Point processes in the analysis of dependent data

Planinić, Hrvoje

University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF MATHEMATICS

Hrvoje Planinić

Point processes in the analysis of dependent data

DOCTORAL THESIS

Zagreb, 2019.

Sveučilište u Zagrebu

PRIRODOSLOVNO - MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Hrvoje Planinić

Točkovni procesi u analizi zavisnih podataka

DOKTORSKI RAD

Zagreb, 2019.

University of Zagreb

FACULTY OF SCIENCE
DEPARTMENT OF MATHEMATICS

Hrvoje Planinić

Point processes in the analysis of dependent data

DOCTORAL THESIS

Supervisors:
prof.dr.sc. Bojan Basrak
prof.dr.sc. Zoran Vondraček

Zagreb, 2019.

Sveučilište u Zagrebu

PRIRODOSLOVNO - MATEMATIČKI FAKULTET
MATEMATIČKI ODSJEK

Hrvoje Planinić

Točkovni procesi u analizi zavisnih podataka

DOKTORSKI RAD

Mentori:
prof.dr.sc. Bojan Basrak
prof.dr.sc. Zoran Vondraček

Zagreb, 2019.

To Marija

Preface

This thesis is centered around applications of the beautiful theory of point processes to problems concerning asymptotic behavior of extremes of a certain class of weakly dependent stationary time series and random fields. As usual in this context, the key role is played by the (compound) Poisson point process. The reason for this can be traced back to the so–called weak law of rare events, a basic principle in probability which states that, in a large family of independent events with each event having a small probability of occurring (i.e. being rare or extreme), the number of events which do occur is approximately Poisson distributed.

The reader is assumed to be familiar with basic facts of the theory of point processes, in particular with parts relevant to extreme value theory. If this is not the case, standard references are the books by Resnick [Res87, Res07]. I owe to them most of my knowledge and intuition on this subject. Other valuable references include Kallenberg [Kal17] (in which general random measures are considered) and Last and Penrose [LP17], or the more classical ones like Kingman [Kin93] and Karr [Kar91]. One should also mention Daley and Vere–Jones [DVJ03].

The thesis is divided into four chapters. Chapter 1 deals with the general theory of Poisson approximation for point processes. The point process framework is then used in Chapter 2 to describe the asymptotic behavior of extremes of a class of weakly dependent stationary time series and random fields which admit a so–called *tail process*. As it turns out, stationary series/fields satisfying the latter condition deserve to be called *regularly varying*. Finally, the tools developed in Chapter 2 are applied to study partial sums and record times of dependent time series, see Chapter 3, and to revisit the classical problem of local sequence alignment, see Chapter 4. Each chapter starts with an introduction which contains a brief motivation as well as description of its main results. Also, some of the proofs are often postponed to the end of the corresponding chapter or section.

Most of the results presented in this thesis can be found in the papers [BPS18, PS18, BP18a, BP18b] written jointly with Bojan Basrak and/or Philippe Soulier.

Acknowledgments

First of all, I sincerely thank my supervisor Bojan Basrak. His guidance made my PhD research pleasant and, at the same time, extremely interesting. I particularly enjoyed, and also largely benefited from, our discussions for which he would find time whenever I knocked on his door.

Furthermore, I am grateful to Zoran Vondraček for giving me the opportunity to start working in academia and for often sparing me long administrative work by kindly taking this painful job upon himself.

Also, I would especially like to thank Philippe Soulier for practically being my co–supervisor.

I would like to thank all the people at the department, it is a pleasure to come to work every day. Special thanks go to my roommate Braslav for numerous and always pleasant, especially non–math, conversations, and to Stjepan, for a number of conferences we attended together - it was always fun.

As for the world outside math, I thank all of my family and friends. It is truly a blessing to be surrounded by such people.

To my parents, my sister, and my brother: Simply, thank you for everything!

Most importantly, there is Marija. It is because of her that my life is filled with joy. Thank you for being my wife.

My final thanks go to dear God, to whom I owe everything.

I gratefully acknowledge the financial support of Croatian Science Foundation under the project 3526 during my PhD research.

Contents

Chapter 1

On (compound) Poisson approximation for point processes

1.1 Introduction

One of the basic principles of probability theory is the so–called weak law of rare events. It states that for independent Bernoulli random variables $X_{n,i}$, $i = 1, \ldots, n$, satisfying $\sup_{i=1,\ldots,n} \mathbb{P}(X_{n,i} = 1) \to 0$ (i.e. events $\{X_{n,i} = 1\}$ are *rare*) and $\sum_{i=1}^{n} \mathbb{P}(X_{n,i} = 1) \to \lambda$, the number of occurrences $\sum_{i=1}^{n} \mathbb{1}_{\{X_{n,i}=1\}}$ is asymptotically Poisson distributed with mean λ. Moreover, such a Poisson approximation holds even in the cases when there is some (but not too much) dependence between $X_{n,i}$'s, see e.g. the important paper by Arratia et al. [AGG89] and examples therein.

