \tag{4.2}$$

where $\stackrel{d}{=}$ denotes equality in distribution, $(S^1, \dots, S^n)$ and $(\widehat{S}^1, \dots, \widehat{S}^n)$ are independent and S^j and $\widehat{S}^j$ are identically distributed for all $j \ge 1$. Therefore $(S^n)_{n \ge 1}$ satisfies recurrence (3.1) choosing

$$K = 2, \quad I_1^{(n)} = \lceil n/2 \rceil, \quad I_2^{(n)} = \lfloor n/2 \rfloor, \quad n_0 = 2,$$
$$A_1^{(n)} = \sqrt{\frac{\lceil n/2 \rceil}{n}}\varphi_{\frac{n}{\lceil n/2 \rceil}}, \quad A_2^{(n)} = \sqrt{\frac{\lfloor n/2 \rfloor}{n}}\psi_{\frac{n}{\lceil n/2 \rceil}}, \quad b^{(n)} = 0.$$

Moreover, as $\varphi_\beta, \psi_\beta$ are deterministic, the operators $A_1^{(n)}$ and $A_2^{(n)}$ are random linear operators in the sense of Definition 2.7; the same holds also for their limits as defined in (4.4). In the following let $\widehat{W} = (\widehat{W}_t)_{t \in [0,1]}$ be a standard Brownian motion, independent of W. Properties of Brownian motion imply

$$W \stackrel{d}{=} \sqrt{\frac{1}{\beta}}\varphi_\beta(W) + \sqrt{\frac{\beta - 1}{\beta}}\psi_\beta(\widehat{W}), \tag{4.3}$$

for any $\beta > 1$. Hence, the Wiener measure $\mathcal{L}(W)$ is a fixed-point of the operator T in (3.2) with

$$K = 2, A_1 = \sqrt{\frac{1}{\beta}}\varphi_\beta, A_2 = \sqrt{\frac{\beta - 1}{\beta}}\psi_\beta, b = 0, \tag{4.4}$$

an illustration thereof is given in the figures 4.1 and 4.2. For $\beta = 2$ the coefficients in (4.2) converge to the ones in (4.3), i.e., as $n \to \infty$,

$$\sqrt{\frac{\lceil n/2 \rceil}{n}} \to \frac{1}{\sqrt{2}}, \quad \sqrt{\frac{\lfloor n/2 \rfloor}{n}} \to \frac{1}{\sqrt{2}},$$

but the coefficients $A_1^{(n)}, A_2^{(n)}$ only converge to A_1, A_2 in the operator norm for n even. Nevertheless, from the point of view of the contraction method this suggests weak convergence of S^n to W.

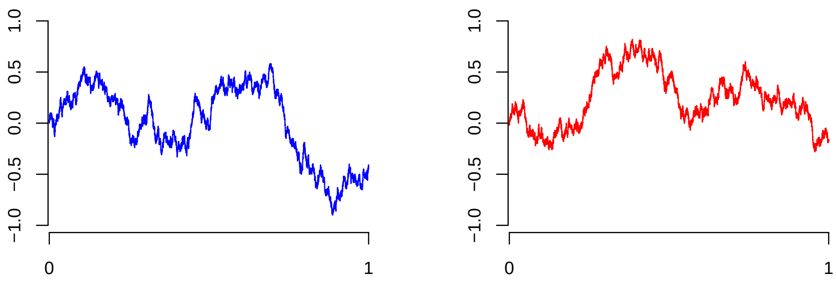


Figure 4.1.: Realizations of independent Brownian motions.

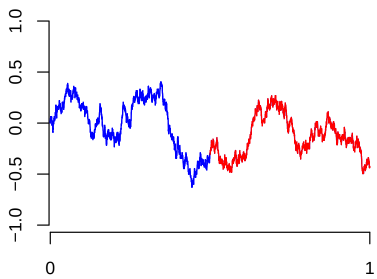


Figure 4.2.: The concatenation in the sense of (4.3) for $\beta = 2$.

The operator T associated with the fixed-point equation (4.3), i.e., with the coefficients in (4.4), satisfies condition **C2** only with $s > 2$. In view of condition **C3** and Lemma 3.8 we need to match the mean and covariance structure. We have $\mathbb{E}\left[S_t^n\right] = 0$ for all $0 \le t \le 1$ and a direct computation yields

$$\operatorname{Cov}(S_s^n, S_t^n) = \begin{cases} s, & \text{for } \lfloor ns \rfloor < \lfloor nt \rfloor, \\ \frac{1}{n}\Big(\lfloor ns \rfloor + (ns - \lfloor ns \rfloor)(nt - \lfloor nt \rfloor)\Big), & \text{for } \lfloor ns \rfloor = \lfloor nt \rfloor. \end{cases} \tag{4.5}$$

Hence, we do not have finite $\zeta_{2+\varepsilon}$ -distance between S^n and W since they do not share their covariance functions. To surmount this problem we consider a suitable linear interpolation of the Brownian motion W. For fixed $n \in \mathbb{N}$ we divide the unit interval into pieces of length $1/n$ and interpolate W linearly between the points $0, 1/n, 2/n, \dots, (n-1)/n, 1$. The interpolated process $W^n = (W_t^n)_{t\in[0,1]}$ is given by

$$W_t^n := W_{\frac{\lfloor nt \rfloor}{n}} + (nt - \lfloor nt \rfloor)\left(W_{\frac{\lfloor nt \rfloor + 1}{n}} - W_{\frac{\lfloor nt \rfloor}{n}}\right), \quad t \in [0,1].$$

We have $\mathbb{E}\left[W_t^n\right] = 0$ and W^n and S^n have the same covariance function (4.5) for all $n \in \mathbb{N}$. Furthermore W^n has the same distributional recursive decomposition (4.2) as S^n.

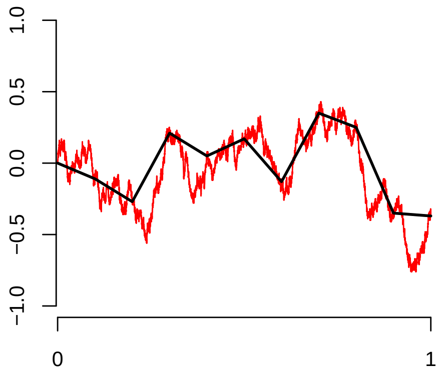


Figure 4.3.: A Brownian motion W and its linearly interpolated version W^n for $n = 10$.

Note that the linearly interpolated version does not differ much from the original one:

Lemma 4.2. *We have* $\|W^n - W\| \to 0$ *as* $n \to \infty$ *almost surely.*

Proof. This directly follows from the uniform continuity of W. For $\varepsilon > 0$ there exists a random $\delta > 0$ such that $|W(t) - W(s)| < \varepsilon$ for any $s, t \in [0,1]$ with $|t - s| < \delta$. An adaption of the triangle inequality gives $\|W^n - W\| < 2\varepsilon$ for any $n > \delta^{-1}$. □

In view of Corollary 2.37 it suffices to prove that S^n and W^n are close with respect to $\zeta_{2+\varepsilon}$. The proof of this runs along the same lines as the one for Proposition 3.2, resp. Corollary 3.5, in fact it is much shorter due to the simple form of the recurrence:

Proposition 4.3. *For any* $\delta < \varepsilon/2$ *we have* $\zeta_{2+\varepsilon}(S^n, W^n) = O(n^{-\delta})$ *as* $n \to \infty$.

Proof. We have

$$\begin{aligned}
\zeta_{2+\varepsilon}(S^n, W^n) &= \zeta_{2+\varepsilon}\Bigg(\sqrt{\frac{\lceil n/2\rceil}{n}}\varphi_{\frac{n}{\lceil n/2\rceil}}\left(S^{\lceil n/2\rceil}\right) + \sqrt{\frac{\lfloor n/2\rfloor}{n}}\psi_{\frac{n}{\lceil n/2\rceil}}\left(\overline{S}^{\lfloor n/2\rfloor}\right), \\
&\qquad \sqrt{\frac{\lceil n/2\rceil}{n}}\varphi_{\frac{n}{\lceil n/2\rceil}}\left(W^{\lceil n/2\rceil}\right) + \sqrt{\frac{\lfloor n/2\rfloor}{n}}\psi_{\frac{n}{\lceil n/2\rceil}}\left(\overline{W}^{\lfloor n/2\rfloor}\right)\Bigg) \\
&\leq \left(\frac{\lceil n/2\rceil}{n}\right)^{1+\varepsilon/2}\zeta_{2+\varepsilon}\left(S^{\lceil n/2\rceil}, W^{\lceil n/2\rceil}\right) \\
&\quad + \left(\frac{\lfloor n/2\rfloor}{n}\right)^{1+\varepsilon/2}\zeta_{2+\varepsilon}\left(S^{\lfloor n/2\rfloor}, W^{\lfloor n/2\rfloor}\right).
\end{aligned}$$

We abbreviate

$$d_n := \zeta_{2+\varepsilon}(S^n, W^n), \quad a_n := \left(\frac{\lceil n/2\rceil}{n}\right)^{1+\varepsilon/2}, \quad b_n := \left(\frac{\lfloor n/2\rfloor}{n}\right)^{1+\varepsilon/2}$$

and note that we have $a_n + b_n \leq 2^{-\varepsilon/2} + C'/n$ for some constant $C' > 0$ and all $n \in \mathbb{N}$. For arbitrary $\delta < \varepsilon/2$ we prove the assertion by induction: Fix $\delta < \delta' < \varepsilon/2$ and choose $m_0 \in \mathbb{N}$ such that $\lfloor n/2\rfloor^{-\delta} \leq (n/2)^{-\delta}2^{\varepsilon/2-\delta'}$ and $1 + 2^{\varepsilon/2}C'/n \leq 2^{\delta'-\delta}$ for all $n \geq m_0$. Furthermore, let

$C > 0$ be large enough such that $d_n \leq Cn^{-\delta}$ for all $1 \leq n \leq m_0$. Then, for $n > m_0$, assuming the claim to be verified for all smaller indices,

$$\begin{aligned} d_n &\leq a_n d_{\lceil n/2 \rceil} + b_n d_{\lfloor n/2 \rfloor} \\ &\leq C\left(a_n (n/2)^{-\delta} + b_n (n/2)^{-\delta} 2^{\varepsilon/2-\delta'}\right) \\ &\leq Cn^{-\delta} 2^{\delta} 2^{\varepsilon/2-\delta'}(a_n + b_n) \\ &\leq Cn^{-\delta}. \end{aligned}$$

The assertion follows. □

Now Donsker's Theorem (Theorem 4.1) follows from Proposition 4.3, Lemma 4.2 and Corollary 2.37. Observe that we could have worked analogously in the framework of càdlàg functions by choosing a constant interpolation between successive points of type i/n.

By Theorem 2.38 and Proposition 2.39 we directly obtain convergence of moments of the supremum if we assume additional moments for the increments.

Corollary 4.4. *Suppose $\mathbb{E}\left[|V_1|^k\right] < \infty$ for an integer $k \geq 3$. Then the first k absolute moments of $\frac{1}{\sqrt{n}} \sup_{0 \leq k \leq n} S_k$ converge to the corresponding moments of $\|W\|$.*

4.2. Characterizing the Wiener measure by a fixed-point property

We reconsider the map T corresponding to the fixed-point equation (4.3) for the case $\beta = 2$:

$$\begin{aligned} &T : \mathcal{M}(\mathcal{C}[0,1]) \to \mathcal{M}(\mathcal{C}[0,1]) \\ &T(\mu) = \mathcal{L}\left(\frac{1}{\sqrt{2}}\varphi_2(Z) + \frac{1}{\sqrt{2}}\psi_2(\overline{Z})\right), \end{aligned} \tag{4.6}$$

where $Z, \overline{Z}$ are independent with distribution $\mathcal{L}(Z) = \mathcal{L}(\overline{Z}) = \mu$. Our discussion above implies that the Wiener measure $\mathcal{L}(W)$ is the unique fixed-point of T restricted to $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ for any $\varepsilon > 0$. Note that $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ is the space of the distributions of all continuous stochastic processes $V = (V_t)_{t \in [0,1]}$ with $\mathbb{E}\left[\|V\|^{2+\varepsilon}\right] < \infty$, $\mathbb{E}[V_t] = 0$ and $\mathrm{Cov}(V_t, V_u) = t \wedge u$ for all $0 \leq t, u \leq 1$. Note that one easily verifies that $T(\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))) \subset \mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ and the last part of the proof of Lemma 3.1 implies that T restricted to $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$ is Lipschitz-continuous with Lipschitz constant at most $L = 2^{-\varepsilon/2} < 1$, hence $\mathcal{L}(W)$ is the unique fixed-point of T in $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$.
We now show that a more general statement is true, the Wiener measure is also, up to multiplicative scaling, the unique fixed-point of T in the larger space $\mathcal{M}(\mathcal{C}[0,1])$. For a related statement, see also Aldous [Ald94, page 528]. The subsequent proof is based on the fact that the centered normal distributions are the only solutions of the fixed-point equation

$$X \stackrel{d}{=} \frac{X + \overline{X}}{\sqrt{2}} \tag{4.7}$$

where $X, \overline{X}$ are independent, identically distributed real-valued random variables as already discussed on page 5 in the introduction.

Theorem 4.5. *Let $X = (X_t)_{t\in[0,1]}$ be a continuous process with $X_0 = 0$. Then $\mathcal{L}(X)$ is a fixed-point of (4.6) if and only if either $X = \mathbf{0}$ a.s. or there exists a constant $\sigma > 0$, such that $(\sigma^{-1}X_t)_{t\in[0,1]}$ is a standard Brownian motion.*

Proof. Let $\mathcal{L}(X)$ be a fixed-point of (4.6) and $\overline{X} = (\overline{X}_t)_{t\in[0,1]}$ be independent of X with the same distribution. The fixed-point property implies

$$X_1 \stackrel{d}{=} \frac{X_1 + \overline{X}_1}{\sqrt{2}},$$

hence $\mathcal{L}(X_1) = \mathcal{N}(0, \sigma^2)$ for some $\sigma^2 \geq 0$, where $\mathcal{N}(0, \sigma^2)$ denotes the centered normal distribution with variance σ^2. This implies

$$X_{1/2} \stackrel{d}{=} \frac{X_1}{\sqrt{2}},$$

hence $\mathcal{L}(X_{1/2}) = \mathcal{N}(0, \sigma^2/2)$. Let $\mathscr{D} = \{m2^{-n} : m, n \in \mathbb{N}_0, m \leq 2^n\}$ by the set of dyadic numbers in $[0, 1]$. By induction, we obtain $\mathcal{L}(X_t) = \mathcal{N}(0, \sigma^2 t)$ for all $t \in \mathscr{D}$. For the distribution of the increments we first obtain

$$X_1 - X_{1/2} \stackrel{d}{=} \frac{X_1}{\sqrt{2}},$$

hence $\mathcal{L}(X_1 - X_{1/2}) = \mathcal{N}(0, \sigma^2/2)$. Again inductively, we obtain $\mathcal{L}(X_1 - X_t) = \mathcal{N}(0, (1-t)\sigma^2)$ for all $t \in \mathscr{D}$. Also by induction, it follows $\mathcal{L}(X_t - X_s) = \mathcal{N}(0, (t-s)\sigma^2)$ for all $s, t \in \mathscr{D}$ with $s < t$. Finally, continuity of X implies the same property for all $s, t \in [0, 1]$. It remains to prove independence of increments. Denoting by $X^{(1)}, X^{(2)}, \ldots$ independent distributional copies of X, we obtain from iterating the fixed-point property

$$(X_t)_{t\in[0,1]} \stackrel{d}{=} \left(2^{-n/2} \sum_{m=1}^{2^n} \mathbf{1}_{\{(m-1)2^{-n} < t \leq m2^{-n}\}} X^{(m)}_{2^n t - m + 1} + \mathbf{1}_{\{m2^{-n} < t\}} X^{(m)}_1\right)_{t\in[0,1]}$$

for all $n \in \mathbb{N}$. Hence, for any dyadic points $0 \leq t_1 < t_2 < \ldots < t_k \leq 1$, choosing n large enough, each $X_{t_{i+1}} - X_{t_i}$ can be expressed as a function of a subset of $X^{(1)}, \ldots, X^{(2^n)}$ these subsets being pairwise disjoint for $i = 0, \ldots, n-1$. Since, $\mathscr{D}$ is dense in $[0, 1]$, this shows that X has independent increments. For $\sigma = 0$ we have $X = \mathbf{0}$ a.s., otherwise $\sigma^{-1}X$ is a standard Brownian motion.
The converse direction of the Theorem is trivial. □

Remark 4.6. Note that we cannot cancel the assumption on continuity of X without replacement, e.g., the process

$$Y_t = \begin{cases} W_t & : \quad t \notin \mathscr{D} \\ 0 & : \quad t \in \mathscr{D} \end{cases}$$

also solves (4.3) and is not a multiple of Brownian motion. However, it would be sufficient to require càdlàg paths, so $\mathcal{C}[0, 1]$ could be replaced by $\mathcal{D}[0, 1]$ in our statement.

Remark 4.7. Our decomposition of Brownian motion in (4.3) is in time. However, equation (4.7) suggests to also investigate a decomposition in space

$$(X_t)_{t\in[0,1]} \stackrel{d}{=} \left(\frac{X_t + \overline{X}_t}{\sqrt{2}}\right)_{t\in[0,1]} \tag{4.8}$$

where $(X_t)_{t\in[0,1]}$ and $(\overline{X}_t)_{t\in[0,1]}$ are independent and identically distributed. Again, equation (4.8) induces a map on $\mathcal{M}(\mathcal{C}[0,1])$ that is a contraction in $\zeta_{2+\varepsilon}$ on the subspace $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$, so the Wiener measure is the only solution in $\mathcal{M}_{2+\varepsilon}(\mathcal{L}(W))$. In this case, we cannot remove the moment assumption as in Theorem 4.5 since any centered, continuous Gaussian process solves equation (4.8). Using (4.7), it is not hard to see that there are no further solutions of (4.8).

5. Analysis of partial match queries

To begin the chapter on partial match retrieval in random quadtrees, let us quickly recap the relevant terms from the introduction. $C_n(s)$ is the number of visited nodes of the retrieval algorithm in the quadtree searching for all items whose first component equals s. Equivalently, we have seen that the quantity coincides with the number of horizontal lines in the partition of the unit square given by the points building up the tree which intersect the vertical line at s. We remember the additive recursion (1.18) that is satisfied by $(C_n(s))_{s\in[0,1]}$ on the level of càdlàg functions.

$$\begin{aligned} C_n(s) \quad &\stackrel{d}{=} \quad 1 + \mathbf{1}_{\{s<U\}} \left[C^{(1)}_{I_1^{(n)}}\left(\frac{s}{U}\right) + C^{(2)}_{I_2^{(n)}}\left(\frac{s}{U}\right)\right] \\ &\qquad + \mathbf{1}_{\{s\geq U\}} \left[C^{(3)}_{I_3^{(n)}}\left(\frac{1-s}{1-U}\right) + C^{(4)}_{I_4^{(n)}}\left(\frac{1-s}{1-U}\right)\right]. \end{aligned} \tag{5.1}$$

Here, (U, V) denote the components of the first inserted point, $I_1^{(n)}, \dots, I_4^{(n)}$ denote the number of points in the subregions and $(C_n^{(1)}), \dots, (C_n^{(4)})$ are independent copies of (C_n), independent of $\left(U, V, I_1^{(n)}, \dots, I_4^{(n)}\right)$. The distribution of the number of points in the subquadrants is

$$\mathcal{L}\left(I_1^{(n)}, \dots, I_4^{(n)}\right) = \text{Mult}(n-1; UV, U(1-V), (1-U)V, (1-U)(1-V)). \tag{5.2}$$

In their analysis of the complexity of partial match retrieval, or as they call it *regionsearch*, Bentley and Stanat [BS75] use the idealized approach of perfect quadtrees in which all subtrees have the same number of nodes. Stochastically, this basically coincides with the assumption that the proportion of nodes in each of the four subtrees converges to $1/4$ as the tree size grows to infinity. By means of the distribution of the split random variable $I^{(n)}$ given in (5.2), we can immediately discard this idea. Comparing their theoretical result with experimental data, the authors observe that their approximation by a term of order $\sqrt{n}$ underestimates the actual costs and give two reasons for this phenomenon. First, based on arguments from [FB74] on the path length of the tree, the number of visited nodes in the random quadtree is larger than in its idealized approximation. Second, they point out that the partitioning of the search space is not well-balanced, or as they call it "checkerboarding" [the distinction between these points is questionable as the first is a result of the second]. However, they neglect the influence of the second point and emphasize the first.
In fact, the results by Flajolet et al. [FGPR93] reveal that it is just the other way around. First, the path length is of the same asymptotic order as for perfect trees, i. e. $\Theta(n \log n)$ and second, higher order asymptotics for the costs are caused by the non-balanced partitioning of the state space. We aim at giving a short heuristic argument for this here. Let ξ be uniform on the unit interval, independent of the quadtree. The relative position of the line at ξ in the both relevant subquadrants that appear by adding the first point in the unit square is again uniform at random. The width of these regions is distributed like a uniform random variable on $[0, 1]$ conditioned to be covered by ξ which gives rise to a size-biased distribution. This implies

$$C_n(\xi) \quad \stackrel{d}{=} \quad 1 + C^{(1)}_{J_1^{(n)}}(\xi) + C^{(2)}_{J_2^{(n)}}(\xi). \tag{5.3}$$

Here, $J_1^{(n)}, J_2^{(n)}$ denote the number of points in the relevant subregions and $(C_n^{(1)}), C_n^{(2)})$ are independent copies of (C_n), independent of $\left(\xi, J_1^{(n)}, J_2^{(n)}\right)$. Moreover ξ and $(J_1^{(n)}, J_2^{(n)})$ are independent. The area of one of the two regions is distributed like a product of a uniform Y (for the height) and an independent size-biased uniform $\sqrt{X}$ (for the width). We thus have $\mathcal{L}(J_1^{(n)}) = \mathrm{Bin}(n-1, \sqrt{X}Y)$ where X, Y are uniform on the unit interval and the same holds for $J_2^{(n)}$. Taking

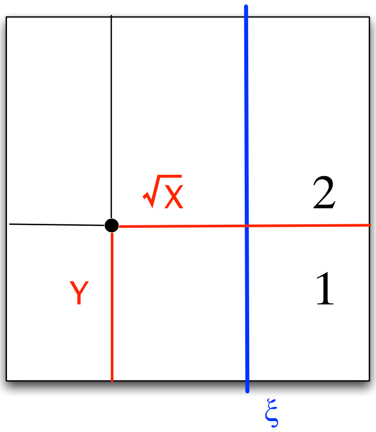


Figure 5.1.: The first split in the quadtree with uniform query line ξ.

expectations in (5.3) and multiplying by $n^{-\gamma}$ with $\gamma > 0$ yields

$$n^{-\gamma}\mathbb{E}\left[C_n(\xi)\right] = n^{-\gamma} + 2\mathbb{E}\left[\left(\frac{J_1^{(n)}}{n}\right)^{\gamma} \frac{C^{(1)}_{J_1^{(n)}}(\xi)}{(J_1^{(n)})^{\gamma}}\right].$$

Assuming that the left hand side of the latter expression converges as $n \to \infty$ and pulling apart the expectation of the product in a non-rigorous way gives

$$1 = \lim_{n\to\infty} 2\mathbb{E}\left[\left(\frac{J_1^{(n)}}{n}\right)^{\gamma}\right] = 2\mathbb{E}\left[(\sqrt{X}Y)^{\gamma}\right] = \frac{4}{(\gamma+1)(\gamma+2)}.$$

Thus, γ has to equal β as defined in (1.15). Note that the same heuristic approach may be applied for the mean of partial match queries in dimension d when s components are fixed, explaining (1.13).

In the first section, we collect all results on the asymptotic behaviour of the sequence $C_n(s)$ we can deduce by applying the contraction method as developed in the previous section. Subsequently, for the remaining of this chapter, we abbreviate that additive recurrences such as (5.1) or fixed-point equations such as (1.20) involving a parameter $s \in [0, 1]$ are to be understood on the level of càdlàg or continuous functions unless stated otherwise.

5.1. Main results and implications

Our main contribution is the following theorem whose proof is an application of Theorem 3.6.

Theorem 5.1. *Let $C_n(s)$ be the cost of a partial match query at a fixed line s in a random two-dimensional quadtree. Then, there exists a random continuous function Z such that, as $n \to \infty$,*

$$\left(\frac{C_n(s)}{K_1 n^\beta}\right)_{s\in[0,1]} \xrightarrow{d} (Z(s))_{s\in[0,1]}, \tag{5.4}$$

where

$$K_1 = \frac{\Gamma(2\beta+2)\Gamma(\beta+2)}{2\Gamma^3(\beta+1)\Gamma^2(\beta/2+1)}$$

is the constant appearing in (1.16). *This convergence in distribution holds in the space $(\mathcal{D}[0,1], d_{sk})$. The distribution of the random function Z is a fixed point of the following functional fixed-point equation*

$$\begin{aligned} Z(s) \stackrel{d}{=} &\mathbf{1}_{\{s<U\}}\left[(UV)^\beta Z^{(1)}\left(\frac{s}{U}\right) + (U(1-V))^\beta Z^{(2)}\left(\frac{s}{U}\right)\right] \\ &+ \mathbf{1}_{\{s\geq U\}}\left[((1-U)V)^\beta Z^{(3)}\left(\frac{s-U}{1-U}\right) + ((1-U)(1-V))^\beta Z^{(4)}\left(\frac{s-U}{1-U}\right)\right], \end{aligned} \tag{5.5}$$

where U and V are independent $[0,1]$-uniform random variables and $Z^{(i)}$, $i = 1,\dots,4$ are independent copies of the process Z, which are also independent of U and V. Furthermore, Z in (5.4) *is the only solution of* (5.5) *such that $\mathbb{E}\left[\|Z\|^2\right] < \infty$ and*

$$\mathbb{E}\left[Z(\xi)\right] = \mathrm{B}(\beta/2+1, \beta/2+1)$$

where ξ is uniformly distributed on the unit interval and $\mathrm{B}(x,y) := \Gamma(x)\Gamma(y)/\Gamma(x+y), x, y > 0$ denotes the Beta function. Additionally, all moments of $\|Z\|$ are finite.

For a simulation of a quadtree with corresponding process rescaled process $C_n(s)$ see figure (5.2).

It turns out that we will make use of the ζ_s metric for $s = 2$; thus, our approach is strong enough to guarantee convergence of the variance of the costs of partial match queries. This settles the open question on the order of the variance for uniform queries.

Theorem 5.2. *Let ξ be uniformly distributed on $[0,1]$, independent of (C_n) and Z, then*

$$\frac{C_n(\xi)}{K_1 n^\beta} \xrightarrow{d} Z(\xi),$$

in distribution with convergence of all moments. In particular

$$\mathbf{Var}\left(C_n(\xi)\right) \sim K_4 n^{2\beta}$$

where

$$\begin{aligned} K_4 &:= K_1^2 \cdot \mathbf{Var}\left[Z(\xi)\right] \\ &= K_1^2\left[\frac{2(2\beta+1)}{3(1-\beta)}(\mathrm{B}(\beta+1,\beta+1))^2 - \left(\mathrm{B}\left(\frac{\beta}{2}+1, \frac{\beta}{2}+1\right)\right)^2\right]. \end{aligned}$$

Numerically,

$$K_4 \approx 0.447363034.$$

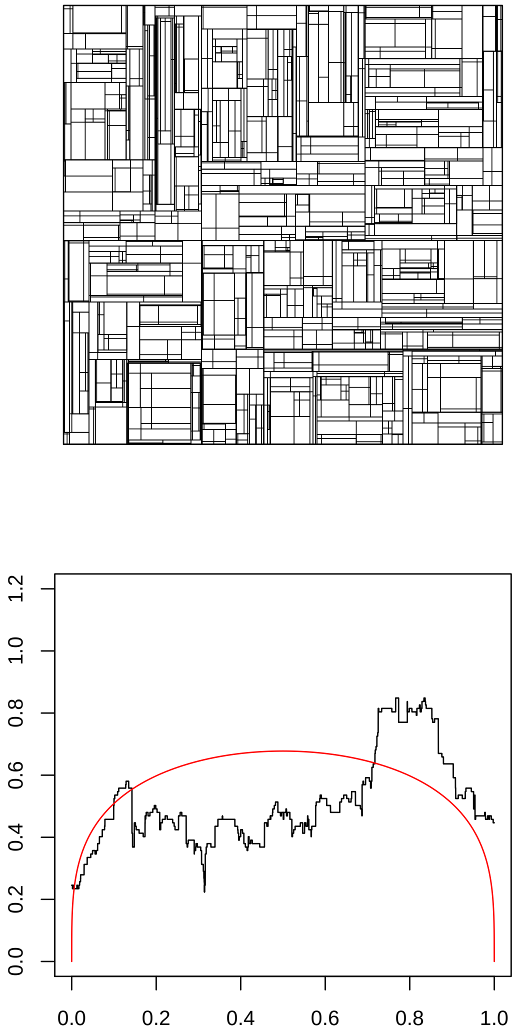


Figure 5.2.: Quadtree for $n = 500$. The lower figure shows $(K_1^{-1} n^{-\beta} C_n(s))_{s\in[0,1]}$ and the limit mean.