For general (independent) random elements $X_{n,i}$, results on convergence in distribution of point processes $\sum_{i=1}^{n} \delta_{X_{n,i}}$ to a so–called Poisson process can be seen as functional extensions of the weak law of rare events. Results of this kind go back to Grigelionis [Gri63] and a very general version of this result, which allows $X_{n,i}$'s to take values in a general Polish space, can be found in Kallenberg [Kal17]. The goal of this chapter is to extend this functional Poisson approximation to the case when $X_{n,i}$'s are dependent. Our motivation for studying point processes on a general, possibly infinite-dimensional, space (and also the reason for the "compound" in the title), actually comes from the problem of obtaining a compound Poisson or Poisson cluster limit for point processes based on a large class of stationary random fields, see Chapter 2.

For standard results from the theory of point processes (and more generally random measures) we will refer to Kallenberg [Kal17] and Resnick [Res87]. Note, even though the latter reference considers only point processes on a locally compact state space, most of the results transfer directly to general Polish case.

The rest of the chapter is organized as follows. In first two sections we discuss the notion of vague convergence of measures and the corresponding notion of convergence in distribution of random measures (and hence point processes). In particular, in Section 1.2 we propose an alternative view on the notion of vague convergence based on the abstract theory of boundedness developed by Hu [Hu66]. We think that such a view is intuitive and helps in clarifying the link between several different notions of convergence of measures used in the literature. In Section 1.3 we show that families of Lipschitz continuous functions determine convergence in distribution of random measures.

In Section 1.4, we present a general Poissonian approximation theorem for point processes on Polish spaces based on points which satisfy a suitable asymptotic (in)dependence condition. Also, we give sufficient conditions in the spirit of [AGG89] under which such dependence assumption is satisfied. These results are similiar to those obtained by Schuhmacher [Sch05]. However, while [Sch05] uses the Stein's method, our approach is based on exploiting the multiplicative strucuture of Laplace functionals of point processes. Ideas similar to ours can be found already in Banys [Ban80].

This chapter is based on the papers [BP18a] and [BP18b].

1.2 Vague convergence of locally finite measures

Let $\mathbb{X}$ be a Polish space, i.e. separable topological space which is metrizable by a complete metric. Denote by $\mathcal{B}(\mathbb{X})$ the corresponding Borel σ–field and choose a subfamily $\mathcal{B}_b(\mathbb{X}) \subseteq \mathcal{B}(\mathbb{X})$ of sets, called *bounded* (Borel) sets of $\mathbb{X}$. When there is no fear of confusion, we will simply write $\mathcal{B}$ and $\mathcal{B}_b$.

A Borel measure μ on $\mathbb{X}$ is said to be *locally (or boundedly) finite* if $\mu(B) < \infty$ for all $B \in \mathcal{B}_b$. Denote by $\mathcal{M}(\mathbb{X}) = \mathcal{M}(\mathbb{X}, \mathcal{B}_b)$ the space of all such measures. For measures $\mu, \mu_1, \mu_2, \ldots \in \mathcal{M}(\mathbb{X})$, we say that μ_n converge *vaguely* to μ and denote this by $\mu_n \xrightarrow{v} \mu$, if as $n \to \infty$,

$$\mu_n(f) = \int f d\mu_n \to \int f d\mu = \mu(f)\,, \tag{1.1}$$

for all bounded and continuous real–valued functions f on $\mathbb{X}$ with support being a bounded set. Denote by $CB_b(\mathbb{X})$ the family of all such functions and by $CB_b^+(\mathbb{X})$ the subset of all nonnegative functions in $CB_b(\mathbb{X})$.

Kallenberg [Kal17, Section 4] develops the theory of vague convergence under the assumption that $\mathcal{B}_b$ is the family all metrically bounded Borel sets w.r.t. a metric generating the topology of $\mathbb{X}$. This notion of convergence was also studied under the name of $w^{\#}$-convergence in Daley and Vere–Jones [DVJ03, Section A2.6].

In the what follows we propose an alternative, and arguably more intuitive, approach to the notion of vague convergence which also clarifies the connection between several well known notions of convergence found in the literature. It is based on the abstract theory of boundedness due to Hu [Hu66] which allows one to characterize all metrizable families of bounded sets.