Concerning the worst-case behaviour, an adaption of Theorem 3.7 and Proposition 2.39 reveals all moments of the supremum to be of the same order as for fixed query lines.

Theorem 5.3. *Let $S_n = \sup_{s\in[0,1]} C_n(s)$. Then, as $n \to \infty$,*

$$\frac{S_n}{K_1 n^\beta} \to S := \sup_{s\in[0,1]} Z(s)$$

in distribution and with convergence of all moments. In particular, $\mathbb{E}[S_n] \sim n^\beta \mathbb{E}[S]$ and $\mathbf{Var}[S_n] \sim n^{2\beta}\mathbf{Var}[S]$. The random variable S satisfies stochastically

$$S \leq \left((UV)^\beta S^{(1)} + (U(1-V))^\beta S^{(2)}\right) \vee \left(((1-U)V)^\beta S^{(3)} + ((1-U)(1-V))^\beta S^{(4)}\right), \quad (5.6)$$

where U and V are independent $[0,1]$-uniform random variables and $S^{(i)}$, $i = 1, \dots, 4$ are independent copies of S, which are also independent of U and V.

The leading constants in the expansion for the mean and the variance for the supremum, that is $\mathbb{E}[S]$ and $\mathbf{Var}[S]$, remain open. The L_p-boundedness of $n^{-\beta}S_n$ implies the corresponding property for $n^{-\beta}C_n(s)$ for fixed s, hence it implies convergence of all moments as stated in the following theorem.

Theorem 5.4. *For all $s \in [0,1]$, we have*

$$\mathbb{E}\left[\left(\frac{C_n(s)}{K_1 n^\beta}\right)^m\right] \to \mathbb{E}\left[Z(s)^m\right],$$

for all $m \in \mathbb{N}$ as $n \to \infty$ where c_m is given in (5.38). *Moreover, for any natural number $\ell > 0$, positions $0 \leq s_1 < \ldots < s_\ell \leq 1$, and $k_1, \ldots, k_\ell \in \mathbb{N}$ one has*

$$\mathbb{E}\left[C_n^{k_1}(s_1)\cdots C_n^{k_\ell}(s_\ell)\right] \sim (K_1 n^\beta)^{\sum_{j=1}^{\ell} k_j} \cdot \mathbb{E}\left[Z^{k_1}(s_1)\cdots Z^{k_\ell}(s_\ell)\right].$$

Finally we note that the one-dimension marginals of the limit process $(Z(s), s \in [0,1])$ are all the same up to a multiplicative constant.

Theorem 5.5. *There is a random variable $Z \geq 0$ such that for all $s \in [0,1]$,*

$$Z(s) \stackrel{d}{=} (s(1-s))^{\beta/2} Z.$$

Z be the unique solution of the fixed-point equation

$$Z \stackrel{d}{=} U^{\beta/2} V^\beta Z + U^{\beta/2}(1-V)^\beta Z',$$

with $\mathbb{E}[Z] = 1$ and $\mathbb{E}\left[Z^2\right] < \infty$, where Z' is an independent copy of Z and (Z, Z') is independent of (U, V).

We immediately proceed to the proof of our main result as an application of the contraction method developed in the previous chapter.

5.2. Proof of the functional limit theorem

Theorem 5.1 can be considered a prototype application for the functional contraction method presented in the previous chapter. The verification of conditions **C1** and **C3** in Theorem 5.1 gives rise to the problems **P2** and **P3a** that have been mentioned in the introduction. We will deal with both of them in subsequent sections and assume for a moment the following two propositions to hold true.

Proposition 5.6. *There exists a continuous solution Z of the fixed-point equation* (5.5) *with $\mathbb{E}[Z(s)] = (s(1-s))^{\beta/2}$ and $\mathbb{E}\left[\|Z\|^2\right] < \infty$. Moreover, all moments of $\|Z\|$ are finite.*

Proposition 5.7. *There exists $\varepsilon > 0$ such that*

$$\sup_{s\in[0,1]} |n^{-\beta}\mathbb{E}[C_n(s)] - \bar{\mu}(s)| = O(n^{-\varepsilon}).$$

Here $\bar{\mu}(s) = K_1(s(1-s))^{\beta/2}$.

Following the heuristics in the introduction we scale the additive recurrence (1.18) by n^β. Let $Q_0(t) := 0$ and

$$Q_n(t) = \frac{C_n(t)}{K_1 n^\beta}, \qquad n \geq 1.$$

The recursive distributional equation then rewrites in terms of Q_n as

$$\begin{aligned}(Q_n(t))_{t\in[0,1]} \stackrel{d}{=} \Bigg(&\mathbf{1}_{\{t<U\}}\left[\left(\frac{I_1^{(n)}}{n}\right)^\beta Q^{(1)}_{I_1^{(n)}}\left(\frac{t}{U}\right) + \left(\frac{I_2^{(n)}}{n}\right)^\beta Q^{(2)}_{I_2^{(n)}}\left(\frac{t}{U}\right)\right]\\ &+ \mathbf{1}_{\{t\geq U\}}\left[\left(\frac{I_3^{(n)}}{n}\right)^\beta Q^{(3)}_{I_3^{(n)}}\left(\frac{t-U}{1-U}\right) + \left(\frac{I_4^{(n)}}{n}\right)^\beta Q^{(4)}_{I_4^{(n)}}\left(\frac{t-U}{1-U}\right)\right]\\ &+ \frac{1}{K_1 n^\beta}\Bigg)_{t\in[0,1]}\end{aligned}$$

where $U, I_1^{(n)}, \ldots, I_4^{(n)}$ are the quantities already introduced in the introduction and $(Q_n^{(1)})_{n\geq 0}$, $\ldots$, $(Q_n^{(4)})_{n\geq 0}$ are independent copies of $(Q_n)_{n\geq 0}$, independent of $(U, V, I_1^{(n)}, \ldots, I_4^{(n)})$. The convergence of the coefficients $(I_j^{(n)}/n)^\beta$ suggests that a limit of $Q_n(t)$ satisfies the fixed-point equation (5.5).

A modified recurrence: Remember from condition **C3** that the rescaled sequence has to have distributions satisfying (2.4) for $n \geq n_0$. As computed later in (5.9) the contraction property **C2** is satisfied for $s = 2$ but not for $s = 1$. Hence, for **C3** to be satisfied, we need to use a scaling that leads to an expectation that is independent of n. This is not the case for $Q_n(t)$. Denoting $\mu_n(t) = \mathbb{E}\left[C_n(t)\right]$, we are naturally led to consider $Y_0(t) := 0$ and

$$Y_n(t) = \frac{C_n(t) - \mu_n(t)}{K_1 n^\beta} = Q_n(t) - h(t) + O(n^{-\varepsilon}), \qquad n \geq 1. \tag{5.7}$$

where the error term is deterministic and uniform in $t \in [0,1]$ by Proposition 5.7. Remember that $h(t)$ was defined as $(t(1-t))^{\beta/2}$. The distributional recursion in terms of Y_n is

$$\begin{aligned}(Y_n(t))_{t\in[0,1]} \stackrel{d}{=} \Bigg(&\mathbf{1}_{\{t<U\}}\left[\left(\frac{I_1^{(n)}}{n}\right)^\beta Y^{(1)}_{I_1^{(n)}}\left(\frac{t}{U}\right) + \left(\frac{I_2^{(n)}}{n}\right)^\beta Y^{(2)}_{I_2^{(n)}}\left(\frac{t}{U}\right)\right]\\ &+ \mathbf{1}_{\{t\geq U\}}\left[\left(\frac{I_3^{(n)}}{n}\right)^\beta Y^{(3)}_{I_3^{(n)}}\left(\frac{t-U}{1-U}\right) + \left(\frac{I_4^{(n)}}{n}\right)^\beta Y^{(4)}_{I_4^{(n)}}\left(\frac{t-U}{1-U}\right)\right]\\ &+ \mathbf{1}_{\{t<U\}}\left[\frac{\mu_{I_1^{(n)}}\left(\frac{t}{U}\right) + \mu_{I_2^{(n)}}\left(\frac{t}{U}\right)}{K_1 n^\beta}\right]\\ &+ \mathbf{1}_{\{t\geq U\}}\left[\frac{\mu_{I_3^{(n)}}\left(\frac{t-U}{1-U}\right) + \mu_{I_4^{(n)}}\left(\frac{t-U}{1-U}\right)}{K_1 n^\beta}\right] + \frac{1-\mu_n(t)}{K_1 n^\beta}\Bigg)_{t\in[0,1]},\end{aligned}$$

where $(Y_n^{(1)})_{n\geq 0}, \ldots, (Y_n^{(4)})_{n\geq 0}$ are independent copies of $(Y_n)_{n\geq 0}$ which are also independent of the vector $(U, V, I_1^{(n)}, \ldots, I_4^{(n)})$. Therefore, any possible limit Y of Y_n should satisfy the following distributional fixed-point equation

$$
\begin{aligned}
(Y(t))_{t\in[0,1]} \stackrel{d}{=} \Bigg(& \mathbf{1}_{\{t<U\}} \left[(UV)^{\beta} Y^{(1)}\left(\frac{t}{U}\right) + (U(1-V))^{\beta} Y^{(2)}\left(\frac{t}{U}\right) \right] \\
& + \mathbf{1}_{\{t\geq U\}} \left[((1-U)V)^{\beta} Y^{(3)}\left(\frac{t-U}{1-U}\right) + ((1-U)(1-V))^{\beta} Y^{(4)}\left(\frac{t-U}{1-U}\right) \right] \\
& + \mathbf{1}_{\{t\geq U\}} h\left(\frac{t-U}{1-U}\right) \left(((1-U)V)^{\beta} + ((1-U)(1-V))^{\beta} \right) - h(t) \\
& + \mathbf{1}_{\{t<U\}} h\left(\frac{t}{U}\right) \left((UV)^{\beta} + (U(1-V))^{\beta} \right) \Bigg)_{t\in[0,1]} .
\end{aligned} \tag{5.8}
$$

Having Theorem 3.6 in mind, we define (random) operators $A_r^{(n)}$, $r = 1, \ldots, 4$, by

$$
A_r^{(n)}(f)(t) = \begin{cases} \mathbf{1}_{\{t<U\}} \left(\dfrac{I_r^{(n)}}{n} \right)^{\beta} f\left(\dfrac{t}{U}\right) & \text{if } r = 1, 2 \\ \mathbf{1}_{\{t\geq U\}} \left(\dfrac{I_r^{(n)}}{n} \right)^{\beta} f\left(\dfrac{t-U}{1-U}\right) & \text{if } r = 3, 4. \end{cases}
$$

Furthermore let $b^{(n)}(t) = \sum_{r=1}^{4} b_r^{(n)}(t) + (1-\mu_n(t))/(K_1 n^{\beta})$ with

$$
b_r^{(n)}(t) = \begin{cases} \mathbf{1}_{\{t<U\}} \cdot \dfrac{\mu_{I_r^{(n)}}\left(\frac{t}{U}\right)}{K_1 n^{\beta}} & \text{if } r = 1, 2 \\ \mathbf{1}_{\{t\geq U\}} \cdot \dfrac{\mu_{I_r^{(n)}}\left(\frac{t-U}{1-U}\right)}{K_1 n^{\beta}} & \text{if } r = 3, 4. \end{cases}
$$

Then the finite-n version of the recurrence relation for $(Y_n)_{n\geq 0}$ is precisely of the form (1.2). We define similarly the coefficients of the limit recursive equation (5.8). Based on the two propositions at the beginning of this section, we will then show that with these definitions, all assumptions **C1**–**C5** are satisfied. The operators $A_1, \ldots, A_4$ are defined by

$$
\begin{aligned}
A_1(f)(t) &= \mathbf{1}_{\{t<U\}} (UV)^{\beta} f\left(\frac{t}{U}\right) & A_2(f)(t) &= \mathbf{1}_{\{t<U\}} (U(1-V))^{\beta} f\left(\frac{t}{U}\right) \\
A_3(f)(t) &= \mathbf{1}_{\{t\geq U\}} ((1-U)V)^{\beta} f\left(\frac{t-U}{1-U}\right) & A_4(f)(t) &= \mathbf{1}_{\{t\geq U\}} ((1-U)(1-V))^{\beta} f\left(\frac{t}{U}\right)
\end{aligned}
$$

and $b(t) = \sum_{r=1}^{4} b_r(t) - h(t)$ with

$$
\begin{aligned}
b_1(t) &= \mathbf{1}_{\{t<U\}} (UV)^{\beta} h\left(\frac{t}{U}\right), & b_2(t) &= \mathbf{1}_{\{t<U\}} (U(1-V))^{\beta} h\left(\frac{t}{U}\right) \\
b_3(t) &= \mathbf{1}_{\{t\geq U\}} ((1-U)V)^{\beta} h\left(\frac{t-U}{1-U}\right), & b_4(t) &= \mathbf{1}_{\{t\geq U\}} ((1-U)(1-V))^{\beta} h\left(\frac{t}{U}\right).
\end{aligned}
$$

The operators $A_1, \ldots, A_4, A_1^{(n)}, \ldots, A_4^{(n)}$ are linear for each n. Moreover, it is immediate to see that they are bounded above by one which implies them to be continuous. Obviously, their norm functions are real-valued random variables. In order to establish them to be random continuous linear operators on $(\mathcal{D}[0,1], d_{sk})$ it remains to check that they are continuous with respect to the Skorokhod topology. To this end, it is sufficient to prove that

$$
d_{sk}(f_n, f) \to 0 \Rightarrow d_{sk}\left(\mathbf{1}_{\{t\leq u\}} f_n\left(\frac{t}{u}\right), \mathbf{1}_{\{t\leq u\}} f\left(\frac{t}{u}\right) \right) \to 0
$$

for any $u \in [0,1]$. This follows easily since $\|f_n(\lambda_n(t)) - f(t)\| \to 0$ with monotonically increasing bijections λ_n on the unit interval such that $\|\lambda_n(t)-t\| \to 0$ implies $\|(\mathbf{1}_{\{\beta_n(t)\le u\}} f_n(\beta_n(t)/u) - \mathbf{1}_{\{t\le u\}} f(t/u)\| \to 0$ where $\beta_n(t) = u\,\lambda_n(t/u)$ for $t \le u$ and $\beta_n(t) = t$ for $t > u$.
We are now ready to check that the conditions **C1–C5** indeed hold.

C3 - Existence of a continuous solution. By Proposition 5.6 we have a continuous solution Z of the fixed-point equation (5.5) with $\mathbb{E}\left[\|Z\|^2\right] < \infty$ and $\mathbb{E}\left[Z(t)\right] = h(t) = (t(1-t))^{\beta/2}$. A proof this the existence of such a process is given in Section 5.3. Hence the function $Y(t) = Z(t) - h(t)$ is a continuous solution of (5.8) with $\mathbb{E}\left[Y(t)\right] = 0$ and $\mathbb{E}\left[\|Y\|^2\right] < \infty$.

C2 - Contraction. An easy computation shows that $\mathbb{E}\left[\|A_r\|^2\right] = \mathbb{E}\left[(UV)^{2\beta}\right] = (2\beta+1)^{-2}$. Thus,

$$L = \sum_{r=1}^{4} \mathbb{E}\left[\|A_r\|^2\right] = \frac{4}{(2\beta+1)^2} < 1. \tag{5.9}$$

In particular, Y is the unique solution of (5.8) with $\mathbb{E}\left[Y(t)\right] = 0$ and $\mathbb{E}\left[\|Y\|^2\right] < \infty$.

C1 and C4 - Convergence of the coefficients. It suffices to focus on the terms

$$\|A_1^{(n)} - A_1\|_2 \qquad \text{and} \qquad \|b_1^{(n)} - b_1\|_2,$$

the remaining terms can obviously be treated in the same way. Establishing the convergence only boils down to verifying that a binomial random variable $\mathrm{Bin}(n,p)$ is properly approximated by np. Using the Chernoff–Hoeffding inequality for the binomial distribution [Hoe63], one easily verifies that for every $\alpha > 0$,

$$\mathbb{E}\left[\left|\frac{\mathrm{Bin}(n,p)}{n} - p\right|^{\alpha}\right] = O(n^{-\alpha/2}), \tag{5.10}$$

uniformly in $p \in [0,1]$. Thus, since $|x^\beta - y^\beta| \le |x-y|^\beta$ for any $x, y \in [0,1]$, we have

$$\|A_1^{(n)} - A_1\|_2 \le \left\|\left(\frac{I_r^{(n)}}{n}\right)^{\beta} - (UV)^{\beta}\right\|_2 = O(n^{-1/2}). \tag{5.11}$$

By Proposition 5.7 we have $\mu_n(t) = K_1 h(t) n^\beta + O(n^{\beta-\varepsilon})$ uniformly in $t \in [0,1]$. Therefore

$$\|b_1^{(n)} - b_1\|_2 \le \left\|\mathbf{1}_{\{t<U\}} h\left(\frac{t}{U}\right)\left(\left(\frac{I_r^{(n)}}{n}\right)^{\beta} - (UV)^{\beta}\right)\right\|_2 + C\left\|\frac{(I_1^{(n)})^{\beta-\varepsilon}}{n^\beta}\right\|_2,$$

for some constant $C > 0$. Since h is bounded, the first summand is $O(n^{-1/2})$ just like in (5.11) above. The second term is trivially bounded by $Cn^{-\varepsilon}$. Overall, we have $\|b_1^{(n)} - b_1\|_2 = O(n^{-\varepsilon})$. Hence, since the coefficients $A_r^{(n)}$ are bounded by one in the operator norm and by distributional properties of $I_1^{(n)}, \ldots, I_4^{(n)}$, condition **C1** is satisfied. Moreover, in **C4**, we may choose $R(n) = Cn^{-\varepsilon}$ for a suitable constant $C > 0$ and $\varepsilon > 0$ as small as we want. By dominated convergence we have

$$\begin{aligned} L^* &= \limsup_{n\to\infty} \mathbb{E}\left[\sum_{r=1}^{4} \|A_r^{(n)}\|^2 \frac{R(I_r^{(n)})}{R(n)}\right] \\ &= 4\mathbb{E}\left[(UV)^{2\beta}(UV)^{-\varepsilon}\right] \\ &= \frac{4}{(2\beta-\varepsilon+1)^2} < 1, \end{aligned}$$

for $\varepsilon > 0$ sufficiently small. This completes the verification of **C4**.

C5 and (3.12) **- Rate of convergence.** Note that Q_n is piecewise constant: $Q_n(t) = Q_n(s)$ for all s, t if no x-coordinate of the first n points lies between s and t. There are n independent points, the probability that there exists two lying within n^{-3} of each other is at most n^{-1}. So **C5** is satisfied with $r_n = n^{-3}$ and $R_n = o(\log^{-2} n) = o(\log^{-2}(1/r_n))$ Therefore, the condition on the rate of convergence is satisfied.

Conclusion: We have shown $(Y_n(s))_{s\in[0,1]} \to (Y(s))_{s\in[0,1]}$ in distribution. By the very definition of Y_n (5.7) and the relation $Y(t) = Z(t) - h(t)$ this implies the functional limit law for $(Q_n(s))_{s\in[0,1]}$. The remaining statements of Theorem 5.1 follow from Proposition 5.6 and the observation in [CJ11, Section 5] that the mean of any process satisfying (5.5) whose mean function is integrable over $[0, 1]$ has to be proportional to $h(s)$. Theorem 5.2 also follows from Theorem 3.6 where the identification of the limit variance is worked out in Section 5.5. Mean convergence of $(\|Y_n\|^2)_{n\geq 1}$ follows from Corollary 3.7 by choosing $r_n = n^{-5}$. Proposition 2.39 implies $\|Y\|$ to have moments of arbitrary order and $\mathbb{E}[\|Y_n\|^\kappa] \to \mathbb{E}[\|Y\|^\kappa]$ for all $\kappa > 0$. As for the process convergence, these results transfer to Q_n and Z and prove Theorem 5.3. Theorem 5.4 follows immediately.

5.3. The limit process

The aim of this section is to prove Proposition 5.6, i.e. the existence of a process Z on the unit interval with continuous paths, that satisfies the distributional fixed point equation (5.5) whose relevant moments match (asymptotically) with the corresponding ones of the rescaled version of $C_n(s)$.

As indicated in the introduction, we will find a representation of Z as an infinite series that converges almost surely. The justification of the point-wise convergence is done by a martingale argument. Showing that the convergence is almost surely uniform allows to deduce that Z has continuous paths.

We identify the nodes of the infinite 4-ary tree with the set of finite words on the alphabet $\{1, 2, 3, 4\}$,

$$\mathcal{T} = \bigcup_{n\geq 0} \{1, 2, 3, 4\}^n.$$

For a node $u \in \mathcal{T}$, we write $|u|$ for its depth, i.e.
the distance between u and the root $\varnothing$. The descendants of $u \in \mathcal{T}$ correspond to all the words in $\mathcal{T}$ with prefix u; in particular, the children of u are $u1, \ldots, u4$. Let $\{U_v, v \in \mathcal{T}\}$ and $\{V_v, v \in \mathcal{T}\}$ be two independent families of i.i.d. $[0, 1]$-uniform random variables. By $\mathcal{C}_0[0, 1]$ we denote the set of continuous functions on the unit interval vanishing at the boundary, i.e. $f(0) = f(1) = 0$ for $f \in \mathcal{C}_0[0, 1]$. Define the continuous operator $G : (0, 1)^2 \times \mathcal{C}_0[0, 1]^4 \to \mathcal{C}_0[0, 1]$ by

$$\begin{aligned} G(x, y, f_1, f_2, f_3, f_4)(s) =& \mathbf{1}_{\{s<x\}} \left[(xy)^\beta f_1\left(\frac{s}{x}\right) + (x(1-y))^\beta f_2\left(\frac{s}{x}\right)\right] \\ &+ \mathbf{1}_{\{s\geq x\}} \left[((1-x)y)^\beta f_3\left(\frac{s-x}{1-x}\right) + ((1-x)(1-y))^\beta f_4\left(\frac{s-x}{1-x}\right)\right]. \end{aligned} \tag{5.12}$$

For every node $u \in \mathcal{T}$, let $Z_0^u = h$ where $h(s) = (s(1-s))^{\beta/2}$ as in (1.17). Then define recursively

$$Z_{n+1}^u = G(U_u, V_u, Z_n^{u1}, Z_n^{u2}, Z_n^{u3}, Z_n^{u4}). \tag{5.13}$$

Starting the iteration with the initial deterministic value h in all nodes at level n, let $Z_n = Z_n^{\varnothing}$ be the value observed at the root of $\mathcal{T}$. As it turns out, for every $s \in [0,1]$, the sequence $(Z_n(s))_{n\geq 0}$ is a non-negative discrete time martingale hence it converges to an integrable limiting random variable almost surely.
It will be convenient to have an explicit representation for Z_n. For $s \in [0,1]$, $Z_n(s)$ is the sum of exactly 2^n terms, each one being the contribution of one of the boxes at level n that is cut by the line at s. Let $\{Q_i^n(s), 1 \leq i \leq 2^n\}$ be the set of rectangles at level n whose first coordinate intersect s. Suppose that the projection of $Q_i^n(s)$ on the first coordinate yields the interval $[\ell_i^n, r_i^n]$. Then

$$Z_n(s) = \sum_{i=1}^{2^n} \mathrm{Leb}(Q_i^n(s))^\beta \cdot h\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right), \tag{5.14}$$

where $\mathrm{Leb}(Q_i^n(s))$ denotes the volume of the rectangle $Q_i^n(s)$. Here, we abbreviate $h(s) = 0$ for $s < 0$ or $s > 1$. The difference between Z_n and Z_{n+1} can easily be expressed in terms of then changes in the boxes $Q_i^n(s)$: We have

$$Z_{n+1}(s) - Z_n(s) = \sum_{i=1}^{2^n} \mathrm{Leb}(Q_i^n(s))^\beta \cdot \left[G(U_i', V_i', h, h, h, h)\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right) - h\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right)\right], \tag{5.15}$$

where U_i', V_i', $1 \leq i \leq 2^n$ are i.i.d. $[0,1]$-uniform random variables. In fact, U_i' and V_i' are some of the variables U_u, V_u for nodes u at level n. Observe that, although $Q_i^n(s)$ is *not* a product of n independent terms of the form UV because of size-biasing, but U_i', V_i' are in fact *unbiased*, i.e. uniform. Let $\mathscr{F}_n$ denote the σ-algebra generated by $\{U_u, V_u : |u| < n\}$. Then the family $\{U_i', V_i' : 1 \leq i \leq 2^n\}$ is independent of $\mathscr{F}_n$.
So, to prove that $Z_n(s)$ is a martingale, it suffices to prove that, for $1 \leq i \leq 2^n$,

$$\mathbf{E}\left[\, G(U_i', V_i', h, h, h, h)\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right) \;\middle|\; \mathscr{F}_n \,\right] = h\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right).$$

Since $U_i', V_i', 1 \leq i \leq 2^n$ are independent of $\mathscr{F}_n$, this reduces to the following lemma.

Lemma 5.8. *For the operator G defined in* (5.12)*, two independent random variables U, V each uniformly distributed on the unit interval, and any $s \in [0,1]$, we have*

$$\mathbb{E}\left[G(U, V, h, h, h, h)(s)\right] = h(s).$$

Proof. Since V and $1-V$ have the same distribution, we have

$$\mathbb{E}\left[G(U,V,h,h,h,h)(s)\right] = 2\mathbb{E}\left[\mathbf{1}_{\{s<U\}}(UV)^\beta h\left(\frac{s}{U}\right)\right] + 2\mathbb{E}\left[\mathbf{1}_{\{s\geq U\}}((1-U)V)^\beta h\left(\frac{1-s}{1-U}\right)\right].$$

Similarly, since U and $1-U$ are both uniform, we clearly have

$$\mathbb{E}\left[G(U,V,h,h,h,h)(s)\right] = f(s) + f(1-s),$$

where we wrote $f(s) = 2\mathbb{E}\left[\mathbf{1}_{\{s<U\}}(UV)^\beta h(s/U)\right]$. To complete the proof, it suffices to compute $f(s)$. We have

$$\begin{aligned}
f(s) = 2\mathbb{E}\left[\mathbf{1}_{\{s<U\}}(UV)^\beta h\left(\frac{s}{U}\right)\right] &= \frac{2}{\beta+1}\mathbb{E}\left[\mathbf{1}_{\{s<U\}}U^\beta\left(\frac{s}{U}\left(1-\frac{s}{U}\right)\right)^{\beta/2}\right] \\
&= \frac{2}{\beta+1}\mathbb{E}\left[\mathbf{1}_{\{s<U\}}s^{\beta/2}(U-s)^{\beta/2}\right] \\
&= \frac{2}{\beta+1}s^{\beta/2}\int_s^1 (x-s)^{\beta/2}dx \\
&= \frac{4}{(\beta+1)(\beta+2)}s^{\beta/2}(1-s)^{\beta/2+1} \\
&= (1-s)h(s),
\end{aligned}$$

where the last line follows since $(\beta+1)(\beta+2) = 4$ by definition of β. It then follows easily that

$$\mathbb{E}\left[G(U,V,h,h,h,h)(s)\right] = (1-s)h(s) + sh(1-s) = h(s),$$

which completes the proof. □

We could now use the martingale convergence theorem to define $Z(s)$ as the limit of $Z_n(s)$ for s fixed. However, since converges only holds almost surely and the unit interval is uncountable, it is not clear that we would thus define a proper limit on a set of $\mathbb{P}$ measure 1. The next proposition which is our main result of the section is proved by means of concentration inequalities and properties of random quadtrees. A simulation of the limit process is presented at the end of the introduction in figure (1.3) on page 12.

Proposition 5.9. *With probability one Z_n converges uniformly to some continuous limit process Z on $[0,1]$.*

It is well-known that (Z_n) has the Cauchy property in $(\mathcal{C}[0,1], \|\cdot\|)$ almost surely if and only if $\sup_{m\geq n}\|Z_m - Z_n\|$ tends to zero in probability as $n\to\infty$. The latter is immediate if we find constants constants $a, b \in (0,1)$ and $C > 0$ such that

$$\mathbf{P}\left(\sup_{s\in[0,1]}|Z_{n+1}(s) - Z_n(s)| \geq a^n\right) \leq C\cdot b^n. \tag{5.16}$$

Completeness of the state space implies the existence of a continuous process Z such that, almost surely, $Z_n \to Z$ uniformly $[0,1]$. We now move on to showing that there exist constants a and b such that (5.16) is satisfied. We start by a bound for a fixed value $s \in [0,1]$. We will then handle the supremum using a large enough number of fixed points in the unit interval and bounding the variations in between. The following Lemma is a necessary tool for the remainder of this section. Its proof relies on the standard Chernoff bound for the exponential distribution; for independent random variables $E_1, \ldots, E_n$ with exponential distribution of unit mean, we have

$$\mathbf{P}\left(\sum_{i=1}^n E_i \geq tn\right) \leq e^{-n(t-\log t-1)} \tag{5.17}$$

for $t > 1$. Analogously, for $0 < t < 1$,

$$\mathbf{P}\left(\sum_{i=1}^n E_i \leq tn\right) \leq e^{-n(t-\log t-1)} \tag{5.18}$$

Lemma 5.10. *Let W_n denote the maximum width of a cell at level n in the construction of the process Z_n. Then, for $c < 1$,*

$$\mathbf{P}\left(W_n \ge c^n\right) \le (4e\log(1/c))^n.$$

Proof. Let U_i, $i \ge 1$ be a family of i.i.d.
$[0,1]$-uniform random variables and E_i, $i \ge 1$, be a family of i.i.d. exponential(1) random variables. Then, the union bound and the Chernoff bound (5.18) for the left tail yield

$$\begin{aligned}\mathbf{P}\left(W_n \ge c^n\right) &\le 4^n \cdot \mathbf{P}\left(\prod_{i=1}^n U_i \ge c^n\right)\\ &= 4^n \cdot \mathbf{P}\left(\sum_{i=1}^n E_i \le n\log(1/c)\right)\\ &\le 4^n \exp(-n(\log(1/c) - 1 - \log\log(1/c)))\\ &\le (4e\log(1/c))^n,\end{aligned}$$

as desired. □

Lemma 5.11. *For every $s \in [0,1]$, any $a \in (0,1)$, and any integer n large enough, we have the bound*

$$\mathbf{P}\left(|Z_{n+1}(s) - Z_n(s)| \ge a^n\right) \le 4(16e\log(1/a))^n.$$

Proof. We use the representation (5.15). As we have already pointed out earlier in Lemma 5.8, for every single rectangle $Q_i^n(s)$ at level n, we have

$$\mathbf{E}\left[\, G(U_i', V_i', h, h, h, h)\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right) - h\left(\frac{s-\ell_i^n}{r_i^n - \ell_i^n}\right) \;\middle|\; \mathscr{F}_n \,\right] = 0.$$

Since $h(x) \le 2^{-\beta}$ for $x \in (0,1)$, conditional on $\mathscr{F}_n$, $Z_{n+1} - Z_n$ is a sum of 2^n *centered*, *bounded* and moreover *independent* terms (but not identically distributed). Moreover, conditional on $\mathscr{F}_n$, the term corresponding to $Q_i^n(s)$ in (5.15) is bounded by

$$\begin{aligned}\mathrm{Leb}(Q_i^n)^\beta \cdot \|G(U_i', V_i', h, h, h, h) - h\| &\le \mathrm{Leb}(Q_i^n)^\beta 2\|h\|\\ &= \mathrm{Leb}(Q_i^n)^\beta 2^{1-\beta}.\end{aligned} \tag{5.19}$$

So when conditioning on $\mathscr{F}_n$, one can bound the variations of $Z_{n+1} - Z_n$ using the Chernoff–Hoeffding inequality [Hoe63]. We have

$$\begin{aligned}\mathbf{P}\left(|Z_{n+1}(s) - Z_n(s)| > a^n\right) &= \mathbb{E}\left[\mathbf{P}(|Z_{n+1}(s) - Z_n(s)| > a^n \mid \mathscr{F}_n)\right]\\ &\le \mathbb{E}\left[2\exp\left(-\frac{a^{2n}}{\sum_{i=1}^{2^n}\mathrm{Leb}(Q_i^n(s))^{2\beta}}\right)\right]\\ &\le 2\exp\left(-a^{-2n}\right) + 2\mathbf{P}\left(\sum_{i=1}^{2^n}\mathrm{Leb}(Q_i^n(s))^{2\beta} > a^{4n}\right),\end{aligned} \tag{5.20}$$

where the precise constant in the exponent in the second inequality can be taken to be one since $2/(2^{1-\beta})^2 > 1$. Now, since $2\beta > 1$ and all the volumes $\mathrm{Leb}(Q_i^n(s))$ are at most one, we have

$$\begin{aligned}\mathbf{P}\left(\sum_{i=1}^{2^n}\mathrm{Leb}(Q_i^n(s))^{2\beta} > a^{4n}\right) &\le \mathbf{P}\left(\sum_{i=1}^{2^n}\mathrm{Leb}(Q_i^n(s)) > a^{4n}\right)\\ &\le \mathbf{P}\left(W_n > a^{4n}\right),\end{aligned} \tag{5.21}$$

where W_n denotes the maximum width of any of the 4^n cells at level n. Indeed, the volume covered by the union of all rectangles $Q_i^n(s)$, $1 \leq i \leq 2^n$ is at most that of a vertical tube of width W_n. Putting together (5.20) and (5.21) yields

$$\begin{aligned}\mathbf{P}\left(|Z_{n+1}(s) - Z_n(s)| \geq a^n\right) &\leq 2\exp(-a^{-2n}) + 2\mathbf{P}\left(W_n > a^{4n}\right) \\ &\leq 2\exp(-a^{-2n}) + 2(16e\log(1/a))^n \\ &\leq 4(16e\log(1/a))^n,\end{aligned}$$

for all n large enough using Lemma 5.10. □

The previous lemma provides good control on pointwise variations of $Z_{n+1} - Z_n$ and we move on to the supremum on $[0, 1]$ now. Let V_n be the set of x-coordinates of the vertical boundaries of all the rectangles at level n and $L_n = \inf\{|x - y| : x, y \in V_n\}$. Subsequently, we need a tail bound for the quantity L_n. Its proof is concerned with the saturation level of a random quadtree. By saturation level we denote the deepest level ℓ in the tree in which all 2^ℓ internal nodes are present. The quantity is studied in [Dev87]; we use arguments resembling ideas from this work to deduce a precise tail bound.

Lemma 5.12. *Let S_k be the saturation level of a random quadtree of size k. Then, for every positive integer $x > 22$, it exists an integer $n_0(x)$ with*

$$\mathbf{P}\left(S_{x^n} < n\right) \leq 4^{n+1}x^{-n/100}, \qquad n \geq n_0(x).$$

Proof. We consider the 4^n possible nodes on level n. By symmetry each of them is occupied by a key with the same probability. Looking at a specific one, e.g. the leftmost, it is obvious that its subtree size is stochastically bounded by $\mathrm{Bin}(x^n; U_1V_1 \cdots U_nV_n) - n$ where $\{U_i, i \geq 1\}$ and $\{V_i, i \geq 1\}$ are independent families of i.i.d. $[0, 1]$-uniform random variables. Then by the union bound applied to the 4^n cells at level n, using Chernoff's inequality, we have

$$\begin{aligned}\mathbf{P}\left(S_{x^n} < n\right) &\leq 4^n \cdot \mathbf{P}\left(\mathrm{Bin}\left(x^n; U_1V_1 \cdots U_nV_n\right) \leq n\right) \\ &\leq 4^n \cdot \exp\left(-(1 - n2^{-n})^2 2^{n+1}\right) + 4^n\mathbf{P}\left(U_1V_1 \cdots U_nV_n \leq \left(\frac{2}{x}\right)^n\right). \end{aligned} \tag{5.22}$$

However, using once again the large deviations principle (5.17) for sums of i.i.d. exponential random variables $E_i, i \geq 1$,

$$\begin{aligned}\mathbf{P}\left(U_1V_1 \cdots U_nV_n \leq \left(\frac{2}{x}\right)^n\right) &= \mathbf{P}\left(\sum_{i=1}^{2n} E_i \geq n\log(x/2)\right) \\ &\leq \exp\left(-2n\left(\frac{\log(x/2)}{2} - 1 - \log\frac{\log(x/2)}{2}\right)\right) \\ &\leq x^{-n/100}, \end{aligned} \tag{5.23}$$

for all $x > 22$ since then $\frac{e^2}{2}\log^2(x/2) \leq x^{99/100}$. Combining (5.22) and (5.23), we obtain

$$\mathbf{P}\left(S_{x^n} < n\right) \leq 4^n\exp\left(-2^{n-1}\right) + 4^n \cdot x^{-n/100} \leq 4^{n+1}x^{-n/100},$$

for $x > 22$ and n large enough. □

Lemma 5.13. *There exists $0 < \gamma_0 < 1$ such that any positive real number $\gamma < \gamma_0$, there exists an integer $n_1(\gamma)$ with*

$$\mathbf{P}(L_n < \gamma^n) \leq 6^{n+1}\gamma^{n/201}, \qquad n \geq n_1(\gamma).$$

Proof. The joint distribution of the x-coordinates of the vertical lines in the tree developed up to level n is complex. In particular, it is *not* that of independent uniform points on $[0,1]$. However, we can use a simple coupling with a family of i.i.d. random points on $[0,1]^2$ that yields a good enough lower bound on L_n.
Let $\xi_i = (U_i, V_i)$, $i \geq 1$ be i.i.d. uniform random points on $[0,1]^2$. Let T_k be the quadtree obtained by inserting the random points ξ_i, $1 \leq i \leq k$, in this order. We write D_i for the depth at which the point ξ_i is inserted where the root has depth zero. Let K_n be the first k for which the tree T_k is complete up to level n, i.e. T_k has 4^n cells at level n and T_{k-1} has less on this level. Then, by definition $\{\xi_i : i \geq 1, D_i < n\}$ has the distribution of the set of points used to construct the process Z_n. Obviously, $\{\xi_i : i \geq 1, D_i < n\} \subseteq \{\xi_i : 1 \leq i \leq K_n\}$ and for any integer x,

$$\begin{aligned}
\mathbf{P}(L_n < \gamma^n) &\leq \mathbf{P}(\exists i, j \leq K_n : i \neq j, |U_i - U_j| < \gamma^n) \\
&\leq \mathbf{P}(\exists i, j \leq x^n : i \neq j, |U_i - U_j| < \gamma^n) + \mathbf{P}(K_n > x^n) \\
&\leq x^{2n} \cdot 2\gamma^n + \mathbf{P}(K_n > x^n),
\end{aligned}$$

by the union bound. The random variable K_n is related to the saturation level as introduced in the previous Lemma. We obtain

$$\mathbf{P}(K_n > x^n) = \mathbf{P}(S_{x^n} < n) \leq 4(4x^{-1/100})^n,$$

as long as $x > 22$ and $n \geq n_0(x)$, compare the conditions of the previous Lemma. It follows readily that

$$\begin{aligned}
\mathbf{P}(L_n < \gamma^n) &\leq 2(x^2\gamma)^n + 4(4x^{-1/100})^n \\
&\leq 6^{n+1}\gamma^{n/201},
\end{aligned}$$

upon choosing $x = \lceil 4^{100/201}\gamma^{-100/201} \rceil$ [that is $x^2\gamma \approx 4x^{-1/100}$] and $\gamma < 4 \cdot 22^{-2.01} =: \gamma_0$ which implies $x > 22$. This completes the proof. $\square$

We continue with the proof of (5.16). For technical reasons, suppose that $1/\gamma$ is an integer. Then, we have

$$\begin{aligned}
\sup_{s\in[0,1]} |Z_{n+1}(s) - Z_n(s)| &\leq \sup_{1\leq i\leq \gamma^{-(n+1)}} \sup_{i\gamma^{n+1}\leq s\leq (i+1)\gamma^{n+1}} |Z_{n+1}(s) - Z_n(s)| \\
&\leq \sup_{1\leq i\leq \gamma^{-(n+1)}} |Z_{n+1}(i\gamma^{n+1}) - Z_n(i\gamma^{n+1})| \\
&\quad +2 \sup_{m\in\{n,n+1\}} \sup_{|s-t|\leq\gamma^{n+1}} |Z_m(s) - Z_m(t)|.
\end{aligned}$$

We first deal with the second term, and suppose that we are on the event that $L_{n+1} \geq (4\gamma)^{n+1}$. Observe that the sieve we used, γ^{n+1}, is much finer than the shortest length of a cell at level $n+1$ which is at least L_{n+1}. We use the representation in (5.14); for $|t-s| \leq \gamma^{n+1}$, the two collections $\{Q_i^n(s), 1 \leq i \leq 2^n\}$ and $\{Q_i^n(t), 1 \leq i \leq 2^n\}$ differ at most on one cell. We obtain, for any

$|s-t| \leq \gamma^{n+1}$,

$$\begin{aligned}
|Z_n(s)-Z_n(t)| &\leq \sum_{i=1}^{2^n} \mathrm{Leb}(Q_i^n(s))^\beta \cdot \left| h\left(\frac{s-\ell_i^n}{r_i^n-\ell_i^n}\right) - h\left(\frac{t-\ell_i^n}{r_i^n-\ell_i^n}\right)\right| + 2\max_i \mathrm{Leb}(Q_i^n(s))^\beta \\
&\leq \sum_{i=1}^{2^n} \mathrm{Leb}(Q_i^n(s))^\beta \cdot 4^{-\beta(n+1)} + 2\max_i \mathrm{Leb}(Q_i^n(s))^\beta \\
&\leq 3W_n^\beta.
\end{aligned}$$

Here, in second inequality follows from the fact that $|h(t)-h(s)| \leq |t-s|^\beta$ for any $s,t \in [0,1]$ and the fact that $L_n \geq (4\gamma)^{n+1}$. The same upper bounds holds for $|Z_{n+1}(s)-Z_{n+1}(t)|$ on $\|s-t\| \leq \gamma^{n+1}$. Thus, by the union bound, for any $\gamma \in (0,1)$,

$$\begin{aligned}
\mathbf{P}\left(\sup_{s\in[0,1]} |Z_{n+1}(s)-Z_n(s)| \geq 2a^n\right) \leq& \gamma^{-n} \sup_{s\in[0,1]} \mathbf{P}\left(|Z_{n+1}(s)-Z_n(s)| \geq a^n\right) \\
&+ \mathbf{P}\left(L_{n+1} < (4\gamma)^{n+1}\right) + \mathbf{P}\left(12W_n^\beta > a^n\right). \quad (5.24)
\end{aligned}$$

We are now ready to complete the proof of Proposition 5.9. From (5.24) and Lemma 5.13,

$$\begin{aligned}
\mathbf{P}\left(\sup_{s\in[0,1]} |Z_{n+1}(s)-Z_n(s)| \geq 2a^n\right) \quad \leq \quad & 4(16e\gamma^{-1}\log(1/a))^n + 6\cdot 16^n\gamma^{n/201} \\
& + (4e\log(12^{1/n}/a)/\beta)^n,
\end{aligned}$$

for all $\gamma < \gamma_0/4$ and $n \geq n_0(\gamma,a)$. Now, first choose $a < 1$ sufficiently close to 1 such that $16(e\log(1/a))^{1/202} < 1/4$ and then $\gamma > 0$ such that $1/\gamma$ is an integer and $\gamma^{1/201} \leq e\gamma^{-1}\log(1/a)$. It follows that, for n sufficiently large,

$$\mathbf{P}\left(\sup_{s\in[0,1]} |Z_{n+1}(s)-Z_n(s)| \geq 2a^n\right) \leq 11\cdot 4^{-n}.$$

Increasing $a < 1$ and C clearly ensures that (5.16) holds with $b = 1/4$ for all $n \geq 1$.
The functions at the four children of the root, $Z_n^1, \dots, Z_n^4$ are distributed as Z_{n-1}; they also converge uniformly to continuous limits denoted $Z^{(1)}, \dots, Z^{(4)}$. The random functions $Z^{(1)}, \dots, Z^{(4)}$ are independent and distributed as Z. Equation (5.13) and independence imply

$$\begin{aligned}
Z(s) =& \mathbf{1}_{\{s<U\}} \left[(UV)^\beta Z^{(1)}\left(\frac{s}{U}\right) + (U(1-V))^\beta Z^{(2)}\left(\frac{s}{U}\right)\right] \\
&+ \mathbf{1}_{\{s\geq U\}} \left[((1-U)V)^\beta Z^{(3)}\left(\frac{s-U}{1-U}\right) + ((1-U)(1-V))^\beta Z^{(4)}\left(\frac{s-U}{1-U}\right)\right],
\end{aligned}$$

almost surely, considered as random continuous paths. In particular, Z solves the distributional fixed-point equation (5.5).
Finally, we look at the moments of $\|Z_n\| = \sup_{s\in[0,1]} |Z_n(s)|$ and $\|Z\| = \sup_{s\in[0,1]} |Z(s)|$.

Proposition 5.14. *For every $p \geq 1$, we have $\mathbb{E}\left[\|Z\|^p\right] < \infty$, and $\mathbb{E}\left[\|Z_n - Z\|^p\right] \to 0$.*

Proof. Let $\Delta_n(x) = \mathbf{P}\left(\|Z_{n+1}-Z_n\| \geq x\right)$ and $a < 1, C > 0$ such that (5.16) is satisfied with $b = 1/4$. Then, by (5.15) and the upper bound (5.19), we have

$$\mathbb{E}\left[\|Z_{n+1}-Z_n\|\right] = \int_0^\infty \Delta_n(x)dx = \int_0^{a^n} \Delta_n(x)dx + \int_{a^n}^{2^{n+1}} \Delta_n(x)dx. \quad (5.25)$$

The first summand is at most a^n, the second one at most $C \cdot 2^{-(n-1)}$ by (5.16). Altogether, there exists $R > 0$ and $0 < q < 1$ with

$$\mathbb{E}\left[\|Z_{n+1} - Z_n\|\right] \leq Rq^n$$

for all n. Furthermore, for any $p \in \mathbb{N}$, our proof also provides (5.16) for a constant $C > 0$ and $b = 4^{-p}$ by increasing the value of a. Therefore, replacing a^n and 2^{n+1} by a^{np} resp. $2^{(n+1)p}$ in (5.25) shows that also the p-th moment of $\|Z_{n+1} - Z_n\|$ is exponentially small in n for any $p > 1$. Then, using Minkowski's inequality

$$\begin{aligned}\mathbb{E}\left[\|Z_n\|^p\right]^{1/p} &= \mathbb{E}\left[\left\|\sum_{k\geq 0}^{n}(Z_k - Z_{k-1}) + h\right\|^p\right]^{1/p} \\ &\leq \mathbb{E}\left[\left(\sum_{k\geq 0}^{n}\|Z_k - Z_{k-1}\| + \|h\|\right)^p\right]^{1/p} \\ &\leq \sum_{k\geq 0}^{n}\mathbb{E}\left[\|Z_k - Z_{k-1}\|^p\right]^{1/p} + \mathbb{E}\left[\|h\|^p\right]^{1/p},\end{aligned}$$

which is uniformly bounded in n. This implies finite moments of $\|Z\|$ or arbitrary order. The L^p convergence follows along similar lines. □

Remark 5.15. It is worth mentioning that we can also consider Z_n as a martingale sequence in the space of continuous functions, that is $\mathbb{E}\left[Z_{n+1}|\mathscr{F}_n\right] = Z_n$ almost surely. This immediately follows from the fact that the relation $\mathbb{E}\left[Y\right](s) = \mathbb{E}\left[Y(s)\right]$ for a continuous process Y extends to conditional expectations, that is $\mathbb{E}\left[Y|\mathscr{F}\right](s) = \mathbb{E}\left[Y(s)|\mathscr{F}\right]$ for any sub-σ-algebra $\mathscr{F}$. For convergence results of martingale sequences in separable Banach spaces we refer to the book of Neveu [Nev75]. It appears that results which provide uniform convergence given L^p boundedness of the norm for some $p > 1$ extend only to reflexive Banach spaces or Hilbert spaces. Nevertheless, we have the Doob representation $Z_n = \mathbb{E}\left[Z|\mathscr{F}_n\right]$. We finally note that, in finding (and even more classifying) fixed-points of the map T in 1.5 in the real-valued case, one main approach to construct solutions as almost sure limits of martingales. These limits also provide series representations for the fixed-points. For more detailed information we refer to [ABM12] and the references therein.

5.4. Uniform convergence of the mean

In this section we prove Proposition 5.7 to hold true. Note that, since $C_n(s)$ is continuous at any fixed $s \in [0,1]$ almost surely, the function $s \to \mathbb{E}\left[C_n(s)\right]$ is continuous for any n.
Following [CJ11], we first prove a poissonized version, the routine depoissonization arguments yielding Proposition 5.7 are presented in Subsection 5.4.3. Consider a Poisson point process with unit intensity on $[0,1]^2 \times [0,\infty)$. The first two coordinates represent the location inside the unit square, whereas the third one represents the time of arrival of the point. Let $P_t(s)$ denote the partial match cost for a query at $x = s$ in the quadtree built from the points arrived by time t.

Proposition 5.16. *There exists $\varepsilon > 0$ such that*

$$\sup_{s\in[0,1]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| = O(t^{-\varepsilon}).$$

The proof of Proposition 5.16 relies crucially on two main ingredients: First, a strengthening of the arguments developed by Curien and Joseph [CJ11], and second, the asymptotic expansion of $\mathbb{E}\left[C_n(\xi)\right]$ for a uniform query line ξ, see (1.14), by Chern and Hwang [CH03]. By symmetry, we write for any $\delta \in (0, 1/2)$

$$\begin{aligned}\sup_{s\in[0,1]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| &= \sup_{s\in[0,1/2]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| \\ &\le \sup_{s\le\delta} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| + \sup_{s\in(\delta,1/2]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right|.\end{aligned} \tag{5.26}$$

The two terms in the right hand side above are controlled by the following lemmas. Their proofs are presented in the following two subsections.

Lemma 5.17. *We have*

$$\sup_{s\le\delta} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| \le 2^{\beta} \sup_{r\ge t/2} r^{-\beta}\mathbb{E}\left[P_r(\delta)\right] + K_1\delta^{\beta/2}.$$

Lemma 5.18. *There exist constants C_1, C_2, η with $0 < \eta < \beta$ and $\gamma \in (0,1)$ such that, for any integer k, and real number $\delta \in (0, 1/2)$ we have, for any real number $t > 0$,*

$$\sup_{s\in[\delta,1/2]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| \le C_1\delta^{-1}(1-\gamma)^k + C_2 k 2^k (\beta-\eta)^{-2k} t^{-\eta}.$$

Before proceeding with the proofs of the lemmas, we indicate how they imply Proposition 5.16. By Lemmas 5.17 and 5.18, we have for any $\delta \in (0, 1/2)$ and natural number $k \ge 0$

$$\sup_{s\in[0,1]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| \le 3K_1\delta^{\beta/2} + 3C_1\delta^{-1}(1-\gamma)^k + 5C_2 k t^{-\eta} 2^k (\beta-\eta)^{-2k}.$$

Choosing $\delta = t^{-\nu}$ and $k = \lfloor \alpha \log t \rfloor$ for $\nu, \alpha > 0$ to be determined, we obtain

$$\begin{aligned}\sup_{s\in[0,1]} \left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s)\right| \quad &\le \quad 3K_1 t^{-\nu\beta/2} + 3C_1 t^{\nu}(1-\gamma)^{\alpha\log t - 1} \\ &\qquad + 5C_2 t^{-\eta}\left[2/(\beta-\eta)^2\right]^{\alpha\log t}\alpha\log t.\end{aligned}$$

First pick $\alpha > 0$ small enough that

$$\alpha \log\left(\frac{2}{(\beta-\eta)^2}\right) < \eta.$$

This α being fixed, choose $\nu > 0$ small enough that $\nu + \alpha\log(1-\gamma) < 0$. The claim follows.

5.4.1. Behavior along the edge: proof of Lemma 5.17

The bound given by Lemma 5.18 blows up as δ approaches zero. However, as we have already noted in the introduction, $C_n(0)$ is asymptotically of smaller order than at any other fixed query line $0 < s \le 1/2$; the case $s = 0$ should therefore not cause any problems at all. This turns out to be true and we will deal with the term involving the values of $s \in [0, \delta]$ by relating the value $\mathbb{E}\left[P_t(s)\right]$ to $\mathbb{E}\left[P_t(\delta)\right]$. The term $\mathbb{E}\left[P_t(\delta)\right]$ will then be shown to be small choosing δ sufficiently small.

The limit first moment $\bar{\mu}(s) = \lim_{t\to\infty}\mathbb{E}\left[P_t(s)\right]$ is monotonic for $s \in [0, 1/2]$. It seems, at least intuitively, that for any fixed real number $t > 0$, $\mathbb{E}\left[P_t(s)\right]$ should also be monotonic for $s \in [0, 1/2]$, but we were unable to prove it. The following weaker version will be sufficient for our purpose.

Proposition 5.19. *For any $s < 1/2$ and $\varepsilon \in [0, 1-2s)$, we have*

$$\mathbb{E}\left[P_t(s)\right] \le \mathbb{E}\left[P_{t(1+\varepsilon)}\left(\frac{s+\varepsilon}{1+\varepsilon}\right)\right].$$

The idea underlying Proposition 5.19 requires to understand what happens to the quadtree by considering a larger point set. For a finite point set $\mathcal{P} \subset [a,b] \times [0,1] \times [0,\infty)$, we let $V(\mathcal{P})$ and $H(\mathcal{P})$ denote, respectively, the set of vertical and horizontal line segments of the quadtree built from $\mathcal{P}$.