1.2.1 The abstract concept of bounded sets

Following [Hu66, Section V.5], we say that a family of sets $\mathcal{B}_b \subseteq \mathcal{B}(\mathbb{X})$ is a (Borel) *boundedness* in $\mathbb{X}$ if (i) $A \subseteq B \in \mathcal{B}_b$ for $A \in \mathcal{B}$ implies $A \in \mathcal{B}_b$; (ii) $A, B \in \mathcal{B}_b$ implies $A \cup B \in \mathcal{B}_b$. A subfamily $\mathcal{C}_b$ of $\mathcal{B}_b$ is called a *basis* of $\mathcal{B}_b$ if every $B \in \mathcal{B}_b$ is contained in some $C \in \mathcal{C}_b$. Finally, boundedness $\mathcal{B}_b$ is said to be *proper* if it is adapted to the topology of $\mathbb{X}$ in the sense that for each $B \in \mathcal{B}_b$ there exists an open set $U \in \mathcal{B}_b$ such that $\overline{B} \subseteq U$, where $\overline{B}$ denotes the closure of B in $\mathbb{X}$.

To be consistent with the existing terminology of [Kal17, p. 19], we say that a boundedness *properly localizes* $\mathbb{X}$ if it is proper and has a countable basis which covers $\mathbb{X}$. If d metrizes $\mathbb{X}$ then the family of all sets in $\mathcal{B}$ with finite d–diameter is an example of a boundedness which properly localizes $\mathbb{X}$. By [Hu66, Corollary 5.12], this turns out to be the only example.

Theorem 1.2.1 ([Hu66, Corollary 5.12])**.** *Boundedness $\mathcal{B}_b$ properly localizes $\mathbb{X}$ if and only if there exists a metric on $\mathbb{X}$ which generates the topology of $\mathbb{X}$ and under which the metrically bounded Borel subsets of $\mathbb{X}$ coincide with $\mathcal{B}_b$.*

Remark 1.2.2. Note that [Hu66, Corollary 5.12] concerns boundednesses which also contain non–Borel subsets of $\mathbb{X}$. We restrict to Borel subsets since we work with Borel measures and it is easily seen that in this setting the conclusion of [Hu66, Corollary 5.12] still holds.

Observe, if $\mathcal{B}_b$ properly localizes $\mathbb{X}$ then one can find a basis $(K_m)_{m\in N}$ of $\mathcal{B}_b$ which consists of open sets and satisfies

$$\overline{K}_m \subseteq K_{m+1}\,, \qquad \text{for all } m \in \mathbb{N}\,. \tag{1.2}$$

Indeed, fix a metric which generates $\mathcal{B}_b$ and take (K_m) to be open balls around some fixed point and with radius strictly increasing to infinity; see also [Hu66, Lemma 5.9] for a direct argument which does not rely on Theorem 1.2.1. Conversely, if (K_m) is a sequence of open sets covering $\mathbb{X}$ and such that (1.2) holds, then the boundedness $\mathcal{B}_b = \{B \in \mathcal{B} : \exists m \in \mathbb{N} \text{ such that } B \subseteq K_m\}$ properly localizes $\mathbb{X}$. We refer to (K_m) as a *proper localizing sequence*.

Observe, by simply choosing a different family of bounded sets in $\mathbb{X}$ one alters the space of locally finite measures $\mathcal{M}(\mathbb{X})$ as well as the corresponding notion of vague convergence defined in (1.1). We mention a couple of important examples.

Example 1.2.3 (**Weak convergence**). By taking $\mathcal{B}_b$ to be the family of all Borel subsets of $\mathbb{X}$, we end up with the usual notion of weak convergence of finite measures on $\mathbb{X}$. Such a family $\mathcal{B}_b$ properly localizes $\mathbb{X}$ since taking $K_m = \mathbb{X}$ for all $m \in \mathbb{N}$ yields a proper localizing sequence.

Observe, it is easy to find a metric which metrizes $\mathbb{X}$ and generates $\mathcal{B}_b$ in this case. Simply choose any bounded metric which generates the topology of $\mathbb{X}$.

Example 1.2.4 (**Vague convergence of Radon measures**). When the space $\mathbb{X}$ is additionaly locally compact, by choosing $\mathcal{B}_b$ as the family of all relatively compact Borel subsets of $\mathbb{X}$ we obtain the well known notion of vague convergence of Radon measures on $\mathbb{X}$ as described in Kallenberg [Kal83] or Resnick [Res87]. Recall, a set is relatively compact if its closure is compact. Note that in this case, since $\mathbb{X}$ is locally compact, second countable and Hausdorff, one can find a sequence $(K_m)_{m\in\mathbb{N}}$ of relatively compact open subsets of $\mathbb{X}$ which cover $\mathbb{X}$ and satisfy (1.2). In particular, these K_m's form a basis for $\mathcal{B}_b$ and hence $\mathcal{B}_b$ properly localizes $\mathbb{X}$.

Example 1.2.5 (**Hult-Lindskog convergence**). Let $(\mathbb{X}', d')$ be a complete and separable metric space. In the theory of regularly varying random variables and processes, the sets of interest, i.e. bounded sets, are usually those which are actually bounded away from some fixed closed set $\mathbb{C} \subseteq \mathbb{X}'$.