Lemma 5.20. *Let $\mathcal{P} = \{p_1, \dots, p_n\}$ be a set of points with $p_i = (x_i, y_i, t_i) \in [a_2, a_3] \times [0,1] \times [0,\infty)$ ordered by their t coordinate, i.e. $t_i \le t_{i+1}$. Additionally we assume $\mathcal{P}$ to be in general position, meaning that all x-coordinates are pairwise different and the same holds true for the y and t coordinates. Furthermore let $\mathcal{Q} = \{p'_1, \dots, p'_m\} \subseteq [a_1, a_2] \times [0,1] \times [0,\infty)$ with $p'_i = (x'_i, y'_i, t'_i)$ again ordered according to their third coordinate such that $\mathcal{P} \cup \mathcal{Q} \subseteq [a_1, a_3] \times [0,1] \times [0,\infty)$ is again in general position. Then we have*

$$H(\mathcal{P} \cup \mathcal{Q}) \supset H(\mathcal{P}) \qquad \textit{and} \qquad V(\mathcal{P} \cup \mathcal{Q}) \subset V(\mathcal{P}).$$

Proof. We assume for a contradiction that the assertion is wrong and focus on the case that $H(\mathcal{P}) \not\subset H(\mathcal{P} \cup \mathcal{Q})$; the other case is handled analogously. Let i_1 be the index of the "first" point in $\mathcal{P}$ such that the horizontal line of p_{i_1} is shorter (at least on the right or left side of the point) in the quadtree built from $\mathcal{P} \cup \mathcal{Q}$ than it was in the one built from $\mathcal{P}$. Here, first refers to the time coordinate t. By construction, there must be an index i_2 such that the vertical line of p_{i_2} blocks the horizontal line of p_{i_1} in $\mathcal{P} \cup \mathcal{Q}$ but not in $\mathcal{P}$. We again choose i_2 such that t_{i_2} is minimal with this property, by construction $t_{i_2} < t_{i_1}$. Repeating the argument gives the existence of an index i_3 and a point p_{i_3} whose horizontal line blocks the vertical line of p_{i_2} in $\mathcal{P} \cup \mathcal{Q}$ but not in $\mathcal{P}$ with $t_{i_3} < t_{i_2}$. This obviously contradicts the choice of i_1. □

Proof of Proposition 5.19. Consider the unit square $[0,1]^2$ and the extended box $[-\varepsilon, 1] \times [0,1]$, and a single Poisson point process on $[-\varepsilon, 1] \times [0,1] \times [0,t]$ with unit intensity. Write $P_t^\varepsilon(s)$ for the number of (horizontal) lines intersecting $\{x = s\}$ in the quadtree formed by the all the points. Similarly, let $P_t(s) = P_t^0(s)$ be the corresponding quantity when the quadtree is formed using only the points falling inside $[0,1]^2$. Then, for this coupling, we have by Lemma 5.20,

$$P_t(s) \le P_t^\varepsilon(s) \stackrel{d}{=} P_{t(1+\varepsilon)}\left(\frac{s+\varepsilon}{1+\varepsilon}\right).$$

Taking expectations completes the proof. □

Proof of Lemma 5.17. We use Proposition 5.19 to relate $\mathbb{E}\left[P_t(s)\right]$ to $\mathbb{E}\left[P_{t'}(\delta)\right]$ for some t'. Choosing $\varepsilon = (\delta - s)/(1-\delta)$ yields $t' = t(1-s)/(1-\delta) \le t(1-\delta)^{-1}$. Thus, for any $\delta \in (0, 1/2)$ and $t > 0$ we have

$$\begin{aligned}
\sup_{s \le \delta} \left| t^{-\beta} \mathbb{E}\left[P_t(s)\right] - \bar{\mu}(s) \right| &\le \sup_{s \le \delta} t^{-\beta} \mathbb{E}\left[P_t(s)\right] + \bar{\mu}(\delta) \\
&\le \sup_{s \le \delta} t^{-\beta} \mathbb{E}\left[P_{t'}(\delta)\right] + \bar{\mu}(\delta) \\
&\le t^{-\beta} \mathbb{E}\left[P_{t/(1-\delta)}(\delta)\right] + \bar{\mu}(\delta) \\
&\le (1-\delta)^{-\beta} \sup_{r \ge t/2} r^{-\beta} \mathbb{E}\left[P_r(\delta)\right] + \bar{\mu}(\delta).
\end{aligned}$$

This completes the proof since $\delta \le \frac{1}{2}$ and $\bar{\mu}(s) \le K_1 \delta^{\beta/2}$. □

5.4.2. Behavior away from the edge: proof of Lemma 5.18

The core of the work is to bound the second term in (5.26) involving $s \in (\delta, 1/2]$. Our approach relies on tightening the arguments developed in [CJ11]. We start by re-explaining their basic ideas. Observe that most of the quantities defined in the remaining of the subsection will depend on s which we will neglect in the notation for the sake of readability.
The first step is to unfold k levels of the fundamental recurrence (1.18) in the Poisson case. Let τ_1 be the arrival time of the first point in the Poisson process and $Q_1 = Q_1(s)$ be the lower of the two rectangles that intersects the line $\{x = s\}$ after inserting the first point. Inductively let $\tau_k = \tau_k(s)$ be the arrival time of the first point of the process in the region Q_{k-1} and Q_k be the lower of the two rectangles that hit the line $\{x = s\}$ at time τ_k. For convenience, set $Q_0 = [0, 1]^2$. Finally, let $\tilde{P}_t$ be an independent copy of the process P_t (set $\tilde{P}_t \equiv 0$ for $t < 0$). At level one, using the horizontal symmetry, we have

$$\mathbb{E}\left[P_t(s)\right] = \mathbf{P}\left(t \geq \tau_1\right) + 2\mathbb{E}\left[\tilde{P}_{\text{Leb}(Q_1)(t-\tau_1)}(\xi_1)\right],$$

where $\xi_1 = \xi_1(s) \in [0, 1]$ denotes the location of the line $\{x = s\}$ relative to the region Q_1. If the interval $[\ell_1, r_1]$ denotes the projection of Q_1 on the first axis, we have

$$\xi_1(s) = \frac{s - \ell_1}{r_1 - \ell_1}.$$

Write $\xi_k = \xi_k(s) \in [0, 1]$ for the location of the line $\{x = s\}$ relatively to the region Q_k, and $M_k = \text{Leb}(Q_k)$. Then, unfolding k levels, we obtain

$$\mathbb{E}\left[P_t(s)\right] = g_k(t) + 2^k\mathbb{E}\left[\tilde{P}_{M_k(t-\tau_k)}(\xi_k)\right], \tag{5.27}$$

where $0 \leq g_k(t) \leq 2^k - 1$. Next, we introduce the inter-arrival times $\zeta'_k = \tau_k - \tau_{k-1}$ with $\zeta'_0 := 0$

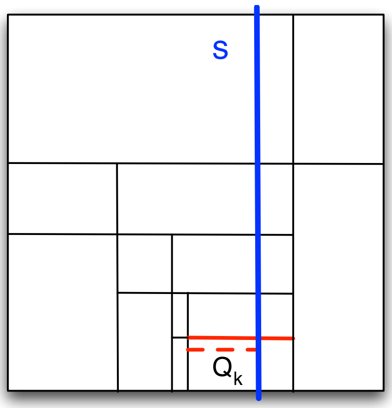


Figure 5.3.: Unfolding k levels of the recursion. $\zeta_k(s)$ equals the quotient of the dashed red line and the solid red line.

and their normalized versions $\zeta_k = \zeta'_k M_{k-1}$ (again $\zeta_0 := 0$). Defining $F_k = M_k\tau_k$, we can rewrite (5.27) as

$$\mathbb{E}\left[P_t(s)\right] = g_k(t) + 2^k\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\right]. \tag{5.28}$$

The sequence $(\zeta_k)_{k\geq 1}$ are i.i.d. exponential random variables with unit mean, also independent of $(\xi_k, Q_k)_{k\geq 1}$. Note that, as we have already seen in Section 5.3, the region Q_k, is not distributed as a typical rectangle at level k; in particular Leb(Q_k) is not distributed as $X_1 Y_1 \cdots X_k Y_k$, for independent $[0,1]$-uniform random variables X_i, Y_i, $i \geq 1$. Intuitively, Q_k should be stochastically larger than a typical cell, since it is conditioned to intersect the line $\{x = s\}$. This is verified by the following lemma. We return to the effect of size-biasing and give more details on the distribution of M_k in Subsection 5.4.4.

Lemma 5.21. *For any $s \in (0,1)$, any integer $k \geq 0$, and $1 \leq i \leq 2^k$, we have, stochastically,*

$$\text{Leb}(Q_k) = M_k \geq X_1 Y_1 \cdots X_k Y_k,$$

where X_i, Y_i, $i \geq 1$ are independent random variables uniform on $[0,1]$.

Proof. Consider one split, at a point (X, Y) uniform inside the unit square. The split creates four new boxes, two of them being hit by s. Let L be the length of these two cells. Their height is either Y or $(1-Y)$, which are both uniform. So it suffices to prove that $L \geq X$ stochastically. By symmetry, it suffices to consider $s \leq 1/2$. We have,

$$L = \mathbf{1}_{\{s \leq X\}} X + \mathbf{1}_{\{s > X\}} (1 - X).$$

Write $F_L(y) = \mathbf{P}(L \leq y)$ and $F_X(y) = \mathbf{P}(X \leq y) = y$. It is then easy to see that

$$F_L(y) = \mathbf{P}(L \leq y) = \begin{cases} 0, & y \leq s \\ y - s, & s \leq y \leq 1 - s \\ 2y - 1, & y \geq 1 - s. \end{cases}$$

Hence, for all $s \in (0, 1/2)$ and all $y \in (0,1)$ we have $F_L(y) \leq y = F_X(y)$. The result follows. $\square$

The second term will be treated using results for the case $s = \xi$, for a uniform random variable ξ independent of everything else. For every $k \geq 1$, the distribution of ξ_k depends on s thus we can not use the result for a uniform query line directly. Curien and Joseph [CJ11] found a very clever way to circumvent this problem. In their Proposition 4.1 they introduce a version of the homogeneous Markov chain $(\xi_k, \mathcal{M}_k)_{k\geq 1}$ where $\mathcal{M}_k := M_k / M_{k-1}$ together with a random time T such that for any $k \in \mathbb{N}$, conditionally on $\{T \leq k\}$, the random variable ξ_k is uniformly distributed on $[0,1]$, independent of $(\mathcal{M}_1, \ldots, \mathcal{M}_k, T)$. Choosing these random variables independent of the process $\tilde{P}_t$ we will use them in the following without changing the notation (F_k can be constructed using $(\mathcal{M}_\ell)_{1\leq \ell \leq k}$ and an additional set of i.i.d. exponential random variables with mean one). The details of the definition of T are not important for us. The only crucial thing is that T has exponential tails. Indeed, we have [CJ11, page 189]

$$\mathbb{E}\left[1.15^T\right] \leq C_4 (s \wedge (1-s))^{-1/2} \leq C_4 \delta^{-1/2}, \tag{5.29}$$

for some constant C_4 in the case of $\delta < s \leq 1/2$. Proceeding as in [CJ11] we do not establish uniform convergence with a suitable rate directly but prove the sequence $(t^{-\beta}\mathbb{E}\left[P_t(s)\right])_{t\geq 0}$ to be Cauchy where we keep track of smaller order terms thoroughly. Using (5.28) and the triangle inequality, we obtain for any t and r such that $r \geq t$,

$$\begin{aligned}
&\left|t^{-\beta}\mathbb{E}\left[P_t(s)\right] - r^{-\beta}\mathbb{E}\left[P_r(s)\right]\right| \qquad (5.30)\\
&\leq 2^k\left|t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\right]\right| + 2^{k+1}r^{-\beta}\\
&\leq 2^k\left|t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T\leq k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T\leq k\}}\right]\right|\\
&+ 2^k\left|t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right]\right| + 2^{k+1}r^{-\beta}.
\end{aligned}$$

To complete the proof of Lemma 5.18, we formulate explicit bounds for the two main terms in (5.30) where we distinguish the cases whether coupling has occurred by level k (i.e. $T \leq k$) or not.

i. No coupling by level k, $T > k$. In this case, we bound the terms roughly. We obtain

$$\begin{aligned}
&2^k\left|t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right]\right|\\
&\leq 2^{k+1}\sup_{u\geq t} u^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k u - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right].
\end{aligned}$$

One then essentially uses the uniform bound $\sup_s \sup_u u^{-\beta}\mathbb{E}\left[P_u(s)\right] \leq C_5$ (see (10) in [CJ11]) Hölder's and Markov's inequalities to make use of a bound that is based on the exponential tails of T. For the details we refer to [CJ11, page 190]. For any $u > 0$ and $s \in (\delta, 1/2]$, one has

$$\begin{aligned}
u^{-\beta}2^k\mathbb{E}\left[\tilde{P}_{M_k u - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right] &\leq C_5 2^k s^{-1/p}\left(\frac{2}{(\beta p+1)(\beta p+2)}\right)^{(k-1)/p}\left(\frac{\mathbb{E}\left[1.15^T\right]}{1.15^k}\right)^{1-1/p}\\
&\leq C_4C_5\delta^{-1/2-1/(2p)}\left(2\left\{\frac{2}{(\beta p+1)(\beta p+2)}\right\}^{1/p}1.15^{1/p-1}\right)^k,
\end{aligned}$$

by the upper bound in (5.29). Choosing p close enough to one that the term in the brackets above is strictly less than one, we obtain for any $s \in (\delta, 1/2]$ and real numbers $t, r > 0$,

$$\begin{aligned}
2^k\left|t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T> k\}}\right]\right| &\leq 2C_4C_5\delta^{-1/2-1/(2p)}(1-\gamma)^k\\
&\leq C_1\delta^{-1}(1-\gamma)^k, \qquad (5.31)
\end{aligned}$$

where C_1 denotes a constant and $\gamma > 0$.

ii. Coupling has occurred before level k, $T \leq k$. In this case, more care has to be taken, we will have to match the first order terms in the expansion (1.14). In what follows, we write $x_+ = x \vee 0$. We start with

$$t^{-\beta}2^k\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T\leq k\}}\right] = 2^k\mathbb{E}\left[\mathbf{1}_{\{T\leq k\}}(M_k - t^{-1}F_k)_+^\beta\theta(M_k t - F_k)\right],$$

where $\theta(x) = x_+^{-\beta}\mathbb{E}\left[P_x(\xi)\right]$ with ξ a $[0,1]$-uniform random variable independent of everything else. The estimate in (1.14) is easily transferred to the poissonized version [the details are similar to the depoissonization in the next subsection], and we have $\theta(x) = \kappa + O(x^{-\eta})$ for any $0 < \eta < \beta$. Therefore

$$\begin{aligned}
&2^k\left|t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T\leq k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T\leq k\}}\right]\right|\\
&\leq 2^k\left|\mathbb{E}\left[\mathbf{1}_{\{T\leq k\}}(M_k - t^{-1}F_k)_+^\beta\theta(M_k t - F_k)\right] - \mathbb{E}\left[\mathbf{1}_{\{T\leq k\}}(M_k - r^{-1}F_k)_+^\beta\theta(M_k r - F_k)\right]\right|\\
&\leq 2^k\mathbb{E}\left[\left|(M_k - t^{-1}F_k)_+^\beta\theta(M_k t - F_k) - (M_k - r^{-1}F_k)_+^\beta\theta(M_k r - F_k)\right|\right]. \qquad (5.32)
\end{aligned}$$

Fix $\eta < \beta$. For $x > 0$, we have, as $x \to \infty$

$$\begin{aligned}
&(M_k - x^{-1}F_k)_+^{\beta} \cdot \theta(M_k x - F_k) \\
&= M_k^{\beta}(1 - O(x^{-1}F_k M_k^{-1}))(\kappa + O(M_k^{-\eta}x^{-\eta})) \\
&= \kappa M_k^{\beta} + O(F_k M_k^{\beta-1}x^{-1}) + O(M_k^{\beta-\eta}x^{-\eta}) + O(F_k M_k^{\beta-1-\eta}x^{-1-\eta}) \\
&= \kappa M_k^{\beta} + O(F_k M_k^{\beta-1}x^{-1}) + O(x^{-\eta}) + O(F_k M_k^{\beta-1-\eta}x^{-1-\eta}),
\end{aligned}$$

since $M_k \in (0,1)$ and $\eta < \beta$, the O terms being deterministic and uniform in $s \in [0,1]$. Going back to (5.32), the terms κM_k^{β} coming from the two terms with t and r cancel out, and there exist constants C_7, C_8 such that, for all t, r large enough such that moreover $t \le r$, we have

$$\begin{aligned}
&2^k \left| t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T\le k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T\le k\}}\right] \right| \\
&\le C_7 2^k \left(t^{-1}\mathbb{E}\left[F_k M_k^{\beta-1}\right] + t^{-\eta} + t^{-1-\eta}\mathbb{E}\left[F_k M_k^{\beta-1-\eta}\right]\right) \\
&\le C_8 2^k t^{-\eta}\mathbb{E}\left[F_k M_k^{\beta-1-\eta}\right].
\end{aligned}$$

Since it will be necessary to choose k tending to infinity with r to control the term in (5.31), it remains to estimate $\mathbb{E}\left[F_k M_k^{\beta-1-\eta}\right]$. By definition of $F_k = M_k \tau_k$, one immediately sees that $F_k \le \sum_{\ell=1}^{k} \zeta_\ell$, where the normalized inter-arrival times ζ_ℓ were defined right after equation (5.27). This rough bound together with Lemma 5.21 implies

$$\begin{aligned}
\mathbb{E}\left[F_k M_k^{\beta-1-\eta}\right] &\le k\mathbb{E}\left[M_k^{\beta-1-\eta}\right] \\
&\le k\mathbb{E}\left[X^{\beta-1-\eta}\right]^{2k} = k(\beta-\eta)^{-2k}.
\end{aligned}$$

Here X denotes a uniform on $[0,1]$. Note that an slightly improved upper bound for moments of M_k is given in [CJ11, Section 4.2], the advance being inessential. We finally obtain

$$2^k \left| t^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k t - F_k}(\xi_k)\mathbf{1}_{\{T\le k\}}\right] - r^{-\beta}\mathbb{E}\left[\tilde{P}_{M_k r - F_k}(\xi_k)\mathbf{1}_{\{T\le k\}}\right] \right| \le C_8 k t^{-\eta} 2^k (\beta-\eta)^{-2k}. \tag{5.33}$$

Putting (5.31) and (5.33) together with (5.30) yields, for any $t, r > 0$ such that $t \le r$

$$\begin{aligned}
\left| t^{-\beta}\mathbb{E}\left[P_t(s)\right] - r^{-\beta}\mathbb{E}\left[P_r(s)\right] \right| &\le C_1\delta^{-1}(1-\gamma)^k + C_8 k 2^k(\beta-\eta)^{-2k}t^{-\eta} + 2^{k+1}t^{-\beta} \\
&\le C_1\delta^{-1}(1-\gamma)^k + C_2 k 2^k(\beta-\eta)^{-2k}t^{-\eta}.
\end{aligned}$$

for some constant C_2. The statement in Lemma 5.18 follows readily from the triangle inequality.

5.4.3. Depoissonization

The depoissonization relies on standard arguments based on the concentration of Poisson random variables and the monotonicity of $\mathbb{E}\left[C_n(s)\right]$ in n for each $s \in (0,1)$.
We first give a standard concentration bound for Poisson distribution that we will use.

Lemma 5.22. *Let N be Poisson(t). Then, there exists $\delta_0 > 0$ such that for every $\delta \in (0, \delta_0)$ and every $t > 0$*

$$\mathbf{P}(|N - t| \ge t\delta) \le 2e^{-t\delta^2/3}.$$

Proof. The standard Chernoff bound for N is

$$\mathbf{P}\left(N > t(1+\delta)\right) \le e^{t\delta - t(1+\delta)\log(1+\delta)}, \quad \text{and} \quad \mathbf{P}\left(N < t(1-\delta)\right) \le e^{-t\delta - t(1-\delta)\log(1-\delta)}.$$

Using $\log(1+x) = x - x^2/2 + O(x^3)$ for $x \to 0$ shows that both tails are bounded by $e^{-t\delta^2/3}$ for sufficiently small δ. This gives the result. □

Recall that we not only need to prove $n^{-\beta}\mathbb{E}\left[C_n(s)\right] \to \bar{\mu}(s)$, uniformly for $s \in (0,1)$, we also want to conserve the polynomial error rate. We first focus on the upper bound. Write $\varepsilon_n = n^{-1/3}$ and let $N \sim \text{Poisson}(n(1+\varepsilon_n))$ be independent of the process building up the discrete quadtree. Then $C_N(s) \stackrel{d}{=} P_{n(1+\varepsilon_n)}(s)$. By monotonicity, we have

$$\begin{aligned}
\mathbb{E}\left[C_n(s)\right] &= \mathbb{E}\left[C_n(s)\mathbf{1}_{\{N\ge n\}}\right] + \mathbb{E}\left[C_n(s)\mathbf{1}_{\{N<n\}}\right] \\
&\le \mathbb{E}\left[C_N(s)\mathbf{1}_{\{N\ge n\}}\right] + \mathbb{E}\left[C_n(s)\mathbf{1}_{\{N<n\}}\right] \\
&\le \mathbb{E}\left[C_N(s)\right] + \mathbb{E}\left[C_n(s)\mathbf{1}_{\{N<n\}}\right] \\
&\le \mathbb{E}\left[C_N(s)\right] + n\mathbf{P}(N<n),
\end{aligned}$$

since $C_n(s) \le n$. For $t = n(1+\varepsilon_n)$ and $\delta = \varepsilon_n/2$, we have $t(1-\delta) = n(1+\varepsilon_n)(1-\varepsilon_n/2) = n(1+\varepsilon_n/2 + o(\varepsilon_n)) \ge n$, for all n large enough. It follows from Lemma 5.22, for all n large enough,

$$\begin{aligned}
\mathbb{E}\left[C_n(s)\right] &\le \mathbb{E}\left[C_N(s)\right] + ne^{-n(1+\varepsilon_n)\varepsilon_n^2/3} \\
&\le \mathbb{E}\left[C_N(s)\right] + e^{-n^{1/3}/12}.
\end{aligned}$$

Therefore, for any $s \in [0,1]$,

$$\begin{aligned}
n^{-\beta}\mathbb{E}\left[C_n(s)\right] - \bar{\mu}(s) \quad &le \quad n^{-\beta}\mathbb{E}\left[C_N(s)\right] - \bar{\mu}(s) + n^{-\beta}e^{-n^{1/3}/12} \\
&= \quad (1+\varepsilon_n)^{\beta}[n(1+\varepsilon_n)]^{-\beta}\mathbb{E}\left[C_N(s)\right] - \bar{\mu}(s) + n^{-\beta}e^{-n^{1/3}/12} \\
&\le \quad [n(1+\varepsilon_n)]^{-\beta}\mathbb{E}\left[C_N(s)\right] - \bar{\mu}(s) \\
&\quad\quad +\varepsilon_n[n(1+\varepsilon_n)]^{-\beta}\mathbb{E}\left[C_N(s)\right] + n^{-\beta}e^{-n^{1/3}/12}. \qquad (5.34)
\end{aligned}$$

Similarly, we can obtain a lower bound using $N' \sim \text{Poisson}(n(1-\varepsilon_n))$, again independent of the discrete process. We obtain

$$\begin{aligned}
\mathbb{E}\left[C_n(s)\right] &= \mathbb{E}\left[C_n(s)\mathbf{1}_{\{N'>n\}}\right] + \mathbb{E}\left[C_n(s)\mathbf{1}_{\{N'\le n\}}\right] \\
&\ge \mathbb{E}\left[C_{N'}(s)\mathbf{1}_{\{N'\ge n\}}\right] \\
&= \mathbb{E}\left[C_{N'}(s)\right] - \mathbb{E}\left[C_{N'}(s)\mathbf{1}_{\{N'<n\}}\right] \\
&\ge \mathbb{E}\left[C_{N'}(s)\right] - n\mathbf{P}\left(N'<n\right).
\end{aligned}$$

We again aim at using Lemma 5.22. Set $t = n(1-\varepsilon_n)$ and $\delta = \varepsilon_n$. Then $t(1+\delta) = n(1-\varepsilon_n^2) \le n$ so that

$$\begin{aligned}
\mathbb{E}\left[C_n(s)\right] &\ge \mathbb{E}\left[C_{N'}(s)\right] - ne^{-n(1-\varepsilon_n)\varepsilon_n^2/3} \\
&\ge \mathbb{E}\left[C_{N'}(s)\right] - e^{-n^{1/3}/12},
\end{aligned}$$

for all n large enough. It follows that, for any $s \in [0,1]$,

$$\begin{aligned} n^{-\beta}\mathbb{E}\left[C_n(s)\right] - \bar{\mu}(s) &\geq (1-\varepsilon_n)[n(1-\varepsilon_n)]^{-\beta}\mathbb{E}\left[C_{N'}(s)\right] - \bar{\mu}(s) - n^{-\beta}e^{-n^{1/3}/12} \\ &\geq [n(1-\varepsilon_n)]^{-\beta}\mathbb{E}\left[C_{N'}(s)\right] - \bar{\mu}(s) - n^{-\beta}e^{-n^{1/3}/12}. \end{aligned} \tag{5.35}$$

Finally, using $C_N(s) \stackrel{d}{=} P_{n(1+\varepsilon_n)}(s), C_{N'}(s) \stackrel{d}{=} P_{n(1-\varepsilon_n)}(s)$, putting (5.34) and (5.35) together, and using Proposition 5.16, we obtain

$$\sup_{s\in[0,1]} \left| n^{-\beta}\mathbb{E}\left[C_n(s)\right] - \bar{\mu}(s) \right| = O(n^{-\varepsilon}),$$

where ε is given in Proposition 5.16. Hence the proof of Proposition 5.7 is complete.

5.4.4. Extensions to the limit mean

The last subsection on the asymptotic behaviour of the mean of $C_n(s)$ is dedicated to a better understanding of Proposition 5.16 and its proof. A key ingredient is the concept of size-biasing as already emphasized in Lemma 5.21. For simplicity, we fix $s \in (0,1)$ and do not face issues of uniformity. Summarizing the approach of Curien and Joseph [CJ11], they first show that $t^{-\beta}\mathbb{E}\left[P_t(s)\right]$ admits a finite limit $\bar{\mu}(s)$ as $t \to \infty$ by verifying the Cauchy property. To infer the shape of the limit, they make use of a fixed-point characterization of $\bar{\mu}(s)$, see [CJ11, page 191]. It is easy to see that this equation is solved by any multiple of $(s(1-s))^{\beta/2}$. Finally, they are able to prove that this fixed-point equation has at most one solution up to a normalization factor. In the last step, using the results for the case of a uniform query line from [FGPR93], Curien and Joseph determine the precise value of $\bar{\mu}(s)$. Our result in this section extends their ideas to uniform convergence on the unit interval. We attempt to give further insight for the occurrence of $\bar{\mu}(s)$ here. Using Lemma 5.18 and the techniques of its proof, there exist constants $\bar{C}_1, \bar{C}_2, \bar{C}_3$ [which now may depend on s] such that for any $t > 0, k \in \mathbb{N}$

$$t^{-\beta}\mathbb{E}\left[P_t(s)\right] = f_1(t) + \kappa 2^k \mathbb{E}\left[M_k^\beta\right],$$

with

$$|f_1(t)| \leq \bar{C}_1 t^{-\beta} 2^k + \bar{C}_2(1-\gamma)^k + \bar{C}_3 k 2^k t^{-\eta}(\beta-\eta)^{-2k}.$$

Hence the term which gives rise to the limit $\bar{\mu}(s)$ is $\kappa 2^k \mathbb{E}\left[M_k^\beta\right]$; more precisely

$$|\kappa 2^k \mathbb{E}\left[M_k^\beta\right] - \bar{\mu}(s)| \leq \bar{C}_3(1-\gamma)^k \qquad \text{and} \qquad \kappa 2^k \mathbb{E}\left[M_k^\beta\right] \to \bar{\mu}(s),$$

as $k \to \infty$. It is much easier to analyze the term M_k in the uniform case $s = U$. Then, using the notation and the result of Proposition 1 in [CJ11], it holds $M_k(U) = \mathcal{M}_1(U)\cdots\mathcal{M}_k(U)$ where $(\mathcal{M}_k(U))_{k\geq 1}$ are i.i.d. random variables with density $2(1-m)\mathbf{1}_{\{m\in[0,1]\}}$. This has already been explained at the beginning of this chapter; the length of the interval covering U has size-biased uniform distribution whereas the height is unbiased. Moreover, the product of two independent random variables with uniform and size-biased uniform distribution on the unit interval has density $2(1-m)\mathbf{1}_{\{m\in[0,1]\}}$. In particular, we have $2^k\mathbb{E}\left[M_k^\beta(U)\right] = 1$ according to the very definition of β. Here we could also turn things around and use $2^k\mathbb{E}\left[M_k^\beta(U)\right] = 1$ as a heuristic for the precise value of β.

Let us return to the non-uniform case where $s \in [0,1]$ is fixed. Denoting $L_k = L_k(s)$ the length of the interval on the x-axis covering s after k iterations, we have $M_k = L_k \prod_{i=1}^k X_i$ where $X_1, \dots, X_k$ are i.i.d. random variables with uniform distribution on $[0,1]$ and independent of L_k. Therefore, the problem of analyzing the mean (or moments) of M_k is actually one-dimensional and we could have focussed on the quantity L_k instead of M_k throughout the proofs of the previous subsections. However, we decided not to do so in order to apply the results from [CJ11] without modifications.

The computation of the limit mean of L_k after rescaling can be worked out as in [CJ11, Section 5] based on its distributional recurrence

$$(L_k(s))_{s\in[0,1]} \stackrel{d}{=} \left(\mathbf{1}_{\{s<X\}} X L_{k-1}^{(1)}\left(\frac{s}{X}\right) + \mathbf{1}_{\{s\geq X\}}(1-X)L_{k-1}^{(2)}\left(\frac{1-s}{1-X}\right)\right)_{s\in[0,1]},$$

where, $L_{k-1}^{(1)}, L_{k-1}^{(2)}$ are independent copies of L_{k-1}, X is uniform on $[0,1]$ and independent of $(L_{k-1}^{(1)}, L_{k-1}^{(2)})$. As a conclusion, we can say that the function $(s(1-s))^{\beta/2}$ results from the difference between the size-biasing effects in $L_k(s)$ and $L_k(U)$, we have

$$\lim_{k\to\infty} \frac{\mathbb{E}\left[L_k^\beta(s)\right]}{\mathbb{E}\left[L_k^\beta(U)\right]} = \frac{(s(1-s))^{\beta/2}}{\mathrm{B}(\beta/2+1, \beta/2+1)},$$

where one should keep in mind that $\mathbb{E}\left[L_k(U)\right] = \left(\frac{2}{3}\right)^k$.

We finally face the scaling behaviour of L_k (or M_k) on the distributional level. Again, we start with the case $s = U$ where U is uniform and independent of the process. Being a product of non-negative i.i.d. random variables with mean one yields

$$\left(\frac{3}{2}\right)^k L_k(U) \to 0 \tag{5.36}$$

almost surely as $k \to \infty$. Choosing the right scaling of $L_k(U)$ still leads to a degenerate limit due to the lack of uniform integrability. To obtain a non-degenerate limit, on might instead consider $\log L_k(U)$. In distribution, $-\log L_k(U)$ equals a sum of independent exponential random variables with parameter 2. This implies

$$\frac{\log L_k(U)}{k} \to -\frac{1}{2}$$

almost surely and

$$\frac{\log L_k(U) + \frac{k}{2}}{\sqrt{k/4}} \to N(0,1)$$

in distribution. The convergence in (5.36) carries over to the case of a fixed $s \in (0,1)$. By independence of U and L_k, for almost every $s \in (0,1)$, we have $\left(\frac{3}{2}\right)^k L_k(s) \to 0$ almost surely.

5.5. The marginals of the limit process

Our main result implies the convergence of the second moment of the discrete towards that of the limit process. This section is devoted to identifying this limit, in particular it provides an explicit expression for the limit variance. The following Proposition is a detailed version of Theorem 5.5 that also covers the additional statements on the variance in Theorem 5.2.

Proposition 5.23. *Let $Z(s)$ be the process constructed in Section 5.3 with mean $h(s)$. Furthermore, let Z be the unique solution of the fixed-point equation*

$$Z \stackrel{d}{=} U^{\beta/2}V^{\beta}Z + U^{\beta/2}(1-V)^{\beta}Z',$$

with $\mathbb{E}[Z]=1$ and $\mathbb{E}\left[Z^2\right]<\infty$, where Z' is an independent copy of Z and (Z,Z') is independent of (U,V). Then, for any $s\in[0,1]$,

$$Z(s) \stackrel{d}{=} Z\cdot(s(1-s))^{\beta/2}. \tag{5.37}$$

The sequence of moments $c_m=\mathbb{E}[Z^m]$ satisfies the recursion

$$c_m = \frac{\beta m+1}{(m-1)\left(m+1-\frac{3}{2}\beta m\right)}\sum_{\ell=1}^{m-1}\binom{m}{\ell}\mathrm{B}(\beta\ell+1,\beta(m-\ell)+1)c_\ell c_{m-\ell}, \tag{5.38}$$

for $m\geq 2$ where $c_1=1$. In particular,

$$\mathbf{Var}\,(Z(s)) = K_2h^2(s) = \left[2\mathrm{B}\,(\beta+1,\beta+1)\,\frac{2\beta+1}{3(1-\beta)}-1\right]h^2(s), \tag{5.39}$$

and for ξ uniformly distributed on $[0,1]$, and independent of Z,

$$\mathbf{Var}\,(Z(\xi)) = K_3 := \frac{2(2\beta+1)}{3(1-\beta)}(\mathrm{B}\,(\beta+1,\beta+1))^2-\left(\mathrm{B}\left(\frac{\beta}{2}+1,\frac{\beta}{2}+1\right)\right)^2. \tag{5.40}$$

Remark 5.24. It's worth noting that the random variable Z also appeared in [NR01] where the false distributional limit law

$$\frac{C_n(\xi)}{\kappa n^\beta}\to Z$$

is stated. In fact, our result reveals that

$$\frac{C_n(\xi)}{\kappa n^\beta}\to Z\frac{(\xi(1-\xi))^{\beta/2}}{B\left(\frac{\beta}{2}+1,\frac{\beta}{2}+1\right)}$$

where ξ is uniform on the unit interval and independent of Z. Thus, compared to the former incorrect result, the limit contains an additional independent multiplicative term $h(\xi)$ scaled to have unit mean.

Proof. The definition of the process $Z(s)$ implies that the second moment $\mu_2(s)=\mathbb{E}\left[Z(s)^2\right]$ satisfies an integral equation. We have

$$\begin{aligned}
\mu_2(s)=&\mathbb{E}\left[Z(s)^2\right] = 2\mathbb{E}\left[\int_s^1 (xY)^{2\beta}Z\left(\frac{s}{x}\right)^2dx\right]+2\mathbb{E}\left[\int_0^s[(1-x)Y]^{2\beta}Z\left(\frac{1-s}{1-x}\right)dx\right]\\
&+2\mathbb{E}\left[\int_s^1 x^{2\beta}[Y(1-Y)]^\beta Z^{(1)}\left(\frac{s}{x}\right)\cdot Z^{(2)}\left(\frac{s}{x}\right)dx\right]\\
&+2\mathbb{E}\left[\int_0^s (1-x)^{2\beta}[Y(1-Y)]^\beta Z^{(1)}\left(\frac{1-s}{1-x}\right)Z^{(2)}\left(\frac{1-x}{1-x}\right)dx\right]\\
=&2\mathbb{E}\left[Y^{2\beta}\right]\left\{\int_s^1 x^{2\beta}\cdot\mu_2\left(\frac{s}{x}\right)dx+\int_0^s(1-x)^{2\beta}\cdot\mu_2\left(\frac{1-s}{1-x}\right)dx\right\}\\
&+2\mathbb{E}\left[[Y(1-Y)]^\beta\right]\cdot\left\{\int_s^1 x^{2\beta}h\left(\frac{s}{x}\right)^2dx+\int_0^s(1-x)^{2\beta}h\left(\frac{1-s}{1-x}\right)^2dx\right\}.
\end{aligned}$$

It now follows that μ_2 satisfies the following integral equation

$$\begin{aligned}\mu_2(s) \quad = \quad & \frac{2}{2\beta+1}\left\{\int_s^1 x^{2\beta}\mu_2\left(\frac{s}{x}\right)dx + \int_0^s (1-x)^{2\beta}\mu_2\left(\frac{1-s}{1-x}\right)dx\right\} \\ & +2B\left(\beta+1,\beta+1\right)\cdot\frac{h^2(s)}{\beta+1}.\end{aligned}$$

One easily verifies that the function f given by $f(s) = c_2h^2(s)$ solves the above equation when c_2 satisfies

$$c_2 = \frac{2}{(2\beta+1)(\beta+1)}c_2 + 2\frac{B\left(\beta+1,\beta+1\right)}{\beta+1}.$$

We obtain, after the simplification using $\beta^2 = 2-3\beta$,

$$c_2 = 2B\left(\beta+1,\beta+1\right)\frac{2\beta+1}{3(1-\beta)}. \tag{5.41}$$

It now suffices to prove that the integral equation for μ_2 admits a unique solution. To this end, we show that the map K defined below is a contraction for the supremum norm:

$$\begin{aligned}Kf(s) = & \frac{2}{2\beta+1}\left\{\int_s^1 x^{2\beta}f\left(\frac{s}{x}\right)dx + \int_0^s (1-x)^{2\beta}f\left(\frac{1-s}{1-x}\right)dx\right\} \\ & + 2B\left(\beta+1,\beta+1\right)\frac{h^2(s)}{\beta+1}.\end{aligned} \tag{5.42}$$

For any two functions f and g from $\mathcal{C}[0,1]$, we have

$$\begin{aligned}&\|Kf-Kg\| \\ &\quad\leq \frac{2}{2\beta+1}\left(\sup_{s\in[0,1]}\left\{\int_s^1 x^{2\beta}dx\right\} + \sup_{s\in[0,1]}\left\{\int_0^s (1-x)^{2\beta}dx\right\}\right)\|f-g\| \\ &\quad= \frac{4}{(2\beta+1)^2}\|f-g\|.\end{aligned}$$

Since $2\beta+1>2$, the operator K is a contraction on $\mathcal{C}[0,1]$ equipped with the supremum norm. Banach fixed point theorem then ensures that the fixed point is unique, which shows that indeed

$$\mathbb{E}\left[Z(s)^2\right] = c_2h^2(s).$$

Then, $K_2 = c_2 - 1$ and by integration

$$\mathbf{Var}\left[Z(\xi)\right] = c_2B\left(\beta+1,\beta+1\right) - \left(B\left(\frac{\beta}{2}+1,\frac{\beta}{2}+1\right)\right)^2.$$

Analogously one shows that the m-th moment of $Z(s)$ is of the form $c_mh^m(s)$ where c_m solves (5.38). The Lipschitz constant of the corresponding operator in (5.42) is $4/(\beta m+1)^2$, hence again smaller than one. This immediately implies that $(c_m)_{m\geq1}$ are the moments of $Z(s)/h(s)$, independently of s.

It only remains to prove that there is only one distribution with these moments. We prove that there exists $A_1>0$ such that

$$c_m \leq A_1^m m^m, \qquad m\geq 1, \tag{5.43}$$

which completes the proof of the proposition by the Carleman condition [see, e.g. Fel71, page 228].
Suppose that (5.43) is satisfied for all $m < m_0$. By Stirling's formula, there exists a constant A_2 such that for all $m \geq 1$ and $1 \leq \ell < m$

$$\binom{m}{\ell} \mathrm{B}(\beta\ell + 1, \beta(m-\ell)+1) \leq \frac{A_2}{m} \left(\frac{\ell^\ell (m-\ell)^{m-\ell}}{m^m} \right)^{\beta-1}.$$

Next, the prefactor in (5.38) is of order $1/m$, hence bounded by A_3/m for some $A_3 > 0$ and all $m > 1$. Using this, the induction hypothesis and $x^x(1-x)^{1-x} \leq 1$ for all $x \in [0,1]$, it follows that

$$\begin{aligned} c_{m_0} &\leq \frac{A_2 A_3}{m_0^2} \sum_{\ell=1}^{m_0-1} \left(\ell^\ell (m_0 - \ell)^{m_0-\ell} \right)^{\beta-1} m_0^{m_0(1-\beta)} c_\ell c_{m_0-\ell} \\ &\leq \frac{A_1^{m_0} A_2 A_3}{m_0^2} \sum_{\ell=1}^{m_0-1} m_0^{\beta m_0} m_0^{m_0(1-\beta)} \\ &\leq A_1^{m_0} m_0^{m_0} \end{aligned}$$

if m_0 is chosen large enough. □

5.6. Partial match queries in random 2-d trees

The random 2-d tree was introduced by Bentley [Ben75] and is used to store two-dimensional data just as the two-dimensional quadtree. It is also called two-dimensional binary search tree since it is binary and mimics the construction rule of binary search tree for two-dimensional data. Our aim in this section is to introduce 2-d trees, and extend to 2-d trees the results for partial match queries in quadtrees we obtained in the previous sections. All the results such as process convergence, convergence of all moments at one fixed or random point or at multiple points and distributional and moment convergence of the supremum can be transferred. We will mainly state the forms of the theorems for 2-d trees and focus on the points that deserve additional verifications.

5.6.1. Constructions and basic properties

Construction of 2-d trees. As in quadtrees, the data field is partitioned recursively, but the splits are only binary; since the data is two-dimensional, one alternates between vertical and horizontal splits, depending on the parity of the level in the tree. More precisely, consider a point sequence $p_1, p_2, \ldots, p_n \in [0,1]^2$. Initially, the root is associated with the whole square $[0,1]^2$. The first point p_1 is stored at the root, and splits *vertically* the unit square in two rectangles, which are associated with the two children of the root. More generally, when i points have already been inserted, the tree has i internal nodes, and $i+1$ (lower level) regions associated with the external nodes, forming a partition of the unit square. When point p_{i+1} is stored in the node, say u, corresponding to the region it falls in, it divides the region in two subrectangles that are associated to the two children of u, which become external nodes. The last partition step depends on the parity of the depth of u in the tree: If u has odd distance to the root we partition horizontally, otherwise vertically. Equally likely, one could start at the root with a horizontal split. Then, splits are performed

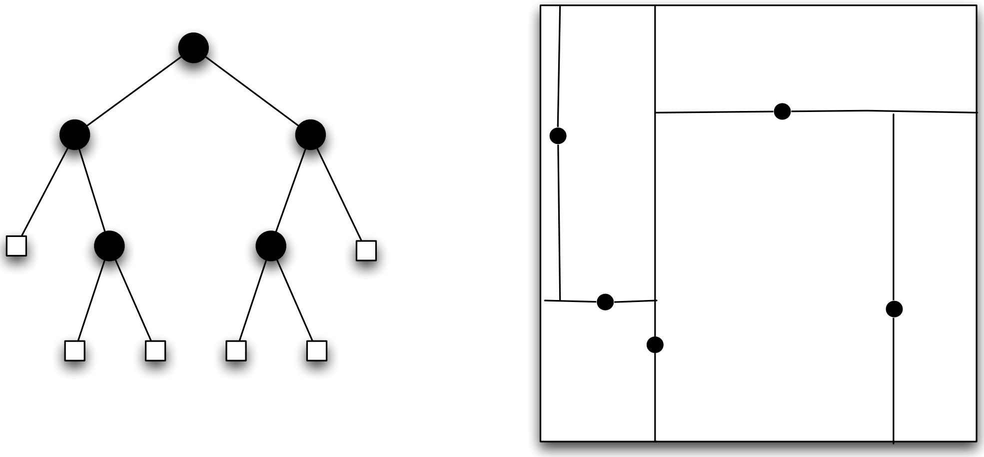


Figure 5.4.: 2-d tree with n nodes and vertical split at the root.

horizontally on levels with even and vertically on levels with odd parity.

Unlike the case of quadtrees, K-d trees remain binary for higher dimensions $d > 2$. Then, during the procedure of inserting a new item in the tree, one compares its s-th component with the corresponding one in an internal node on level ℓ if $\ell = s + kd$ where $k \in \mathbb{N}$.

Partial match queries. From now on, we assume that the data consists of a set of independent random points, uniformly distributed on the unit square. Unlike in the case of quadtrees, the *direction* of a query line with respect to the direction of the root does matter. Let $C_n^{=}(t)$ and $C_n^{\perp}(t)$ denote the number of nodes visited by a partial match for a query at position $t \in [0, 1]$ when the directions of the split at the root and the query are parallel and perpendicular, respectively. Subsequently, we will analyze both quantities synchronously as far as possible. We will always consider directions with respect to the query line, and although some of the expressions (for the sizes of the regions for instance) will be symmetric, we keep them distinct for the sake of clarity. Without loss of generality we assume the query line to be vertical, and that the direction of the cut at the root may change.
As in a quadtree, a node is visited by a partial match query if and only if it is inserted in a subregion that intersects the query line. Unfortunately, these nodes are not easily identifiable *after* the insertion of n points: the value of the quantity $C_n^{=}(s)$ is obtained by adding twice the number of lines intersecting the query line at s to the number of boxes which are intersected by the query line and about to be split perpendicular to the line in the next step [that is, the depth of the corresponding external nodes in the tree has odd parity].

Recursive decompositions. Let (U, V) be the first point which partitions the unit square. By construction, since the directions of the partitioning lines alternate, both processes $C_n^{=}(t)$ and $C_n^{\perp}(t)$ are coupled. When the query line is perpendicular to the split direction, the recursive search proceeds in both subregions whose sizes we denote by N_n and D_n, and we have

$$C_n^{\perp}(s) \stackrel{d}{=} 1 + C_{N_n}^{(=,1)}(s) + C_{S_n}^{(=,2)}(s). \tag{5.44}$$

When the query line and the first split at the root are parallel, only one of the subregions (of sizes L_n and R_n) is relevant for the remainder of the retrieval algorithm and we have

$$C_n^{\perp}(s) \stackrel{d}{=} 1 + \mathbf{1}_{\{s<U\}} C_{L_n}^{(=,1)}\left(\frac{s}{U}\right) + \mathbf{1}_{\{s\geq U\}} C_{R_n}^{(=,2)}\left(\frac{s-U}{1-U}\right). \tag{5.45}$$

Here $(C_n^{(=,1)})_{n\geq 0}, (C_n^{(=,2)})_{n\geq 0}$ are independent copies of $(C_n^{=})_{n\geq 0}$, independent of (N_n, S_n) in (5.44) and $(C_n^{(\perp,1)})_{n\geq 0}, (C^{(\perp,2)})_{n\geq 0}$ are independent copies of $(C_n^{\perp})_{n\geq 0}$, independent of (L_n, R_n) in (5.45). Moreover, here and in the following distributional recurrences and fixed-point equations involving a parameter $s \in [0,1]$ are to be understood on the level of càdlàg or continuous functions unless stated otherwise.
As in the case of partial match in random quadtrees, the expected value at a random uniform query line ξ, independent of the tree, is of order n^{β} for the same constant β defined in (1.15), and we have

$$\mathbb{E}\left[C_n^{=}(\xi)\right] \sim \kappa_{=} n^{\beta}, \qquad \mathbb{E}\left[C_n^{\perp}(\xi)\right] \sim \kappa_{\perp} n^{\beta},$$

for some constants $\kappa_{=} > 0, \kappa_{\perp} > 0$. This was first proved by Flajolet and Puech [FP86]. A more detailed analysis by Chern and Hwang [CH06] shows that

$$\mathbb{E}\left[C_n^{=}(\xi)\right] = \kappa_{=} n^{\beta} - 2 + O(n^{\beta-1}), \qquad \kappa_{=} = \frac{13(3-5\beta)}{4} \cdot \frac{\Gamma(2\beta+2))}{\Gamma^3(\beta+1)}, \tag{5.46}$$

$$\mathbb{E}\left[C_n^{\perp}(\xi)\right] = \kappa_{\perp} n^{\beta} - 3 + O(n^{\beta-1}), \qquad \kappa_{\perp} = \frac{13(2\beta-1)}{2} \cdot \frac{\Gamma(2\beta+2))}{\Gamma^3(\beta+1)}. \tag{5.47}$$

Observe that $\kappa_{=} = \frac{1}{2} 13(3-5\beta)\kappa$ and $\kappa_{\perp} = 13(2\beta-1)\kappa$, where κ is the leading constant for $\mathbb{E}\left[C_n(\xi)\right]$ in the case of quadtrees defined in (1.15). Note that both $\kappa_{=}$ and $\kappa_{\perp}$ are larger than κ.

Two-step recursions and limit behaviour. For our purposes, and although yielding more complex expressions, it is more convenient to expand the recursion one more level. Thus we obtain recursive relations that only involve quantities of the same type, only $(C_n^{=})_{n\geq 0}$ or only $(C_n^{\perp})_{n\geq 0}$. This follows since in both cases each of the first two subregions at the root is eventually split, and this gives rise two a partition into four regions at level two of the tree. Let (U_ℓ, V_ℓ) and (U_r, V_r) be respectively the first points on each side (left and right) of the first cut, when it is parallel to the query line. Let also (U_u, V_u) and (U_d, V_d) be the first points on each side of the cut (up and down) when it is perpendicular to the query line. Note that U, V_ℓ, V_r are independent and uniform on $[0,1]$, and so are V, U_u and U_d.
Let $I_{=,1}^{(n)}, \ldots, I_{=,4}^{(n)}$ and $I_{\perp,1}^{(n)}, \ldots, I_{\perp,4}^{(n)}$ denote the number of data points falling in these regions when the root and the query line are parallel and perpendicular respectively. The distributions of $I_{=,1}^{(n)}, \ldots, I_{=,4}^{(n)}$ and $I_{\perp,1}^{(n)}, \ldots, I_{\perp,4}^{(n)}$ are slightly more involved than in the case of quadtrees. One has e.g. given the values of U, V_ℓ, V_r

$$I_{=,1}^{(n)} \stackrel{d}{=} \mathrm{Bin}((\mathrm{Bin}(n-1;U)-1)_+, V_\ell),$$

and given V, U_d, U_u

$$I_{\perp,1}^{(n)} \stackrel{d}{=} \mathrm{Bin}((\mathrm{Bin}(n-1;V)-1)_+, U_d),$$

where the inner and outer binomials are independent. Analogous expressions hold true for the remaining quantities.

Substituting (5.44) and (5.45) into each other gives

$$\begin{aligned} C_n^{=}(s) \stackrel{d}{=} 1 + \mathbf{1}_{\{s<U\}} &\left[\mathbf{1}_{\{L_n>0\}} + C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right) + C^{(=,2)}_{I^{(n)}_{=,2}}\left(\frac{s}{U}\right) \right] \\ + \mathbf{1}_{\{s\ge U\}} &\left[\mathbf{1}_{\{R_n>0\}} + C^{(=,3)}_{I^{(n)}_{=,3}}\left(\frac{s-U}{1-U}\right) + C^{(=,4)}_{I^{(n)}_{=,4}}\left(\frac{s-U}{1-U}\right) \right] \end{aligned} \tag{5.48}$$

and

$$\begin{aligned} C_n^{\perp}(s) \stackrel{d}{=} 1 + \mathbf{1}_{\{S_n>0\}} + \mathbf{1}_{\{N_n>0\}} + \mathbf{1}_{\{s<U_d\}} C^{(\perp,1)}_{I^{(n)}_{\perp,1}}\left(\frac{s}{U_d}\right) + \mathbf{1}_{\{s<U_u\}} C^{(\perp,2)}_{I^{(n)}_{\perp,2}}\left(\frac{s}{U_u}\right) \\ + \mathbf{1}_{\{s\ge U_d\}} C^{(\perp,3)}_{I^{(n)}_{\perp,3}}\left(\frac{s-U_d}{1-U_d}\right) + \mathbf{1}_{\{s\ge U_u\}} C^{(\perp,4)}_{I^{(n)}_{\perp,4}}\left(\frac{s-U_u}{1-U_u}\right) \end{aligned} \tag{5.49}$$

where $(C_n^{(=,i)})_{n\ge 0}$, $i = 1, \dots, 4$ are independent copies of $(C_n^{=})_{n\ge 0}$, which are also independent of the family $(U, I^{(n)}_{=,1}, I^{(n)}_{=,2}, I^{(n)}_{=,3}, I^{(n)}_{=,4})$ in (5.48), and $(C_n^{(\perp,i)})_{n\ge 0}$, $i = 1, \dots, 4$ are independent copies of $(C_n^{\perp})_{n\ge 0}$, which are also independent of $(U_d, U_u, I^{(n)}_{\perp,1}, I^{(n)}_{\perp,2}, I^{(n)}_{\perp,3}, I^{(n)}_{\perp,4})$ in (5.49). Asymptotically, any limit $Z^{=}(s)$ of $n^{-\beta} C_n^{=}(s)$ should satisfy the following fixed-point equation

$$\begin{aligned} Z^{=}(s) \stackrel{d}{=} &\mathbf{1}_{\{s<U\}} \left[(UV_\ell)^\beta Z^{(=,1)}\left(\frac{s}{U}\right) + (U(1-V_\ell))^\beta Z^{(=,2)}\left(\frac{s}{U}\right) \right] \\ &+ \mathbf{1}_{\{s\ge U\}} \left[((1-U)V_r)^\beta Z^{(=,3)}\left(\frac{s-U}{1-U}\right) + ((1-U)(1-V_r))^\beta Z^{(=,4)}\left(\frac{s-U}{1-U}\right) \right], \end{aligned} \tag{5.50}$$

where $Z^{(=,i)}$, $i = 1, \dots, 4$ are independent copies of $Z^{=}$, independent of (U, V_ℓ, V_r). Note that, even though (5.50) resembles very much the corresponding fixed-point equation for quadtrees, it is different from (1.20). Likewise any limit of $n^{-\beta} C_n^{\perp}(s)$ should satisfy

$$\begin{aligned} Z^{\perp}(s) \stackrel{d}{=} &\mathbf{1}_{\{s<U_d\}} (U_d V)^\beta Z^{(\perp,1)}\left(\frac{s}{U_d}\right) + \mathbf{1}_{\{s<U_u\}} (U_u(1-V))^\beta Z^{(\perp,2)}\left(\frac{s}{U_u}\right) \\ &+ \mathbf{1}_{\{s\ge U_d\}} ((1-U_d)V)^\beta Z^{(\perp,3)}\left(\frac{s-U_d}{1-U_d}\right) \\ &+ \mathbf{1}_{\{s\ge U_u\}} ((1-U_u)(1-V))^\beta Z^{(\perp,4)}\left(\frac{s-U_u}{1-U_u}\right), \end{aligned} \tag{5.51}$$

where $Z^{(\perp,i)}$, $i = 1, \dots, 4$ are independent copies of $Z^{\perp}$, independent of (U_d, U_u, V). Moreover, according to (5.44) and (5.45), we expect a connection between these two limits. This will be stated in the first result of the next section and always allows us to focus on $C_n^{=}(s)$ first. Results for $C_n^{\perp}(s)$ can then be deduced easily afterwards.

5.6.2. The conditions to use the contraction argument

Existence of continuous limit processes. The two main difficulties in proving the functional limit theorem for partial match queries in quadtrees where the existence of a continuous limit process and uniform convergence of the mean after rescaling. We address these issues in the present subsection. The first results is the analogue of Proposition 5.6 for 2-d trees where we also include the precise values for the limit variance.

Proposition 5.25. *There exist two random continuous processes $Z^{=}, Z^{\perp}$ with $\mathbb{E}\left[\|Z^{=}\|^2\right] < \infty$, $\mathbb{E}\left[\|Z^{\perp}\|^2\right] < \infty$ and $\mathbb{E}\left[Z^{=}(\xi)\right] = \mathbb{E}\left[Z^{\perp}(\xi)\right] = \mathrm{B}(\beta/2+1, \beta/2+1)$ such that $Z^{=}$ satisfies*

(5.50) and $Z^\perp$ satisfies (5.51). The laws of $Z^=$ and $Z^\perp$ are both unique under these constraints. Moreover $\mathbb{E}\left[Z^=(s)\right] = \mathbb{E}\left[Z^\perp(s)\right] = h(s)$ and the suprema of both processes have finite absolute moments of all orders. Additionally,

- *we have*
$$\frac{2}{\beta+1} Z^\perp(s) \stackrel{d}{=} V^\beta Z^{(=,1)}(s) + (1-V)^\beta Z^{(=,2)}(s) \tag{5.52}$$
and
$$\frac{\beta+1}{2} Z^=(s) \stackrel{d}{=} \mathbf{1}_{\{s<U\}} U^\beta Z^{(\perp,1)}\left(\frac{s}{U}\right) + \mathbf{1}_{\{s\geq U\}} (1-U)^\beta Z^{(\perp,2)}\left(\frac{s-U}{1-U}\right), \tag{5.53}$$

- *for every fixed $s \in [0,1]$, $Z^=(s)$ is distributed like $Z(s)$ where Z is the process constructed in Section 5.3. In particular, $\mathbf{Var}\left[Z^=(s)\right]$ is given in (5.39) and $\mathbf{Var}\left[Z^\perp(s)\right] = K_2^\perp h^2(s)$, where*
$$K_2^\perp = \left(\frac{2c_2}{2\beta+1}\left(\frac{\beta+1}{2}\right)^2 + 2\mathrm{B}\left(\beta+1,\beta+1\right)\left(\frac{\beta+1}{2}\right)^2 - 1\right), \tag{5.54}$$
and c_2 is defined in (5.41),

- *if ξ is uniform on $[0,1]$ and independent of $Z^=, Z^\perp$, then $\mathbf{Var}\left[Z^=(\xi)\right]$ is given by (5.40) and*
$$\begin{aligned}\mathbf{Var}\left[Z^\perp(\xi)\right] = K_3^\perp = &\left(\frac{2c_2}{2\beta+1} + 2\mathrm{B}\left(\beta+1,\beta+1\right)\right)\left(\frac{\beta+1}{2}\right)^2 \mathrm{B}(\beta+1,\beta+1) \\ &- \left(\mathrm{B}\left(\frac{\beta}{2}+1,\frac{\beta}{2}+1\right)\right)^2. \end{aligned}\tag{5.55}$$

Proof. The fixed-point equation (5.50) is very similar to that in (5.5), and we use the approach that has proved fruitful in Section 5.3. Define the continuous operator $G^= : [0,1]^3 \times \mathcal{C}_0[0,1]^4 \to \mathcal{C}_0[0,1]$ by

$$\begin{aligned} G^=(x,y,z,f_1,f_2,f_3,f_4)(s) \quad = \quad & \mathbf{1}_{\{s<x\}}\left[(xy)^\beta f_1\left(\frac{s}{x}\right) + (x(1-y))^\beta f_2\left(\frac{s}{x}\right)\right] \\ &+\mathbf{1}_{\{s\geq x\}}\left[((1-x)z)^\beta f_3\left(\frac{s-x}{1-x}\right)\right] \\ &+\mathbf{1}_{\{s\geq x\}}\left[((1-x)(1-z))^\beta f_4\left(\frac{s-x}{1-x}\right)\right]. \end{aligned}$$

Then let (as in Section 5.3)

$$Z_{n+1}^{=,u} = G^=(U_u, V_u, W_u, Z_n^{=,u1}, Z_n^{=,u2}, Z_n^{=,u3}, Z_n^{=,u4}), \qquad Z_0^{=,u} = h(s),$$

for all $u \in \mathcal{T}$, where $\{U_v, v \in \mathcal{T}\}, \{V_v, v \in \mathcal{T}\}$ and $\{W_v, v \in \mathcal{T}\}$ are three independent families of i.i.d. $[0,1]$-uniform random variables. Lemma 5.11 remains true for $Z_n^= := Z_n^{=,\emptyset}$ since $W_n^=$ equals W_n in distribution where W_n appears in (5.21). Since also $L_n^=$ and L_n (appearing in Lemma 5.13) coincide in distribution, (5.16) holds true for $Z_n^=$ and therefore Proposition 5.9 remains valid. The existence of all moments of $\|Z^=\|$ follows in the same way. Finally, note that $Z_n^=(s)$ is distributed as $Z_n(s)$ for all fixed n, s, hence the one-dimensional distributions of $Z^=$

and Z coincide. It is now easy to see that $Z^\perp$ defined by (5.52) solves (5.51). The uniqueness of $Z^=(s)$ (resp. $Z^\perp(s)$) follows by contraction with respect to the ζ_2 metric when fixing the mean to be $h(s)$. The improvement of the uniqueness statement is obtained as in the quadtree case based on the arguments in [CJ11, Section 5]. Finally, the variance of $Z^\perp(s)$ can be computed as in Section 5.5 but it is much easier to use (5.52), we omit the calculations. □

Uniform convergence of the mean. Comparing construction and recurrence for partial match queries in 2-d trees and quadtrees it seems very likely that this quantities are not only of the same asymptotic order in the case of a uniform query but also closely related for fixed $s \in [0,1]$ and $n \in \mathbb{N}$. This can be formalized by the following Lemma.

Lemma 5.26. *For any $s \in [0,1]$ and $n \in \mathbb{N}$ we have*

$$\frac{1}{5}\mathbb{E}\left[C_n(s)\right] \le \mathbb{E}\left[C_n^=(s)\right] \le 2\mathbb{E}\left[C_n(s)\right].$$

Proof. We prove both bounds by induction on n using the recursive decompositions (1.18), (5.48). Both inequalities are obviously true for $n = 0, 1$. Assume that the assertions were true for all $m \le n-1$ and $s \in [0,1]$. We start with the upper bound which is easier. By (5.48), we have

$$\begin{aligned}\mathbb{E}\left[C_n^=(s)\right] \le 2 &+ \mathbb{E}\left[\mathbf{1}_{\{s<U\}}\left[C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right) + C^{(=,2)}_{I^{(n)}_{=,2}}\left(\frac{s}{U}\right)\right]\right] \\ &+ \mathbb{E}\left[\mathbf{1}_{\{s\ge U\}}\left[C^{(=,3)}_{I^{(n)}_{=,3}}\left(\frac{s-U}{1-U}\right) + C^{(=,4)}_{I^{(n)}_{=,4}}\left(\frac{s-U}{1-U}\right)\right]\right].\end{aligned}$$

Hence, it suffices to show that

$$\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right] \le 2\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(1)}_{I^{(n)}_{1}}\left(\frac{s}{U}\right)\right].$$

This can be done in two steps. First, by conditioning on $I^{(n)}_{=,1}$ and U, using the induction hypothesis, we have

$$\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right] \le 2\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right].$$

Finally, conditioning on U, $I^{(n)}_{=,1}$ is stochastically smaller than $I^{(n)}_1$ which gives

$$\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right] \le 2\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(1)}_{I^{(n)}_{1}}\left(\frac{s}{U}\right)\right].$$

by monotonicity of $n \to \mathbb{E}\left[C_n(s)\right]$. For the lower bound, note that

$$\begin{aligned}\mathbb{E}\left[C_n^=(s)\right] \ge 1 &+ \mathbb{E}\left[\mathbf{1}_{\{s<U\}}\left[C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right) + C^{(=,2)}_{I^{(n)}_{=,2}}\left(\frac{s}{U}\right)\right]\right] \\ &+ \mathbb{E}\left[\mathbf{1}_{\{s\ge U\}}\left[C^{(=,3)}_{I^{(n)}_{=,3}}\left(\frac{s-U}{1-U}\right) + C^{(=,4)}_{I^{(n)}_{=,4}}\left(\frac{s-U}{1-U}\right)\right]\right].\end{aligned}$$

Therefore, it is enough to prove

$$\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right] \ge \frac{1}{5}\left(\mathbb{E}\left[\mathbf{1}_{\{s<U\}}C^{(1)}_{I^{(n)}_{1}}\left(\frac{s}{U}\right)\right] - 1\right).$$

This can be done as for the upper bound. First, by the induction hypothesis, we have

$$\mathbb{E}\left[\mathbf{1}_{\{s<U\}} C^{(=,1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right] \geq \frac{1}{5}\mathbb{E}\left[\mathbf{1}_{\{s<U\}} C^{(1)}_{I^{(n)}_{=,1}}\left(\frac{s}{U}\right)\right].$$

The result follows as for the upper bound by the fact that $I^{(n)}_{=,1}$ is stochastically larger than $(I^{(n)}_1 - 1)^+$ and $C^{(1)}_{(I^{(n)}_1-1)^+} \geq C^{(1)}_{I^{(n)}_1} - 1$. □

Recalling (5.46) and (5.47), it is natural to introduce the constants

$$K^{=}_1 = \frac{\kappa^{=}}{\mathrm{B}\left(\frac{\beta}{2}+1, \frac{\beta}{2}+1\right)}, \qquad K^{\perp}_1 = \frac{\kappa^{\perp}}{\mathrm{B}\left(\frac{\beta}{2}+1, \frac{\beta}{2}+1\right)} \qquad \text{with} \qquad K^{\perp}_1 = \frac{2}{\beta+1}K^{=}_1, \tag{5.56}$$

and the functions $\bar{\mu}^{\perp}(s) = K^{\perp}_1 h(s)$, and $\bar{\mu}^{=}(s) = K^{=}_1 h(s)$.

Proposition 5.27. *There exists $\varepsilon_{=} > 0$ such that*

$$\sup_{s\in[0,1]} |n^{-\beta}\mathbb{E}\left[C^{=}_n(s)\right] - \bar{\mu}^{=}(s)| = O(n^{-\varepsilon_{=}}),$$

and the analogous result holds true for $\mathbb{E}\left[C^{\perp}_n(s)\right]$.

We proceed as in Section 5.4 by considering the continuous-time process $P^{=}_t(s)$. Again the proof runs along very similar lines as in the quadtree case. Thus we only give a brief sketch that focuses on the few locations where the arguments need to be modified.

Sketch of proof. The first step is to prove pointwise convergence which is done as in [CJ11]. By Lemma 5.26, using a Poisson(t) number of points, we have

$$\frac{1}{5}\mathbb{E}\left[P_t(s)\right] \leq \mathbb{E}\left[P^{=}_t(s)\right] \leq 2\mathbb{E}\left[P_t(s)\right]. \tag{5.57}$$

Let $\tau^{=}_1$ be the arrival time of the first point which gives rise to a horizontal partitioning line that intersects the query line $\{x = s\}$, and let $Q^{=}_1 = Q^{=}_1(s)$ be the lower of the two rectangles created by this cut. Let $\xi^{=}_1 := \xi^{=}_1(s)$ be the relative position of the query line s within the rectangle $Q^{=}_1$ and $M^{=}_1 = \mathrm{Leb}(Q^{=}_1)$. Denoting τ the arrival time of the first point in the process, we have

$$\mathbb{E}\left[P^{=}_t(s)\right] = \mathbf{P}\left(t \geq \tau\right) + \mathbf{P}\left(t \geq \tau^{=}_1\right) + 2\mathbb{E}\left[\tilde{P}^{=}_{M^{=}_1(t-\tau^{=}_1)}(\xi^{=}_1)\right],$$

where $(\tilde{P}^{=}(t))_{t\geq 0}$ denotes an independent copy of $(P^{=}(t))_{t\geq 0}$ and $\tilde{P}^{=}(t) = 0$ for $t < 0$. Similarly, let $\tau^{=}_k$ be the arrival time of the first point which cuts $Q^{=}_{k-1}$ perpendicularly to the query line. Let $Q^{=}_k$ be the lower of the two rectangles created by this cut, and let $\xi^{=}_k$ be the position of the query line s relative to the rectangle $Q^{=}_k$. With this notation and $M^{=}_k = \mathrm{Leb}(Q^{=}_k)$, we have

$$\mathbb{E}\left[P^{=}_t(s)\right] = g^{=}_k(t) + 2^k\mathbb{E}\left[\tilde{P}^{=}_{M^{=}_k(t-\tau^{=}_k)}(\xi^{=}_k)\right],$$

where $0 \leq g^{=}_k(t) \leq 2^{k+1}$.

We need to modify the inter-arrival times $\zeta'^{=}_k = \tau^{=}_k - \tau^{=}_{k-1}$. We can split $\zeta'^{=}_k$ in the time it takes for the first vertical point to fall in $Q^{=}_{k-1}$ which we denote by $\zeta'^{=,1}_k$ and the remaining time $\zeta'^{=,2}_k$. The normalized versions of the inter-arrival times with unit mean are

$$\begin{aligned}
\zeta^{=,1}_k &= \zeta'^{=,1}_k \cdot M^{=}_{k-1}, \\
\zeta^{=,2}_k &= \left(\frac{\xi^{=}_k}{\xi^{=}_{k-1}}\mathbf{1}_{\{\xi^{=}_k<\xi^{=}_{k-1}\}} + \frac{\xi^{=}_{k-1}}{\xi^{=}_k}\mathbf{1}_{\{\xi^{=}_k\geq\xi^{=}_{k-1}\}}\right)\zeta'^{=,2}_k \cdot M^{=}_{k-1} \geq \zeta'^{=,2}_k \cdot M^{=}_{k-1}.
\end{aligned}$$

Write $\mathcal{M}_k^= = M_k^=/M_{k-1}^=$. Observe that, given $\mathcal{M}_0^=, \ldots, \mathcal{M}_k^=$, the random variable $F_k^= = M_k^= \cdot \tau_k^=$ is not independent of $(\xi_\ell^=)_{0\le \ell\le k}$, a property which is used in [CJ11] and in the proof of Lemma 5.18 in this thesis. However we can use the trivial lower bound $0 \le F_k$ and the upper bound obtained by bounding $\zeta'^{=,2}_k$ from above by $\zeta_k^{=,2}/M_{k-1}^=$. Then, using almost sure monotonicity of $P_t(s)$ (in t) and (5.57) to transform bounds for the mean in the quadtree to bounds in the 2-d tree (and vice versa), it is easy to see that the techniques of Section 4 in [CJ11] work equally well in this case. The limit $\bar\mu^=(s)$ is identified as in Section 5 of [CJ11] since both limits $\bar\mu$ and $\bar\mu^=$ satisfy the same fixed-point equation.

The generalization to uniform convergence with polynomial rate can be worked out as in Section 5.4 using the modifications we have described above, e.g. the behaviour along the edge is controlled by Lemma 5.17 and 5.26. The constants appearing in the course of Section 5.4 need to be modified, but $\varepsilon_=$ may be chosen to equal the value of ε in Proposition 5.16. The depoissonization of Subsection 5.4.3 goes through without any modification.

Finally, we indicate how to proceed with $\mathbb{E}\left[C_n^\perp(s)\right]$. The arguments above can be used to treat uniform convergence of $n^{-\beta}\mathbb{E}\left[C_n^\perp(s)\right]$ on $[0,1]$; we present a direct approach relying on (5.44). We have

$$\begin{aligned}
n^{-\beta}\mathbb{E}\left[C_n^\perp(s)\right] &= n^{-\beta} + 2n^{-\beta}\mathbb{E}\left[C_{S_n}^=(s)\right] \\
&= n^{-\beta} + 2\int_0^1 \sum_{k=0}^{n-1} \left(\bar\mu^=(s) + O(k^{-\varepsilon_=})\right) \frac{k^\beta}{n^\beta} \mathbf{P}\left(\mathrm{Bin}(n-1,v)=k\right) dv \\
&= n^{-\beta} + 2\bar\mu^=(s) \cdot \frac{\mathbb{E}\left[\mathrm{Bin}(n-1,V)^\beta\right]}{n^\beta} + O(n^{-\beta}\mathbb{E}\left[\mathrm{Bin}(n-1,V)^{\beta-\varepsilon_=}\right]) \\
&= \bar\mu^\perp(s) + O(n^{-\varepsilon_=}),
\end{aligned}$$

uniformly in $s \in [0,1]$ using Minkowski's inequality, the concentration result for the binomial distribution in (5.10), (5.56) for the first term and Jensen's inequality for the second. $\square$

5.6.3. Statement of the result

We are finally ready to state the version of our main result for 2-d trees. It is proved along the same lines we used for the case of quadtrees, and we omit the details.

Theorem 5.28. *Let $Z^=$ and $Z^\perp$ be the processes in Proposition 5.25. Then*

$$\left(\frac{C_n^=(s)}{K_1^= n^\beta}\right)_{s\in[0,1]} \to (Z^=(s))_{s\in[0,1]}, \qquad \left(\frac{C_n^\perp(s)}{K_1^\perp n^\beta}\right)_{s\in[0,1]} \to \left(Z^\perp(s)\right)_{s\in[0,1]},$$

in distribution in $(\mathcal{D}[0,1], d_{sk})$. Here $K_1^=$ and $K_1^\perp$ are defined in (5.56)*. Moreover*

$$n^{-\beta}\mathbb{E}\left[C_n^=(s)\right] \to K_1^=[s(1-s)]^{\beta/2}, \quad n^{-2\beta}\mathbf{Var}\left[C_n^=(s)\right] \to (K_1^=)^2 K_2 [s(1-s)]^\beta,$$

and

$$n^{-\beta}\mathbb{E}\left[C_n^\perp(s)\right] \to K_1^\perp[s(1-s)]^{\beta/2}, \quad n^{-2\beta}\mathbf{Var}\left[C_n^\perp(s)\right] \to (K_1^\perp)^2 K_2^\perp [s(1-s)]^\beta,$$

where K_2 is in (5.39) *and $K_2^\perp$ in* (5.54)*. If ξ is uniformly distributed on $[0,1]$, independent of $(C_n^=)_{n\ge0}, (C_n^\perp)_{n\ge0}$ and $Z^=, Z^\perp$, then*

$$\frac{C_n^=(\xi)}{K_1^= n^\beta} \xrightarrow{d} Z^=(\xi), \qquad \frac{C_n^\perp(\xi)}{K_1^\perp n^\beta} \xrightarrow{d} Z^\perp(\xi),$$

with convergence of the first two moments in both cases. In particular

$$\mathbf{Var}\left[C_n^{=}(\xi)\right] \sim K_4^{=} n^{2\beta}, \qquad \mathbf{Var}\left[C_n^{\perp}(\xi)\right] \sim K_4^{\perp} n^{2\beta},$$

where

$$K_4^{=} = (K_1^{=})^2 K_3 \approx 0.69848, \qquad K_4^{\perp} = (K_1^{\perp})^2 K_3^V \approx 0.77754,$$

with K_3 *in* (5.40) *and* $K_3^{\perp}$ *in* (5.55).

For simulations of both versions of 2-d trees with corresponding $C_n^{=}$ and $C_n^{\perp}$ see figure 5.5 and 5.6 on pages 103 resp. 104.

Note that since $Z^{=}(s)$ equals $Z(s)$ in distribution for fixed $s \in [0,1]$, thus we can characterize $Z^{=}(s)$ as in (5.37). (5.52) together with Proposition 5.25 implies that for fixed $s \in [0,1]$

$$Z^{\perp}(s) \stackrel{d}{=} Z^{\perp} \cdot (s(1-s))^{\beta/2},$$

with

$$Z^{\perp} = \frac{\beta+1}{2}\left(V^{\beta} Z + (1-V)^{\beta} Z'\right),$$

where Z' is an independent copy of Z, Z being defined in Proposition 5.23, and V is uniform on $[0,1]$ and independent of (Z, Z'). In particular, we have

$$\mathbb{E}\left[(Z^{\perp})^m\right] = \left(\frac{\beta+1}{2}\right)^m \sum_{\ell=0}^{m} \binom{m}{\ell} \mathrm{B}(\beta\ell+1, \beta(m-\ell)+1) c_{\ell} c_{m-\ell},$$

for $m \geq 2$ where $c_m = \mathbb{E}\left[Z^m\right]$ satisfies recursion (5.38) and $c_0 = c_1 = 1$.
As in the quadtree case, it is possible to give convergence of mixed moments of arbitrary order, compare Theorem 5.4. Distributional and moment convergence of the suprema of the processes after rescaling follows similarly.

5.7. Open problems

We have given a functional limit law for the cost of the partial match retrieval problem in random quadtrees and 2-d trees. Moreover, we obtained a description of the limiting distribution for fixed s in terms of a single distribution whose moments can be computed recursively. Solving several open problems, our results naturally give rise to further studies:

Covariances. Our results imply $n^{-2\beta}\mathrm{Cov}[C_n(t), C_n(s)] \to \mathrm{Cov}[Z(t), Z(s)]$. Hence, the covariance function of the process Z is of interest. We do not provide any information on this quantity in the thesis.

Path properties. We know the paths of Z to be continuous. By construction, the paths of the sequence Z_n in Section 5.3, whose uniform limit is Z, are locally β'-Hölder continuous for any $\beta' < \beta/2$ and the same holds for the mean of Z. Hence, also in the context of a related result obtained in [CLG11] that will be discussed in the next section, we conjecture that the sample paths of Z are almost surely locally β'-Hölder continuous for any $\beta' < \beta/2$.

The supremum. Consider the limiting random variable S of the supremum in Theorem 5.3. Trivially, S is non-negative and our techniques imply S to have moments of all orders. However, we do not provide further information about S except for the distributional inequality (5.6) and the trivial lower bound $\mathbb{E}[S^m] \geq c_m 2^{-\beta m}$ where c_m is given in (5.38). In particular, $\mathbb{E}[S]$ is unknown to us and we believe that this problem is hard. To this end, a comparison to the analogous question for the FIND process, introduced by Grübel and Rösler [GR96], seems appropriate [see [Dev84] and [Dev01] for further results on the worst-case of FIND].

Finally, one is immediately led to ask for similar results in higher dimensions. The results on mean convergence for a uniform query line in [FGPR93], [CH03] in the case of quadtrees and [FP86] and [CH06] for K-d trees for higher dimensions appear to be sufficiently strong for this purpose. We believe that our method can be applied essentially in the same manner to the case where, in trees of dimension d, $d-1$ components are left open and only one component is fixed. Fixing more, say s, parameters leads to considering functions from the unit cube $[0,1]^s$ to $\mathbb{R}$; thus, a generalization of the contraction method to more intricate function spaces is necessary.

5.8. Random recursive triangulations

Curien and Le Gall [CLG11] consider a stochastic process in which chords (straight connections) are inserted between points on the unit circle with circumference 2π. In each step, two points on the circle are chosen uniformly at random and become connected by a chord if it does not intersect any other existing one. In the case of a crossing with a present chord we reject the points and do not insert anything. Let N_n be the number of inserted chords at time n, i.e. after n drawings of uniform point pairs where $N_0 = 0$. By L_n we denote the union of all inserted chords by time n considered as a subset of the unit disk in the complex plane. The authors introduce $L_\infty = \overline{\bigcup_{n\geq 1} L_n}$ as an infinite geodesic lamination and investigate its Hausdorff dimension, various other geometric properties and approximations by discrete triangulations of polygons. More interestingly in our context, they also consider the random variable $H_n(x,y)$ which counts the number of intersections of chords in L_n with the chord from x to y for fixed x, y on the sphere. Without loss of generality one may fix one point to be one and consider $H_n(1,s)$ as a process where s ranges over all points on the unit sphere. This immediately connects the problem with the partial match retrieval algorithm in quadtrees. For the sake of convenience, we state the main result from [CLG11] on $H_n(1,x)$, Theorem 1, together with parts of Proposition 4.1. Subsequently, for the sake of comparing with the partial match problem, it is appropriate to identify the unit sphere with the unit interval by $x \to e^{2i\pi x}, x \in [0,1]$. We then let $H_n(x) := H_n(0,x)$ where $H_n(x)$ is extended to a càdlàg function in those finitely many points where the quantity is not well-defined.

Theorem 5.29 (Theorem 1[CLG11]).
- *Almost surely,*

$$n^{-1/2} N_n \to \sqrt{\pi} \tag{5.58}$$

  as $n \to \infty$.

- *There exists a random process $(M(s), s \in [0,1])$ which is (locally) β'-Hölder continuous for every $\beta' < \beta$ such that for every $s \in [0,1]$,*

$$n^{-\beta/2} H_n(s) \to M(s), \tag{5.59}$$

in probability. Moreover, $\mathbb{E}\left[M(s)\right] = C_L(s(1-s))^\beta$ *for some* $C_L > 0$.

Proving only pointwise convergence, process convergence is conjectured in Section 4 of [CLG11]. The proof of their main result is based on fragmentation processes [for the convergence] and suitable upper bounds on the moments [for the Hölder continuity via the Kolmogorov-Chentsov theorem]. It is worthwhile noting that, since N_n is order $\sqrt{n}$, $H_n(x)$ is of order N_n^β which resembles the behaviour in quadtrees [there, n points give rise to n horizontal and n vertical lines, rejection does not take place; thus N_n is to be identified with n]. Before discussing a strengthening of (5.59) we mention a simple observation concerning the sequence (N_n) that allows a refinement of (5.58) based on the work of Bai et al. [BHLT01].

Let U be the length of the arc connecting the first inserted point pair in an arbitrary direction. Then $(N_n)_{n\geq 0}$ satisfies the following recursion

$$N_n \stackrel{d}{=} N^{(1)}_{I_1^{(n)}} + N^{(2)}_{I_2^{(n)}} + 1, \qquad n \geq 1,$$

where the random sequences $(N_n^{(1)})_{n\geq 0}, (N_n^{(2)})_{n\geq 0}$ are independent copies of $(N_n)_{n\geq 0}$, and independent of $(I_1^{(n)}, I_2^{(n)}, I_3^{(n)})$. Additionally, given U,

$$\left(I_1^{(n)}, I_2^{(n)}, I_3^{(n)}\right) \stackrel{d}{=} \text{Multi}(n-1; U^2, (1-U)^2, 2U(1-U)).$$

Note that N_n satisfies the same recursive decomposition as the number of maxima in the *unit* triangle [that is the triangle with corners $(0,0)$, $(0,1)$ and $(1,0)$]; thus both distributions are equal for all n. In [BHLT01], the authors give exact formulas for the mean and the second moment together with first order asymptotics of all higher moments which imply asymptotic normality of N_n after rescaling. We quote their Theorem 3 here for the sake of completeness.

Theorem 5.30. *Mean and Variance of* N_n *satisfy*

$$\begin{aligned} \mathbb{E}\left[N_n\right] &= \frac{\sqrt{\pi} n!}{\Gamma(n+1/2)} - 1 = \sqrt{\pi n} - 1 + O(n^{-1/2}), \\ \mathbf{Var}\left(N_n\right) &= \sigma^2 \sqrt{n} - \frac{\pi}{4} + O(n^{-1/2}), \end{aligned}$$

with $\sigma^2 = \sqrt{(2\log 2 - 1)\sqrt{\pi}}$. N_n *satisfies asymptotic normality, i.e.*

$$\frac{N_n - \sqrt{\pi n}}{\sigma n^{1/4}} \xrightarrow{d} N(0,1),$$

where the limit holds with convergence of all moments.

Remark 5.31. Using the asymptotics for mean and variance it is possible to give a considerably shorter proof of the central limit theorem by the contraction method based on the ζ_3 metric; the details have been worked out in [NR04b].

It seems obvious from the recursive construction of the insertion process that $H_n(s)$ satisfies an additive recurrence. To this end, let $0 \leq U_1 \leq U_2 \leq 1$ be the values of the feet of the first inserted chord, where we use the notation in [CLG11] and denote the points on the unit disk which are connected by a chord by its feet. Let S^+ be the arc connecting U_1, U_2 clockwise and S^- be the arc connecting them counterclockwise. Observe that (U_1, U_2) has density $2\mathbf{1}_{\{0\leq u_1\leq u_2\leq 1\}}$ and

$1-(U_2-U_1)$, the length of S^+ has size-biased uniform distribution which is the distribution of $U^{1/2}$ for a uniform U. Furthermore, let I_n^+ be the number of attempted insertions of chords in S^+ and I_n^- the corresponding quantity in S^-. By F_n we denote the number of unsuccessful insertions of chords due to one foot falling in S^+ and the other one in S^-. Then, given (U_1,U_2),

$$(I_n^+, I_n^-, F_n) \stackrel{d}{=} \text{Multi}(n-1; U, (1-\sqrt{U})^2, 2\sqrt{U}(1-\sqrt{U})),$$

where $U := (1-(U_2-U_1))^2$ is uniformly distributed on $[0,1]$. The construction provides the following recursive decomposition

$$\begin{aligned} C_n(s) &\stackrel{d}{=} \mathbf{1}_{\{s\le U_1\}} C^{(1)}_{I_n^+}\left(\frac{s}{1-(U_2-U_1)}\right) + \mathbf{1}_{\{s>U_2\}} C^{(1)}_{I_n^+}\left(\frac{s-(U_2-U_1)}{1-(U_2-U_1)}\right) \\ &\quad + \mathbf{1}_{\{U_1<s\le U_2\}}\left(1 + C^{(1)}_{I_n^+}\left(\frac{U_1}{1-(U_2-U_1)}\right) + C^{(2)}_{I_n^-}\left(\frac{s-U_1}{U_2-U_1}\right)\right), \end{aligned} \tag{5.60}$$

for $s\in[0,1]$. For fixed n, we can consider $(C_n(s))_{s\in[0,1]}$ as a process with càdlàg paths and (5.60) remains true on the level of càdlàg functions. Letting $Q_n(s) := n^{-\beta/2}C_n(s)$ and introducing linear operators $A_1^{(n)}$ and $A_2^{(n)}$

$$\begin{aligned} A_1^{(n)} f(s) &= \left(\frac{I_n^+}{n}\right)^{\beta/2}\left(\mathbf{1}_{\{s\le U_1\}} f\left(\frac{s}{(1-(U_2-U_1)}\right) + \mathbf{1}_{\{s>U_2\}} f\left(\frac{s-(U_2-U_1)}{(1-(U_2-U_1)}\right)\right. \\ &\quad \left. + \mathbf{1}_{\{U_1<s\le U_2\}} f\left(\frac{U_1}{1-(U_2-U_1)}\right)\right), \\ A_2^{(n)} f(s) &= \mathbf{1}_{\{U_1<s\le U_2\}}\left(\frac{I_n^-}{n}\right)^{\beta/2} f\left(\frac{s-U_1}{U_2-U_1}\right), \end{aligned}$$

yields

$$Q_n(s) \stackrel{d}{=} A_1^{(n)} Q^{(1)}_{I_n^+}(s) + A_2^{(n)} Q^{(2)}_{I_n^-}(s) + b^{(n)}(s)$$

with $b^{(n)}(s) = \mathbf{1}_{\{U_1<s\le U_2\}} n^{-\beta/2}$. This suggests that any limit $Q(s) = \lim_n Q_n(s)$ satisfies

$$Q(s) \stackrel{d}{=} A_1 Q^{(1)}(s) + A_2 Q^{(2)}(s) \tag{5.61}$$

as process in $\mathcal{D}[0,1]$ with

$$\begin{aligned} A_1 f(s) &= (1-(U_2-U_1))^\beta\left(\mathbf{1}_{\{s\le U_1\}} f\left(\frac{s}{1-(U_2-U_1)}\right) + \mathbf{1}_{\{s>U_2\}} f\left(\frac{s-(U_2-U_1)}{1-(U_2-U_1)}\right)\right. \\ &\quad \left. + \mathbf{1}_{\{U_1<s\le U_2\}} f\left(\frac{U_1}{1-(U_2-U_1)}\right)\right), \\ A_1 f(s) &= \mathbf{1}_{\{U_1<s\le U_2\}} (U_2-U_1)^\beta f\left(\frac{s-U_1}{U_2-U_1}\right), \end{aligned}$$

In Section 8, Curien and Le Gall observe that the limit process M satisfies fixed-point equation (5.61); moreover they ask to what extent the distribution of M is characterized by (5.61). We lack knowledge about the supremum of M. As in the partial match case, it is very likely that $\|M\|$ has (at least) finite second moment, which, by

$$\mathbb{E}\left[\|A_1\|^2 + \|A_2\|^2\right] = \mathbb{E}\left[1-(U_2-U_1)^{2\beta}\right] + \mathbb{E}\left[(U_2-U_1)^{2\beta}\right] < 1,$$

would imply M to be the only solution (modulo multiplicative constants) of (5.61) with continuous (or only càdlàg) sample paths, square integrable norm and mean $(s(1-s))^\beta$ at point s.

As in the quadtree case, applying the contraction method, more precisely Theorem 3.6, to achieve distributional convergence of $(n^{-\beta/2}H_n(s))_{s\in[0,1]}$ boils down to establishing the following two results: First, it is necessary to prove $\mathbb{E}\left[\|M\|^2\right] < \infty$; second, we need a uniform polynomial rate of convergence for the mean of $H_n(s)$ after rescaling. The similarities between the recursions in the present and in the quadtree case propose that both problems could be solved by means of our approach in Sections 5.3 and 5.4, where it seems that technicalities are more involved here. Remembering that the rate of convergence for the mean in quadtrees has been transferred from the analogous result at a uniform query line (1.14), the following theorem could play a major role. Its proof is based on the use of generating functions as in [FGPR93] and [CH03]; it is not given here.

Theorem 5.32. *Let ξ be independent of the process and uniformly distributed on $[0,1]$. Let $\mu_n = \mathbb{E}\left[H_n(\xi)\right], \alpha = \beta/2 + 1$ and $\bar{\alpha} = \frac{-\sqrt{17}+1}{4}$. Then*

$$\begin{aligned} \mu(n) &= \frac{\sqrt{\pi}}{4}\sum_{k=1}^{n}\binom{n}{k}(-1)^{k+1}\frac{\Gamma(k-\alpha+1)\Gamma(k-\bar{\alpha}+1)}{k!\Gamma(k+3/2)\Gamma(2-\alpha)\Gamma(2-\bar{\alpha})} \\ &= C_u n^{\beta/2} + O(1), \end{aligned}$$

with

$$C_u = \frac{\sqrt{\pi}\Gamma(2\alpha - 1/2)}{2\Gamma(\alpha)\Gamma^2(\alpha+1/2)}.$$

Remark 5.33. In [CLG11], the authors show

$$\frac{H_n(\xi)}{n^{\beta/2}} \to T,$$

where $T > 0$ almost surely and convergence holds in mean and almost surely. Thus we have identified the mean of T and moreover, we have also found the value of C_L in Theorem 5.29,

$$C_L = \frac{C_u}{\mathrm{B}(\frac{\beta}{2}+1, \frac{\beta}{2}+1)} = \frac{\sqrt{\pi}\Gamma(\beta+1)\Gamma(\beta+3/2)}{2\Gamma^3(\beta/2+1)\Gamma^2(\beta/2+3/2)} \approx 1.292574852.$$

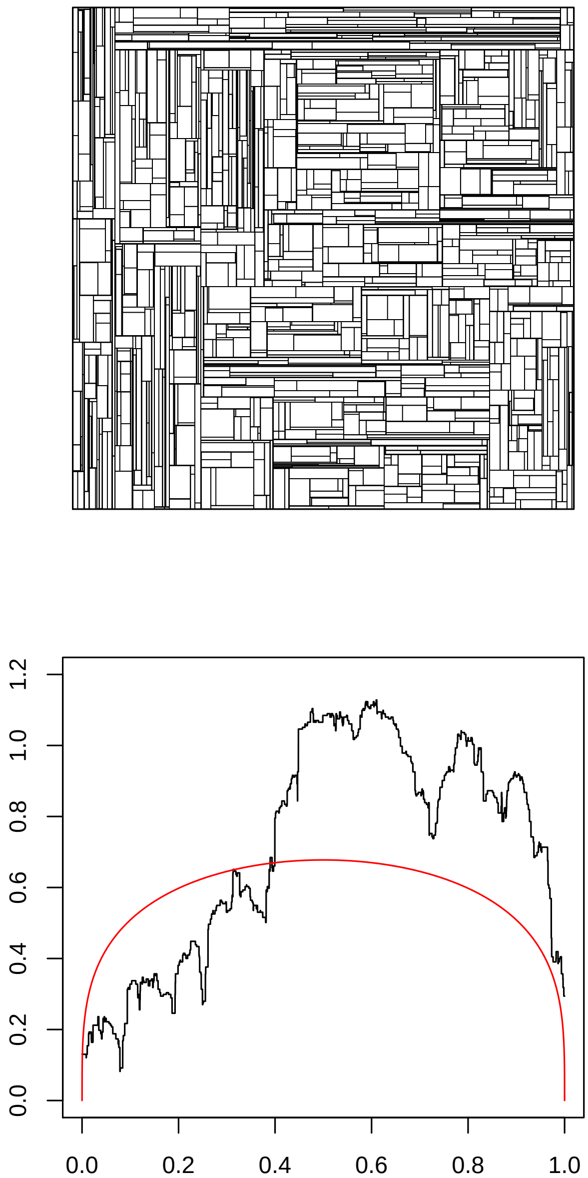
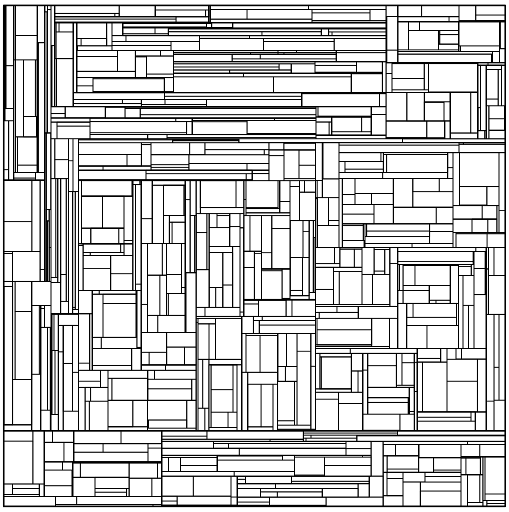


Figure 5.5.: 2-d tree with vertical split at the root and $n = 1000$. The lower figure shows $\big((K_1^{=})^{-1} n^{-\beta} C_n^{=}(s)\big)_{s \in [0,1]}$ and the limit mean.



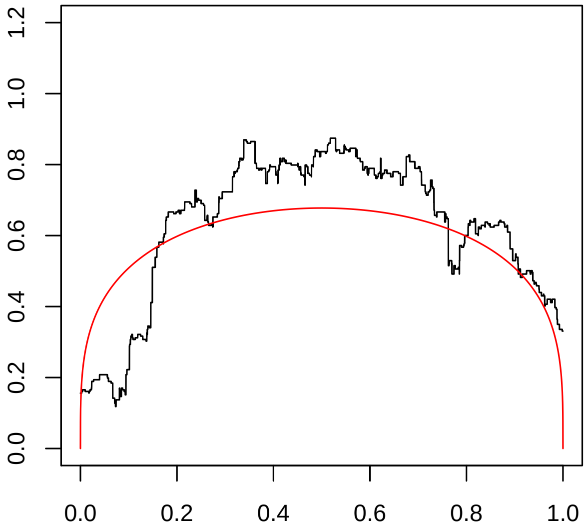


Figure 5.6.: 2-d tree with horizontal split at the root and $n = 1000$.The lower figure shows $\big((K_1^{\perp})^{-1}n^{-\beta}C_n^{\perp}(s)\big)_{s\in[0,1]}$ and the limit mean.

A. Appendix

The first Lemma is very elementary however practical and used several times throughout the thesis. It is a slight generalization of Theorem 1.2 in [Bil68].

Lemma A.1. *Let (S, d) be a metric space. For any $0 < s \leq 1$, closed set C and $\varepsilon > 0$, there exists a real-valued function $g : S \to [0, 1]$ with $g(x) = 1$ for $x \in C$, $g(x) = 0$ for $d(x, C) \geq \varepsilon$ and $|g(x) - g(y)| \leq \varepsilon^{-s}\|x - y\|^s$ for all $x, y \in S$. The function $f : S \to [-1, 1], f = 2g - 1$ satisfies $f(x) = 1$ for $x \in C$ and $f(x) = -1$ for $d(x, C) \geq \varepsilon$ and $|f(x) - f(y)| \leq 2\varepsilon^{-s}\|x - y\|^s$.*

Proof. The function $g(x) = \max\left(0, 1 - \left(\frac{d(x,C)}{\varepsilon}\right)^s\right)$ has the desired properties. □

The following Lemmas A.2 to A.5 all concern properties of the metrics ℓ_s and κ_s. The statement of Lemma A.2 may be found in several references, e. g. [BF81, Lemma 8.1]. Lemma A.3 is given in the same paper, however the proof presented here is based on arguments from [DR85]. The idea of the proof of Lemma A.5 is taken from [Dud76, Section 20].

Lemma A.2. *Let B be separable and μ, ν be probability measures on B with finite absolute moment of order s. Then there exists random variables X, Y with $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$ and $\ell_s(\mu, \nu) = \|X - Y\|_s$.*

Proof. Let T be the set of probability measures on B^2 with marginals μ and ν. Let $\varepsilon > 0$ be arbitrary. Since μ and ν are tight, we can find compact sets K, L with $\mu(K^c), \nu(L^c) < \varepsilon$. Thus $\varrho((K \times L)^c) < 2\varepsilon$ for any $\varrho \in T$ and T is tight. Any accumulation point of T in the weak topology has to have marginals μ and ν which shows that T is closed. Prokhorov's theorem implies compactness of T. The map $f : T \to \mathbb{R}^+ : f(\varrho) = \|X - Y\|_s$ with $\mathcal{L}(X, Y) = \varrho$ is continuous and therefore it attains its infimum on T. □

Lemma A.3. *Let B be separable, $(X_n), X$ be B-valued random variables, $s > 0$ and $\mathbb{E}[\|X_n\|^s]$, $\mathbb{E}[\|X\|^s] < \infty$ for all n. Then $\ell_s(X_n, X) \to 0$ implies $X_n \to X$ in distribution and $\mathbb{E}[\|X_n\|^s] \to \mathbb{E}[\|X\|^s]$. The converse is true as well.*

Proof. Assume $\ell_s(X_n, X) \to 0$. Let $\varepsilon > 0$ and $X_n^{(\varepsilon)}, X^{(\varepsilon)}(n)$ be random variables with $\mathcal{L}(X_n^{(\varepsilon)}) = \mathcal{L}(X_n)$ and $\mathcal{L}(X^{(\varepsilon)}(n)) = \mathcal{L}(X)$ for all n such that

$$\|X_n^{(\varepsilon)} - X^{(\varepsilon)}(n)\|_s \leq \varepsilon$$

for all $n \geq N_0 = N_0(\varepsilon)$. As discussed right before Theorem 2.23, Lemma A.1 together with the proof of the Portementeau Lemma implies that distributional convergence $X_n \to X$ is equivalent to convergence of $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ for all bounded s-Hölder continuous functions $f : B \to \mathbb{R}$ where $s \in (0, 1]$ is fixed. Let $s > 1$ and f be bounded and K-Lipschitz continuous. Then

$$\mathbb{E}\left[|f(X_n^{(\varepsilon)}) - f(X^{(\varepsilon)}(n))|\right] \leq K\mathbb{E}\left[\|X_n^{(\varepsilon)} - X^{(\varepsilon)}(n)\|\right] \leq K\|X_n^{(\varepsilon)} - X^{(\varepsilon)}(n)\|_s \leq K\varepsilon$$

for all $n \geq N_0$. By the triangle inequality, we have $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ hence $X_n \to X$ in distribution. For $s < 1$ the proof runs along the same lines using bounded s-Hölder continuous functions. Next, we have

$$\|X_n\|_s \leq \|X_n^{(\varepsilon)} - X^{(\varepsilon)}(n)\|_s + \|X\|_s \leq \varepsilon + \|X\|_s$$

and analogously, $\|X\|_s \leq \varepsilon + \|X_n\|_s$ hence

$$|\|X_n\|_s - \|X\|_s| < \varepsilon$$

for all $n \geq N_0$. This gives $\mathbb{E}[\|X_n\|^s] \to \mathbb{E}[\|X\|^s]$ as $n \to \infty$.
Now, suppose $X_n \to X$ in distribution and $\mathbb{E}[\|X_n\|^s] \to \mathbb{E}[\|X\|^s]$. By separability, using Skorokhod's representation theorem, we may choose $(X_n), X$ such that $X_n \to X$ almost surely. Let $q_s = \max(1, 2^{s-1})$ and $\Delta_n = \|X_n - X\|$. Then, for any $0 < \varepsilon < 1$,

$$\begin{aligned}
\mathbb{E}[\|X_n - X\|^s] &\leq \varepsilon^s + \mathbb{E}\left[\|X_n - X\|^s 1_{\{\Delta_n > \varepsilon\}}\right] \\
&\leq \varepsilon^s + q_s\left(\mathbb{E}\left[(\|X_n\|^s + \|X\|^s)\, 1_{\{\Delta_n > \varepsilon\}}\right]\right) \\
&= \varepsilon^s + q_s\left(\mathbb{E}[\|X_n\|^s - \|X\|^s] + 2\mathbb{E}\left[\|X\|^s 1_{\{\Delta_n > \varepsilon\}}\right]\right. \\
&\quad \left. - \mathbb{E}\left[(\|X_n\|^s - \|X\|^s) 1_{\{\Delta_n \leq \varepsilon\}}\right]\right). \qquad \text{(A.1)}
\end{aligned}$$

Then, the bounded convergence Theorem implies $\mathbb{E}\left[\|X\|^s 1_{\{\Delta_n > \varepsilon\}}\right] \to 0$ and the term in (A.1) to vanish as $n \to \infty$. This implies the assertion. □

Lemma A.4. *Let B be separable and $s > 0$. Then the topologies induced by ℓ_s, κ_s on $\mathcal{M}_s(B)$ are equal.*

Proof. Let $\ell_s(X_n, X) \to 0$. Again, we may choose $X_n \to X$ almost surely. For $s > 1$ and Hölder's inequality gives

$$\kappa_s(X_n, X) \leq \mathbb{E}\left[\|X(\|X_n\|^{s-1} - \|X\|^{s-1})\|\right] + (\mathbb{E}[\|X_n\|^s])^{1-1/s}(\mathbb{E}[\|X_n - X\|^s])^{1/s}.$$

By Lemma A.3, we have $\mathbb{E}[\|X_n\|^s] \to \mathbb{E}[\|X\|^s]$. Thus, by arguments as in the proof of Lemma A.3, we see that $\kappa_s(X_n, X)$ tends to zero as $n \to \infty$. For $s \leq 1$, it is not hard to see that for there exists a constant $\bar{q} = \bar{q}(s)$ such that for all $x, y \in B$,

$$\|x\|x\|^{s-1} - y\|y\|^{s-1}\| \leq \bar{q}\|x - y\|^s.$$

A detailed proof of this inequality in contained in the proof of Lemma 2.18. This proves $\kappa_s(X_n, X) \to$ in the case $s < 1$.
Conversely, let $\kappa_s(X_n, X) \to 0$. By Lemma 2.10, $X_n\|X_n\|^{s-1} \to X\|X\|^{s-1}$ in distribution and $\mathbb{E}[\|X_n\|^s] \to \mathbb{E}[\|X\|^s]$. Let Y_n, Y be random variables with $\mathcal{L}(Y_n) = \mathcal{L}(X_n\|X_n\|^{s-1}), \mathcal{L}(Y) = \mathcal{L}(X\|X\|^{s-1})$ for all n and $Y_n \to Y$ almost surely. Then $Y_n/\|Y_n\|^{(s-1)/s} \to Y/\|Y\|^{(s-1)/s}$ almost surely, hence $X_n \to X$ in distribution. This shows $\ell_s(X_n, X) \to 0$ by Lemma 2.10 and completes the proof. □

Lemma A.5. *Let X, Y be real-valued random variable with $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$. Then*

$$\zeta_1(X, Y) = \ell_1(X, Y) = \int_0^1 |F_X^{-1}(u) - F_Y^{-1}(u)| du = \int_{-\infty}^{\infty} |F_X(u) - F_Y(u)| du.$$

Proof. By definition, we have $\zeta_1(X,Y) \le \ell_1(X,Y) \le \int_0^1 |F_X^{-1}(u) - F_Y^{-1}(u)|du$. The last equality in the assertion is easily seen by geometric arguments, hence it suffices to show that $\zeta_1(X,Y) = \int_{-\infty}^{\infty} |F_X(u) - F_Y(u)|du$. Let $\mu = \mathbb{P}_X - \mathbb{P}_Y$ and $F(u) = \mu((-\infty, u]) = F_X(u) - F_Y(u)$. Since μ has finite first moment, we have $uF_\mu(u) \to 0$ for $u \to \pm\infty$. For any 1-Lipschitz function f, using partial integration, this yields

$$\mathbb{E}\left[f(X) - f(Y)\right] = \int f(u)d\mu(u) = -\int F_\mu(u)f'(u)du.$$

It is well-known that any function f on the real line is Lipschitz with constant K if and only if f is differentiable almost everywhere with $|f'(u)| \le K$ for almost all u. Now, let $h : \mathbb{R} \to \{-1, 1\}$ be defined by $h(u) = 1$ for $F_X(u) \le F_Y(u)$ and $h(u) = -1$ otherwise. Let $f(x) = \int_0^x h(u)du$. Then, f is differentiable almost everywhere and $|f'(u)| = |h(u)| \le 1$ for almost all u. Hence, f is 1-Lipschitz and

$$\mathbb{E}\left[f(X) - f(Y)\right] = \int |F_\mu(u)|du.$$

This proves the assertion. □

The last Lemmas is concerned with the geometry of $\mathcal{C}[0,1]$.

Lemma A.6. *The function $\nu : \mathcal{C}[0,1] \to \mathbb{R}, \nu(x) = \|x\|$ is nowhere differentiable.*

Proof. The norm function is easily seen to be non-differentiable at zero in any Banach space. Moreover, the relation $\nu(\lambda x) = \lambda\nu(x)$ for all $\lambda > 0$ implies that we may restrict ourselves to the unit sphere. Let $x \in \mathcal{C}[0,1]$ with $\|x\| = 1$ and (ε_n) be a sequence of real numbers with $\varepsilon_n \downarrow 0$ as $n \to \infty$. Suppose ν was differentiable at x. For any $h \in \mathcal{C}[0,1]$, denote

$$\Delta(h) = \frac{\|x+h\| - 1 - D\nu(x)(h)}{\|h\|}.$$

By the Riesz representation theorem there exists a finite signed measure $\mu = \mu_x$ on the unit interval such that $D(\nu(x))(h) = \int h(t)d\mu(t)$. We first assume that there exists $t^* > 0$ with $x(t^*) = 1$. Then it is possible to choose $a < t^*$ such that $x|_{[a,t^*)} > 0$ and $\mu(\{a\}) = 0$. For all n large enough, define $h_n \in \mathcal{C}[0,1]$ by $h_n(t) = \varepsilon_n$ for $t \in [a, t^*]$, $h_n(t) = 0$ for $t \le a - \varepsilon_n$ or $t > t^* + \varepsilon_n$ and linear in between. This implies

$$\Delta(h_n) = \frac{\varepsilon_n - \int_{[a-\varepsilon_n,a]} h_n(t)d\mu(t) - \int_{(a,t^*]} h_n(t)d\mu(t) - \int_{(t^*,t^*+\varepsilon_n]} h_n(t)d\mu(t)}{\varepsilon_n}.$$

By σ-continuity the first and third integrals are of order $o(\varepsilon_n)$, hence $\Delta(h_n) \to 0$ implies $\mu([a,t^*]) = 1$. The same arguments also imply $\mu([b,t^*]) = 1$ for any $a < b < t^*$ with $\mu(\{b\}) = 0$, in particular we have $\mu([a,b]) = 0$ for these values of b which gives $\mu|_{[a,t^*]} = \delta_{t^*}$.
Next, choose a sequence (a_n) from $[0,1]$ with $a < a_n < t^*$ such that $x|_{[a_n,t^*]} \in [1-\varepsilon_n, 1]$. Define a sequence of continuous functions h_n by $h_n(t) = 1 - x(t) + \varepsilon_n$ for $t \in [a_n, (a_n+t^*)/2]$, $h_n(t) = 0$ for $t \le a$ or $t \ge t^*$ and linear in between. Then, for all n large enough, $\|x + h_n\| \ge 1 + \varepsilon_n$, $\|h_n\| \le 2\varepsilon_n$ and thus $\Delta(h_n) \ge 1/2$ while $h_n \to 0$. This contradicts the differentiability of ν at x.
All remaining cases follow from two observations: First, any function f between two Banach spaces with $f(x) = f(-x)$ is differentiable at x if and only if it is at $-x$ and second, any function $f : \mathcal{C}[0,1] \to \mathbb{R}$ with $f(x) = f(\bar{x})$, where $\bar{x}(t) = x(1-t)$, is differentiable at x if it is at $\bar{x}$. □

Bibliography

[ABM12] G. Alsmeyer, J. D. Biggins, and M. Meiners. The functional equation of the smoothing transform. *to appear in Ann. Probab.*, 2012.

[Ald94] D. Aldous. Recursive self-similarity for random trees, random triangulations and Brownian excursion. *Ann. Probab.*, 22(2):527–545, 1994.

[Bar90] A. D. Barbour. Stein's method for diffusion approximations. *Probab. Theory Relat. Fields*, 84(3):297–322, 1990.

[Ben75] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18:509–517, 1975.

[BF81] P. J. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.*, 9(6):1196–1217, 1981.

[BHLT01] Z.-D. Bai, H.-K. Hwang, W.-Q. Liang, and T.-H. Tsai. Limit theorems for the number of maxima in random samples from planar regions. *Electron. J. Probab.*, 6:no. 3, 41 pp. (electronic), 2001.

[Bil68] P. Billingsley. *Convergence of probability measures.* New York-London-Sydney-Toronto: John Wiley and Sons, Inc. XII, 253 p. , 1968.

[Bil99] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[BJ09] A. D. Barbour and S. Janson. A functional combinatorial central limit theorem. *Electron. J. Probab.*, 14:2352–2370, 2009.

[BNS12] N. Broutin, R. Neininger, and H. Sulzbach. Partial match queries in random quadtrees. In Y. Rabani, editor, *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1056–1065, 2012.

[Bou73] P. Boulicaut. Convergence cylindrique et convergence etroite d'une suite de probabilités de Radon. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 28:43–52, 1973.

[BR85] V. Bentkus and A. Rachkauskas. Estimates of distances between sums of independent random elements in Banach spaces. *Theory Probab. Appl.*, 29:50–65, 1985.

[BS75] J. L. Bentley and D. F. Stanat. Analysis of range searches in quad trees. *Inf. Process. Lett.*, 3(6):170–173, 1975.

[Car71] H. Cartan. *Differential calculus*. Paris: Hermann, Publishers in Arts and Science; London: Kershaw Publishing Company, Ltd., 166 p. , 1971.

[CH03] H. H. Chern and H. K. Hwang. Partial match queries in random quadtrees. *SIAM J. Comput.*, 32:904–915, 2003.

[CH06] H. H. Chern and H. K. Hwang. Partial match queries in random k-d trees. *SIAM J. Comput.*, 35:1440–1466, 2006.

[CJ11] N. Curien and A. Joseph. Partial match queries in two-dimensional quadtrees: a probabilistic approach. *Adv. Appl. Probab.*, 43(1):178–194, 2011.

[CLG11] N. Curien and J.-F. Le Gall. Random recursive triangulations of the disk via fragmentation theory. *Ann. Prob.*, 39(6):2224–2270, 2011.

[Dal56] G. Dall'Aglio. Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Scuoloa Norm. Sup. Pisa (3)*, 10:35–74, 1956.

[Dev84] L. Devroye. Exponential bounds for the running time of a selection algorithm. *J. Comput. System Sci.*, 29(1):1–7, 1984.

[Dev87] L. Devroye. Branching processes in the analysis of the heights of trees. *Acta Informatica*, 24(3):277–298, 1987.

[Dev01] L. Devroye. On the probabilistic worst-case time of "find". *Algorithmica*, 31(3):291–303, 2001.

[DJN08] M. Drmota, S. Janson, and R. Neininger. A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.*, 18(1):288–333, 2008.

[DL90] L. Devroye and L. Laforest. An analysis of random d-dimensional quad trees. *SIAM J. Comput.*, 19(5):821–832, 1990.

[Don51] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.*, 6:12 p., 1951.

[Don52] M. D. Donsker. Justification and extension of Doob's heuristic approach to the Komogorov-Smirnov theorems. *Ann. Math. Statist.*, 23:277–281, 1952.

[Doo49] J. L. Doob. Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.*, 20:393–403, 1949.

[DR85] M. Denker and U. Rösler. Some contributions to Chernoff-Savage theorems. *Statist. Decisions*, 3(1-2):49–75, 1985.

[Dud66] R. M. Dudley. Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Ill. J. Math.*, 10:109–126, 1966.

[Dud67] R. M. Dudley. Measures on non-separable metric spaces. *Ill. J. Math.*, 11:449–453, 1967.

[Dud76] R. M. Dudley. *Probabilities and metrics*. Matematisk Institut, Aarhus Universitet, Aarhus, 1976.

[DZC04] L. Devroye and C. Zamora-Cura. Expected worst-case partial match in random quadtries. *Discrete Appl. Math.*, 141:103–117, 2004.

[EK46] P. Erdős and M. Kac. On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.*, 52:292–302, 1946.

[EK47] P. Erdős and M. Kac. On the number of positive sums of independent random variables. *Bull. Amer. Math. Soc.*, 53:1011–1020, 1947.

[Enf73] Per Enflo. A counterexample to the approximation problem in Banach spaces. *Acta Mathematica*, 130:309–317, 1973.

[ER07] K. Eickmeyer and L. Rüschendorf. A limit theorem for recursively defined processes in L^p. *Statist. Decisions*, 25(3):217–235, 2007.

[FB74] R. A. Finkel and J. L. Bentley. Quad trees, a data structure for retrieval on composite keys. *Acta Informatica*, 4:1–19, 1974.

[Fel71] W. Feller. *An Introduction to Probability Theory and its Applications*, volume II. Wiley, New York, 3rd edition, 1971.

[FGPR93] P. Flajolet, G. H. Gonnet, C. Puech, and J. M. Robson. Analytic variations on quadtrees. *Algorithmica*, 10:473–500, 1993.

[FJ02] J. A. Fill and S. Janson. Quicksort asymptotics. *J. Algorithms*, 44(1):4–28, 2002.

[FL94] P. Flajolet and T. Lafforgue. Search costs in quadtrees and singularity perturbation asymptotics. *Discrete Comput. Geom.*, 12(2):151–175, 1994.

[For49] R. Fortet. Quelques travaux récents sur le mouvement brownien. *Ann. Inst. H. Poincaré*, 11:175–226, 1949.

[FP86] P. Flajolet and C. Puech. Partial match retrieval of multidimensional data. *Jounal of the ACM*, 33(2):371–407, 1986.

[GL80] E. Giné and J. R. León. On the central limit theorem in Hilbert space. *Stochastica*, 4(1):43–71, 1980.

[GR96] R. Grübel and U. Rösler. Asymptotic distribution theory for Hoare's selection algorithm. *Adv. Appl. Probab.*, 28(1):252–269, 1996.

[Gro53] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

[HJP76] J. Hoffmann-Jørgensen and G. Pisier. The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probab.*, 4(4):587–599, 1976.

[Hoa61] C. A. R. Hoare. Partition: Algorithm 63, Quicksort: Algorithm 64, and Find: Algorithm 65. *Communications of the ACM*, 4(7):321–322, 1961.

[Hoa62] C. A. R. Hoare. Quicksort. *Comput. J.*, 5:10–15, 1962.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963.

[Jai77] N. C. Jain. Central limit theorem and related questions in Banach space. In *Probability (Proc. Sympos. Pure Math., Vol. XXXI, Univ. Illinois, Urbana, Ill., 1976)*, pages 55–65. Amer. Math. Soc., Providence, R.I., 1977.

[Jam77] G. I. Jamukov. Estimates for generalized Dudley metrics in spaces of finite-dimensional distributions. *Theory Probab. Appl.*, 22(3):579–584, 1977. With notes by V. M. Zolotarev.

[JK] S. Janson and S. Kaijser. Higher moments of Banach space valued random variables. *in preparation*.

[JN08] S. Janson and R. Neininger. The size of random fragmentation trees. *Probab. Theory Related Fields*, 142(3-4):399–442, 2008.

[Kac46] M. Kac. On the average of a certain Wiener functional and a related limit theorem in calculus of probability. *Trans. Amer. Math. Soc.*, 59:401–414, 1946.

[Kol31] A. Kolmogorov. Eine Verallgemeinerung des Laplace-Liapounoffschen Satzes. *Izv. Akad. Nauk SSSR, Otd. Mat. Estest. Nauk, VII. Ser.*, 7:959–962, 1931.

[Kol35] A. Kolmogorov. La transformation de Laplace dans les espaces linéaires. *C. R. Acad. Sci., Paris*, 200:1717–1718, 1935.

[KPS96] S. Kwapień, M. Pycia, and W. Schachermayer. A proof of conjecture of Bobkov and Houdre. *Electron. Comm. Probab.*, 1:no. 2, 7–10 (electronic), 1996.

[KR58] L. V. Kantorovič and G. Š. Rubinšteĭn. On a space of completely additive functions. *Vestnik Leningrad. Univ.*, 13(7):52–59, 1958.

[Kwa72] S. Kwapień. Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients. *Studia Math.*, 44:583–595, 1972.

[Led84] M. Ledoux. A remark on the central limit theorem in Banach spaces. In *Probability theory on vector spaces, III (Lublin, 1983)*, volume 1080 of *Lecture Notes in Math.*, pages 144–151. Springer, Berlin, 1984.

[LL65] R. G. Laha and E. Lukacs. On a linear form whose distribution is identical with that of a monomial. *Pacific J. Math.*, 15:207–214, 1965.

[LT91] Michel Ledoux and Michel Talagrand. *Probability in Banach spaces. Isoperimetry and processes.* Berlin etc.: Springer-Verlag, 1991.

[Maj78] P. Major. On the invariance principle for sums of independent identically distributed random variables. *J. Multivariate Anal.*, 8(4):487–517, 1978.

[Mar49] A. M. Mark. Some probability limit theorems. *Bull. Amer. Math. Soc.*, 55:885–900, 1949.

[Mou53] E. Mourier. Éléments aléatoires dans un espace de Banach. *Ann. Inst. Henri Poincaré*, 13:161–244, 1953.

[MS48] E. Marczewski and R. Sikorski. Measures in non-separable metric spaces. *Colloq. Math.*, 1:133–139, 1948.

[Nei01] R. Neininger. On a multivariate contraction method for random recursive structures with applications to quicksort. *Random Structures and Algorithms*, 19(3-4):498–524, 2001.

[Nei04] R. Neininger. Stochastische Analyse von Algorithmen, Fixpunktgleichungen und ideale Metriken. *Habilitation Thesis*, 2004.

[Nev75] J. Neveu. *Discrete-parameter martingales*. North-Holland Publishing Co., Amsterdam, revised edition, 1975.

[NR01] Ralph Neininger and Ludger Rüschendorf. Limit laws for partial match queries in quadtrees. *Ann. Appl. Probab.*, 11(2):452–469, 2001.

[NR02] R. Neininger and L. Rüschendorf. Rates of convergence for Quicksort. *J. Algorithms*, 44(1):52–62, 2002.

[NR04a] R. Neininger and L. Rüschendorf. On the contraction method with degenerate limit equation. *Ann. Probab.*, 32(3B):2838–2856, 2004.

[NR04b] R. Neininger and L. Rüschendorf. A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.*, 14(1):378–418, 2004.

[Pac05] J. Pachl. The sum of two measurable functions. page p. 4, 2005.

[Pes95] W. R. Pestman. Measurability of Linear Operators in the Skorokhod Topology. *Bull. Belg. Math. Soc*, 2:381–388, 1995.

[Pis11] G. Pisier. Grothendieck's theorem, past and present. page p. 87, 2011.

[Pit99] B. Pittel. Normal convergence problem? Two moments and a recurrence may be the clues. *Ann. Appl. Probab.*, 9(4):1260–1302, 1999.

[Pól23] G. Pólya. Herleitung des Gaußschen Fehlergesetzes aus einer Funktionalgleichung. *Math. Z.*, 18(1):96–108, 1923.

[Pol79] D. Pollard. Weak convergence on nonseparable metric spaces. *J. Austral. Math. Soc. Ser. A*, 28(2):197–204, 1979.

[Pro53] Y. V. Prohorov. Probability distributions in functional spaces. *Uspehi Matem. Nauk (N.S.)*, 8(3(55)):165–167, 1953.

[Pro56] Y. V. Prohorov. Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.*, 1:177–238, 1956.

[Rac91] S. T. Rachev. *Probability metrics and the stability of stochastic models*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1991.

[Rég89] M. Régnier. A limiting distribution for quicksort. *RAIRO Inform. Théor. Appl.*, 23:335–343, 1989.

[Riv76] Ronald L. Rivest. Partial-match retrieval algorithms. *SIAM J. Comput.*, 5(1):19–50, 1976.

[RR95] S. T. Rachev and L. Rüschendorf. Probability metrics and recursive algorithms. *Adv. Appl. Probab.*, 27(3):770–799, 1995.

[RR01] U. Rösler and L. Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29:3–33, 2001.

[Rüs06] L. Rüschendorf. On stochastic recursive equations of sum and max type. *J. Appl. Probab.*, 43(3):687–703, 2006.

[Rös91] U. Rösler. A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.*, 25:85–100, 1991.

[Rös92] U. Rösler. A fixed point theorem for distributions. *Stochastic Processes and their Applications*, 42(2):195 – 214, 1992.

[Rös01] U. Rösler. On the analysis of stochastic divide and conquer algorithms. *Algorithmica*, 29:238–261, 2001.

[Sam90a] H. Samet. *Applications of Spatial Data Structures: Computer Graphics, Image Processing, and GIS*. Addison-Wesley, Reading, MA, 1990.

[Sam90b] H. Samet. *The Design and Analysis of Spatial Data Structures*. Addison-Wesley, Reading, MA, 1990.

[Sam06] H. Samet. *Foundations of multidimensional and metric data structures*. Morgan Kaufmann, 2006.

[Sch95] W. Schachinger. The variance of a partial match retrieval in a multidimensional symmetric trie. *Random Structures Algorithms*, 7(1):81–95, 1995.

[Sch00] W. Schachinger. Limiting distributions for the costs of partial match retrievals in multidimensional tries. In *Proceedings of the Ninth International Conference "Random Structures and Algorithms" (Poznan, 1999)*, volume 17, pages 428–459, 2000.

[SD69] V. Strassen and R. M. Dudley. The central limit theorem and ε-entropy. In *Probability and Information Theory (Proc. Internat. Sympos., McMaster Univ., Hamilton, Ont., 1968)*, pages 224–231. Springer, Berlin, 1969.

[Sen81] V. V. Senatov. Some lower bounds for the rate of convergence in the central limit theorem in Hilbert space. *Dokl. Akad. Nauk SSSR*, 256(6):1318–1322, 1981.

[Sen83] V. V. Senatov. On a relation between Lévy-Prohorov metrics and ideal metrics. In *Stability problems for stochastic models (Moscow, 1982)*, volume 982 of *Lecture Notes in Math.*, pages 213–216. Springer, Berlin, 1983.

[Sen84] V. V. Senatov. An estimate of the Lévy-Prokhorov metric. *Theory Probab. Appl.*, 29(1):109–114, 1984.

[Sko56] A. V. Skorohod. Limit theorems for stochastic processes. *Theory Probab. Appl.*, 1:261–290, 1956.

[Tal78] M. Talagrand. Les boules peuvent-elles engendrer la tribu borélienne d'un espace métrisable non séparable? In *Séminaire Choquet, 17e année (1977/78), Initiation à l'analyse, Fasc. 2*, pages Comm. No. 5, 2. Secrétariat Math., Paris, 1978.

[Var62] S. Varadhan. Limit theorems for sums of independent random variables with values in a Hilbert space. *Sankhyā, Ser. A*, 24:213–238, 1962.

[Woy80] W. A. Woyczyński. On Marcinkiewicz-Zygmund laws of large numbers in Banach spaces and related rates of convergence. *Probab. Math. Statist.*, 1(2):117–131, 1980.

[Yur75] V.V. Yurinski. A smoothing inequality for estimates of the Lévy-Prohorov distance. *Theory Probab. Appl.*, 20:1–10, 1975.

[Zol76] V. M. Zolotarev. Approximation of distributions of sums of independent random variables with values in infinite-dimensional spaces. *Theory Probab. Appl.*, 21(1976):721–737, 1976.

[Zol77] V. M. Zolotarev. Ideal metrics in the problem of approximating distributions of sums of independent random variables. *Theory Probab. Appl.*, 22(1977):433–449, 1977.

[Zol78] V. M. Zolotarev. Metric distances in spaces of random variables and their distributions. *Math. USSR, Sb.*, 30(1976):373–401, 1978.

[Zol79a] V. M. Zolotarev. Ideal metrics in the problems of probability theory and mathematical statistics. *Aust. J. Stat.*, 21:193–208, 1979.

[Zol79b] V. M. Zolotarev. Properties and relations of certain types of metrics. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 87:18–35, 206, 212, 1979. Studies in mathematical statistics, 3.

[Zol84] V. M. Zolotarev. Probability metrics. *Theory Probab. Appl.*, 28:278–302, 1984.

[Zol97] V. M. Zolotarev. *Modern theory of summation of random variables*. Modern Probability and Statistics. VSP, Utrecht, 1997.

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit einer Verallgemeinerung der Kontraktionsmethode auf Zufallsgrößen mit Werten in unendlichdimensionalen topologischen Vektorräumen. Dabei liefert die Kontraktionsmethode einen Zugang um das asymptotische Verhalten von Folgen von Zufallsvariablen zu analysieren, die einer Rekursion der folgenden Art unterliegen

$$Y_n \stackrel{d}{=} \sum_{r=1}^{K} A_r(n) Y^{(r)}_{I_r^{(n)}} + b(n), \quad n \geq n_0. \tag{A.2}$$

Hierbei sind einige Annahmen zu treffen: $(Y_j^{(r)})_{j\geq 0}$ für $r = 1, \dots, K$ sind unabhängige Kopien der Folge $(Y_n)_{n\geq 0}$, $K \geq 1$ und $n_0 \geq 0$ sind natürliche Zahlen, $I^{(n)} = (I_1^{(n)}, \dots, I_K^{(n)})$ ist zufälliger Vektor mit Werten in $\{0, \dots, n\}$ und die zufälligen Folgen $(Y_j^{(1)})_{j\geq 0}, \dots, (Y_j^{(K)})_{j\geq 0}$ und $(A_1(n), \dots, A_K(n), b(n), I^{(n)})$ sind stochastisch unabhängig. Für reellwertige Folgen von Zufallsvariablen (Y_n) findet sich in der Literatur eine Vielzahl von Beispielen für ebensolche Rekursionen, die zumeist auf dem Gebiet der probabilistischen Analyse von Algorithmen oder im Studium von zufälligen Bäumen auftreten. In dieser Arbeit betrachten wir Rekursionen vom Typ (A.2) für Folgen (Y_n) mit Werten in einem topologischen Vektorraum B, wobei wir hauptsächlich an dem Fall $\mathcal{C}[0,1]$, der stetigen Funktionen auf dem Einheitsintervall ausgestattet mit der Supremumsnorm $\|f\| = \sup_{x\in[0,1]} |f(x)|$, oder $\mathcal{D}[0,1]$, der rechtsstetigen Funktionen mit linken Grenzwerten auf dem Einheitsintervall ausgestattet mit der Skorohod Topologie, interessiert sind. In diesem Fall sind $A_1(n), \dots, A_K(n)$ zufällige Endomorphismen auf dem Zustandsraums B. In der Regel geht man von der Folge (Y_n) durch Zentrierung und Normierung auf eine skalierte Größe X_n über, die nach Konstruktion eine ähnliche Rekursion erfüllt, welche sich typischerweise direkt aus (A.2) bestimmen lässt.

$$X_n \stackrel{d}{=} \sum_{r=1}^{K} A_r^{(n)} X^{(r)}_{I_r^{(n)}} + b^{(n)}, \quad n \geq n_0. \tag{A.3}$$

Die Kontraktionsmethode stellt einen Zugang zur Verteilungskonvergenz von (X_n) dar. Sie wurde in ihrer grundlegenden Form in der Pionierarbeit von Rösler [Rös91] über die Anzahl der Schlüsselvergleiche des randomisierten Quicksort-Algorithmus entwickelt. Die Methode beruht dabei maßgeblich auf der Beobachtung, dass die modifizierten Koeffizienten in (A.3) resp. die Folge $b^{(n)}$ in einem geeigneten Sinne gegen zufällige Operatoren konvergieren resp. eine Zufallsvariable konvergieren,

$$A_r^{(n)} \to A_r, \quad b^{(n)} \to b. \tag{A.4}$$

Diese Erkenntnis lässt unmittelbar vermuten, dass ein möglicher Grenzwert X von (X_n) der stochastischen Fixpunktgleichung

$$X \stackrel{d}{=} \sum_{r=1}^{K} A_r X^{(r)} + b \tag{A.5}$$

genügt. Im Kontext der Kontraktionsmethode wird der Gedankengang umgekehrt: Zunächst wird die Existenz einer (unter Nebenbedingungen) eindeutigen Lösung der obigen Fixpunktgleichung bewiesen und im zweiten Schritt Verteilungskonvergenz von X_n gegen eben diese Verteilung gezeigt. Zu diesem Zweck betrachtet man Lösungen von (A.5) als Fixpunkte der folgenden Abbildung

$$T : \mathcal{M}(B) \to \mathcal{M}(B)$$
$$T(\mu) = \mathcal{L}\left(\sum_{r=1}^{K} A_r Z^{(r)} + b\right), \tag{A.6}$$

wobei $(A_1, \ldots, A_K, b), Z^{(1)}, \ldots, Z^{(K)}$ unabhängig sind und $Z^{(1)}, \ldots, Z^{(K)}$ Verteilung μ tragen. Mit $\mathcal{M}(B)$ wird dabei die Menge aller Wahrscheinlichkeitsmaße auf B bezeichnet. Die Umsetzung dieser Heuristik basiert nun auf der Wahl eines passenden Teilraums $\mathcal{M}'(B)$ von $\mathcal{M}(B)$, auf dem T zur Selbstabbildung wird, und einer geeigneten Metrik auf $\mathcal{M}'(B)$ von $\mathcal{M}(B)$, bezüglich der T eine kontrahierende Abbildung ist. Ist die Metrik vollständig auf dem betrachteten Teilraum, so impliziert dies direkt die Existenz eines Fixpunkts von T nach dem Fixpunktsatz von Banach. Bei ausreichender Güte der Konvergenz (A.4) lässt sich im nächsten Schritt mit ähnlichen Argumenten die Konvergenz der Verteilung von X_n in der gewählten Metrik gegen den Fixpunkt zeigen. Impliziert die Konvergenz in der gewählten Metrik schwache Konvergenz, so ist die gewünschte Verteilungskonvergenz von X_n nachgewiesen und ihr Grenzwert über (A.5) charakterisiert. An dieser Stelle muss erwähnt werden, dass, abhängig von der jeweiligen Anwendung, asymptotische Resultate über den Erwartungswert oder gar die Varianz von Y_n a priori bekannt sein müssen um die Kontraktionsmethode erfolgreich umzusetzen.
Die Kontraktionsmethode wurde zunächst hauptsächlich auf den minimalen ℓ_p Metriken basierend entwickelt [Rös92, RR95, Rös01, RR01, Nei01] [für eine Definition dieser Metriken, siehe Seite 18]. In den letzten Jahren haben sich andere Metriken als vorteilhafter erwiesen, allen voran die Klasse der idealen Metriken ζ_s, die Ende der siebziger Jahre in einer Reihe von Arbeiten von Zolotarev [Zol76, Zol78] eingeführt wurde. In ihrem Zusammenhang spielt der Reichtum an reellwertigen Fréchet differenzierbaren Funktionen auf B eine entscheidende Rolle. Mit der Zerlegung $m = \lceil s \rceil - 1, \alpha = s - m$, definiert man für separable Banachräume zunächst

$$\mathcal{F}_s := \{f : B \to \mathbb{R} : \|D^m f(x) - D^m f(y)\| \leq \|x - y\|^\alpha \, \forall\, x, y \in B\}$$

und setzt dann

$$\zeta_s(\mu, \nu) = \sup_{f \in \mathcal{F}_s} |\mathbb{E}\left[f(X) - f(Y)\right]|,$$

für $\mu, \nu \in \mathcal{M}(B)$ wobei X und Y B-wertige Zufallsvariablen mit Verteilungen μ respektive ν sind. ζ_s Metriken wurden im Kontext der Kontraktionsmethode bereits in [RR95, RR01] verwendet, ihre Nützlichkeit beruht dabei maßgeblich auf der Idealeigenschaft, cf. 2.5. Ein vollständiger systematischer Zugang, der ihre Vorteile im Vergleich zu den minimalen ℓ_p Metriken hervorhebt, wurde von Neininger und Rüschendorf [NR04b] bereitgestellt. In separablen Hilberträumen wurde die Kontraktionsmethode auf ζ_s Metriken beruhend in [DJN08] eingeführt.

In dieser Arbeit wird die Kontraktionsmethode basierend auf den Zolotarev Metriken in erster Linie für separable Banachräume entwickelt. Zudem beschreiben wir einen Zugang im Fall von nicht-separablen Banachräumen, die mit einer kleineren σ-Algebra ausgestattet werden. In diesem

Fall sind wir hauptsächlich am Raum der càdlàg Funktionen, ausgestattet mit der Skorohod Topologie, interessiert. Dabei diskutieren wir ausführlich sämtliche relevante Eigenschaften der Zolotarev Metrik und klassifizieren Teilräume von $\mathcal{M}(B)$, in denen der ζ_s Abstand zwischen zwei Wahrscheinlichkeitsmaßen endlich ist. Geeignete obere Schranken, die es erlauben aus der Konvergenz in einer anderen Metrik auf Konvergenz in ζ_s zu schließen, sind in der Literatur bekannt und werden hier nur vorgestellt. Anders als in Hilberträumen, in denen Konvergenz in der Zolotarev Metrik schwache Konvergenz impliziert, diskutieren wir Banachräume mit schlechten geometrischen Eigenschaften, in denen sich im Zusammenhang mit dem zentralen Grenzwertsatz Folgen von Zufallsvariablen konstruieren lassen, die einerseits in ζ_s andererseits aber nicht in Verteilung konvergieren. Ein explizites Gegenbeispiel findet sich dabei beispielsweise in einer Arbeit von Strassen und Dudley [SD69] im Fall von $\mathcal{C}[0,1]$. Basierend auf einem Resultat von Barbour [Bar90] im Kontext der Stein'schen Methode für Diffusionen, lässt sich unter schwachen Zusatzbedingungen an die Pfade der betrachteten Folge und die Güte der ζ_s Konvergenz Verteilungskonvergenz ableiten. Letzteres ist für Folgen $\mathcal{D}[0,1]$-wertiger Zufallsvariablen darauf beschränkt, dass ihr Grenzwert stetige Pfade besitzt. Anders als in Hilberträumen ist Vollständigkeit der Metriken vom ζ_s Typ nicht bekannt; die kontrahierende Eigenschaft von T liefert demnach nur die Eindeutigkeit eines Fixpunkts, dessen Existenz anderweitig garantiert werden muss. Wir gehen darauf in den Anwendungen ein. Unser Hauptresultat des theoretischen Teil der Arbeit, Theorem 3.6, beschreibt abschließend einen funktionalen Grenzwertsatz für Rekursionen vom Typ (A.3) deren Grenzwert über (A.5) und Zusatzbedingungen charakterisiert werden kann. Dafür müssen einige technische Bedingungen erfüllt sein, wir skizzieren hier nur die wesentlichen:

- Kontraktionseigenschaft von T mittels $\sum_{r=1}^{K} \mathbb{E}\left[\|A_r\|^s\right] < 1$, wobei $\|\cdot\|$ die Operatornorm eines stetigen linearen Operators beschreibt.

- $\mathbb{E}\left[\|A_r^{(n)} - A\|^s\right] \to 0$ für $r = 1, \ldots, K$ und $\mathbb{E}\left[\|b^{(n)} - b\|^s\right] \to 0$, jeweils mit geeigneter Rate $R(n)$. In Anwendungen ist jede polynomielle Rate ausreichend.

- Existenz einer Lösung X von (A.5) mit $\zeta_s(X_n, X_m), \zeta_s(X_n, X) < \infty$ für alle n, m. Für letztere Bedingung beachte man Lemma 3.8.

- Existenz einer Folge stetiger (resp. càdlàg) Funktionen h_n mit $\|h_n - h\| \to 0$ und $h \in \mathcal{C}[0,1]$, so dass $X_n - h_n$ für $n \to \infty$ mit hoher Wahrscheinlichkeit auf Intervallen der Mindestlänge r_n linear (resp. konstant) ist. r_n und R_n stehen dabei im Zusammenhang

$$R(n) = o\left(\frac{1}{\log^m(1/r_n)}\right).$$

Für eine präzise Formulierung sämtlicher Bedingungen und des Theorems verweisen wir auf Abschnitt 3.1.

In Anwendungen spielt das asymptotische Verhalten von $\|Y_n\|$ eine Rolle. Prozesskonvergenz von (X_n) impliziert die Verteilungskonvergenz $\|X_n\| \to \|X\|$ für $n \to \infty$, unter schwachen zusätzlichen Bedingungen zeigen wir zudem Konvergenz aller Momente von $\|X_n\|$.

Im zweiten Teil der Arbeit beschreiben wir zwei unterschiedliche Anwendungen der entwickelten Methode. Zum einen eröffnet uns der Zugang der Kontraktionsmethode die Möglichkeit eines kurzen direkten Beweises von Donsker's Invarianzprinzip. Zum anderen geben wir einen funktionalen Grenzwertsatz im Kontext von partiellen Suchabfragen in zufälligen Bäumen, genauer Quadrantenbäumen und 2-d Bäumen, welcher einige offene Probleme über die Komplexität dieses Suchproblems löst. Der Beweis des Satzes von Donsker ist dabei aus Sicht der Kontraktionsmethode intuitiv und beruht auf einer zeitlichen Zerlegung der Prozesse, die zu einer interessanten Charakterisierung der Brownschen Bewegung führt. Wir zeigen, dass unter allen stetigen Prozessen mit $B_0 = 0$ und $\mathbb{E}[B_1] = 1$ die Brownsche Bewegung durch die stochastische Fixpunktgleichung

$$(B_t)_{t\in[0,1]} = \left(\sqrt{\frac{1}{2}}\left[\mathbf{1}_{\{t\le 1/2\}}B_{2t} + \mathbf{1}_{\{t>1/2\}}(B_1 + B_{2t-1})\right]\right)_{t\in[0,1]}$$

ausgezeichnet ist. Im Beweis des Grenzwertsatzes führen wir eine linearisierte Version der Brownschen Bewegung ein um die Endlichkeit der betrachteten ζ_s Abstände zu gewährleisten. Dies verkompliziert die Argumente unwesentlich.

Im letzten Kapitel der Arbeit werden zunächst Quadrantenbäume eingeführt, die als vergleichsbasierte Datenstruktur bei der Verarbeitung hochdimensionaler Datensätze fungieren. Quadranten- wie auch K-d Bäume lassen sich als mehrdimensionale Verallgemeinerung des Binärsuchbaums verstehen. Wir behandeln den Fall zweidimensionaler Daten, die in dem probabilistischen Modell durch unabhängige, uniform auf dem Einheitsquadrat verteilte Zufallsvariable realisiert werden. Das Ziel einer partiellen Suchabfrage in einem Baum der Größe n besteht darin, sämtliche Daten auszulesen, deren erste Komponente fixiert ist, während die zweite beliebige Werte annehmen kann (oder umgekehrt). Wird der Abfragewert ξ selbst zufällig, unabhängig vom Baum und uniform auf dem Einheitsintervall gewählt, so ist aus der Pionierarbeit von Flajolet et al. [FGPR93] auf diesem Gebiet bekannt, dass

$$\mathbb{E}[C_n(\xi)] \sim \kappa\, n^{\beta}, \qquad \beta = \frac{\sqrt{17}-3}{2}, \tag{A.7}$$

für eine Konstante $\kappa > 0$, siehe (1.15) auf Seite 10. Hierbei bezeichnet $C_n(s)$ die Anzahl besuchter Knoten einer Suche nach sämtlichen Einträgen mit erster Komponente s. Die rekursive Konstruktion des Baums, welche im Kapitel 1 ausführlich erläutert wird, erlaubt es eine Rekursion für die Anzahl besuchter Knoten bei der Ausführung der Suchabfrage auf dem Niveau von càdlàg Funktionen aufzustellen. Basierend auf dem Hauptresultat 3.6 zeigen wir einen funktionalen Grenzwertsatz für $(C_n(s))_{s\in[0,1]}$ nach Skalierung, die Charakterisierung des stetigen Grenzprozess $Z = \left(Z(s)_{s\in[0,1]}\right)$ wird dabei durch $\mathbb{E}[Z(s)] = (s(1-s))^{\beta/2}$, $\mathbb{E}\left[\|Z\|^2\right] < \infty$ und folgende stochastische Fixpunktgleichung geliefert:

$$\begin{aligned} Z(s)_{s\in[0,1]} \quad \stackrel{d}{=} \quad & \mathbf{1}_{\{s<U\}}\left[(UV)^{\beta}Z^{(1)}\left(\frac{s}{U}\right) + (U(1-V))^{\beta}Z^{(2)}\left(\frac{s}{U}\right)\right] \\ & +\mathbf{1}_{\{s\ge U\}}\left[((1-U)V)^{\beta}Z^{(3)}\left(\frac{1-s}{1-U}\right) + ((1-U)(1-V))^{\beta}Z^{(4)}\left(\frac{1-s}{1-U}\right)\right]. \end{aligned}$$

Eine Simulation des Grenzprozesses ist auf Seite 12 zu finden. Um unseren Zugang zu verwirklichen, konstruieren wir die Lösung Z als gleichmäßigen Grenzwert von stetigen Prozessen, welche punktweise die Martingaleigenschaft erfüllen. Dieser Teil der Arbeit beruht wesentlich auf Konzentrationsungleichungen vom Chernoff-Hoeffding Typ und geometrischen Eigenschaften von Quadrantenbäumen. Weiter verwenden wir ein Ergebnis von Chern und Hwang [CH03] über die Konvergenzrate in (A.7) und ein Resultat von Curien und Joseph [CJ11] über die Asymptotik des Erwartungswert

$$\mathbb{E}\left[C_n(s)\right] \sim K_1(s(1-s))^{\beta/2}n^{\beta},$$

mit $K_1 > 0$. Eine Verfeinerung letzterer asymptotischen Entwicklung erweist sich als notwendig um die gleichmäßige Konvergenz der Koeffizienten und des additiven Terms in der Rekursion zu garantieren. Sämtliche wesentliche Resultate werden in Abschnitt 5.1 aufgelistet, u.a. charakterisieren wir die eindimensionalen Randverteilungen von Z und zeigen die Konvergenz der Verteilung und aller Momente des Supremums des skalierten Prozesses, welches lange offene Fragen über das worst-case Verhalten des Algorithmus löst. Als einfache Folgerungen lösen wir damit zusätzlich die offenen Fragen nach Grenzwertsätzen und dem asymptotischen Verhalten der Varianz von $C_n(s)$ bei festem s und zufälligem $s = \xi$. Im hinteren Teil der Arbeit übertragen wir die Methode auf den verwandten Fall der 2-d Bäume, die Hauptresultate sind dabei von ähnlicher Natur. Im letzten Abschnitt erläutern wir ein offenes Problem aus der Arbeit von Curien und Le Gall [CLG11] im Kontext von zufälligen rekursiven Triangulierungen, das sich womöglich durch unsere Methoden lösen lässt. Wir geben dabei einige, die Arbeit [CLG11] ergänzende, Resultate, ein Beweis des angestrebten funktionalen Grenzwertsatzes erfordert weitere technische Abschätzungen.

Wesentliche Teile der Analyse der Komplexität von partiellen Suchabfragen in Quadrantenbäumen wurden bereits als “extended abstract“ in [BNS12] veröffentlicht.

Lebenslauf

Dipl.-Math. Henning Sulzbach
Mörfelder Landstrasse 121
60598 Frankfurt
Email: sulzbach@math.uni-frankfurt.de

Persönliche Angaben:

Geburtsdatum: 11. 10. 1983
Geburtsort: Frankfurt am Main
Staatsangehörigkeit: deutsch

Ausbildung:

07/1993 - 06/2002	Augustinergymnasium in Friedberg Abitur: Gesamtnote 1.3
10/2002 - 02/2007	Studium der Mathematik auf Diplom mit Nebenfach Theoretische Physik an der J. W. Goethe-Universität Frankfurt.
02/2005	Vordiplom “mit Auszeichnung”
02/2007	Diplom “mit Auszeichnung”, Diplomarbeit unter Betreuung von Prof. Neininger mit dem Titel “Das Profil zufälliger Binärsuchbäume”
seit 03/2007	Doktorand und wissenschaftlicher Mitarbeiter bei Prof. Neininger am Fachbereich Informatik und Mathematik der J. W. Goethe-Universität Frankfurt.

Akademische Lehrer:

Professoren und Hochschuldozenten der Goethe-Universität Frankfurt: Prof. Bieri, Prof. Burde, Prof. Johannson, HD Dr. Geiger, Prof. Kersting, Prof. Neininger, Prof. Reichert, Prof. Pflaum, Prof. Wakolbinger, Prof. Weth.

Bisherige Veröffentlichungen:

A functional limit law for the profile of plane-oriented recursive trees, *Fifth Colloquium on Mathematics and Computer Science*,339-350, Discrete Math. Theor. Comput. Sci. Proc., AI, *Assoc. Discrete Math. Theor. Comput. Sci., Nancy*, 2008.

Partial match queries in random quadtrees. joint work with N. Broutin and R. Neininger, *Proceedings of ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 1056–1065, 2012.

Besuchte Konferenzen und Workshops:

04/2007	Modern Perspectives in Real and Stochastic Analysis - Kaiserslautern
04/2007	Premières Journées Statistiques du Sud - Nizza
03/2008	8th German open conference on probability and statistics - Aachen, Vortrag mit dem Titel "On the profile of random trees"
09/2008	Fifth Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities - Blaubeuren, Vortrag mit dem Titel "The profile of random plane oriented recursive trees"
05/2009	Workshop: Combinatorics, Randomization, Algorithms and Probability - Montreal
05/2009	Fields-MITACS Summer School in Applied Probability, Ottawa
03/2010	9th German open conference on probability and statistics - Leipzig, Vortrag mit dem Titel "A new proof for Donsker's invariance principle "
03/2010	YEP VII: Probability, random trees and algorithms - Eindhoven, Vortrag mit dem Titel "A new proof of Donsker?s invariance principle"
09/2010	Limit behaviour of random graphs and related processes - Freiburg, Vortrag mit dem Titel "On contraction method in function spaces"
03/2011	6th Annual Workshop on Probabilistic Combinatorics and WVD - Barbados
04/2011	Mini-Workshop: Random Trees, Information and Algorithms - Oberwolfach, Vortrag mit dem Titel "A process convergence result for partial match queries in random quadtrees"
12/2011	Conference in the memory of Philippe Flajolet - Paris
03/2012	12th Latin American Congress of Probability and Mathematical Statistics - Viña del Mar, Vortrag mit dem Titel "Probabilistic analysis of a search tree problem"

Forschungsaufenthalte:

18. 10. - 22. 10. 2010	Centre de recherche INRIA - Paris, Rocquenquert
28. 11. - 16. 12. 2011	Centre de recherche INRIA - Paris, Rocquenquert
10. 04. - 26. 04. 2012	Centre de recherche INRIA - Paris, Rocquenquert