More precisely, assume that $\mathbb{X}$ is of the form $\mathbb{X} = \mathbb{X}' \setminus \mathbb{C}$ equipped with the subspace topology and set $\mathcal{B}_b$ to be the class of all Borel sets $B \subseteq \mathbb{X}$ such that for some $\epsilon > 0$, $d'(x, \mathbb{C}) > \epsilon$ for all $x \in B$, where $d'(x, \mathbb{C}) = \inf\{d'(x, z) : z \in \mathbb{C}\}$. In this way, we obtain the notion of the so–called $\mathbb{M}_{\mathbb{O}}$-convergence (where $\mathbb{O} = \mathbb{X}$) as discussed in Lindskog et al. [LRR14] and originally introduced by Hult and Lindskog [HL06]. Observe, such $\mathcal{B}_b$ properly localizes $\mathbb{X}$ since one can take $K_m = \{x \in \mathbb{X} : d'(x, \mathbb{C}) > 1/m\}$, $m \in \mathbb{N}$, as a proper localizing sequence.

Remark 1.2.6. As observed by [Kal17, p. 125], under the notation of the previous example, one metric d which is topologically equivalent to d' and generates $\mathcal{B}_b$ is given by

$$d(x, y) = (d'(x, y) \wedge 1) \vee |1/d'(x, \mathbb{C}) - 1/d'(y, \mathbb{C})| \ , \ x, y \in \mathbb{X} .$$

In fact, this construction illustrates the basic idea in the proof of Theorem 1.2.1, see [Hu66, Theorem 5.11].

Remark 1.2.7. By the proof of [Hu66, Theorem 5.11], if $\mathbb{X}$ is completely metrizable and $\mathcal{B}_b$ properly localizes $\mathbb{X}$, one can assume that the metric which generates $\mathcal{B}_b$ is also complete.

In the rest of the chapter we will always assume that the space $\mathbb{X}$ is properly localized by the given family of bounded sets $\mathcal{B}_b$. In this case, Theorem 1.2.1 guarantees existence of at least one metric which generates $\mathcal{B}_b$ and this enables us to directly translate the results of [Kal17] to results concerning vague convergence of locally finite measures on an arbitrary properly localized space. In particular, by the so–called Portmanteau theorem (see [Kal17, Lemma 4.1]), $\mu_n \xrightarrow{v} \mu$ in $\mathcal{M}(\mathbb{X})$ is equivalent to convergence

$$\mu_n(B) \to \mu(B) \tag{1.3}$$

for all $B \in \mathcal{B}_b$ with $\mu(\partial B) = 0$, where ∂B denotes the boundary of the set B.

1.2.2 Vague convergence of point measures

Denote by δ_x the Dirac measure concentrated at $x \in \mathbb{X}$. A *(locally finite) point measure* on $\mathbb{X}$ is a measure $\mu \in \mathcal{M}(\mathbb{X})$ which is of the form $\mu = \sum_{i=1}^{K} \delta_{x_i}$ for some $K \in \{0, 1, \dots\} \cup \{\infty\}$ and (not necessarily distinct) points $x_1, x_2, \dots, x_K$ in $\mathbb{X}$. Note that by definition at most finitely many x_i's fall into every bounded set $B \in \mathcal{B}_b$. Denote by $\mathcal{M}_p(\mathbb{X})$ the space of all point measures on $\mathbb{X}$. Equivalently, one can define point measures as integer–valued measures in $\mathcal{M}(\mathbb{X})$, see e.g. [Res87, Exercise 3.4.2].

The following result, which is a simple consequence of (1.3), characterizes vague convergence in the case of point measures; for the proof see [Res87, Proposition 3.13]. cf. also [EMD16, Lemma 2.1]. It is fundamental when applying continuous mapping arguments to results on convergence in distribution of point processes (i.e. random point measures), see e.g. the proof of [Res07, Theorem 7.1].

Proposition 1.2.8. *Let $\mu, \mu_1, \mu_2, \dots \in \mathcal{M}_p(\mathbb{X})$ be point measures. Then $\mu_n \xrightarrow{v} \mu$ implies that for every $B \in \mathcal{B}_b$ such that $\mu(\partial B) = 0$ there exist $k, n_0 \in \mathbb{N}$ and points $x_i^{(n)}, x_i$, $n \geq n_0, i = 1, \dots, k$, in B such that for all $n \geq n_0$,*

$$\mu_n|_B = \sum_{i=1}^{k} \delta_{x_i^{(n)}} \quad \textit{and} \quad \mu|_B = \sum_{i=1}^{k} \delta_{x_i} \,,$$

and for all $i = 1, \dots, k$,

$$x_i^{(n)} \to x_i \quad \textit{in} \quad \mathbb{X} \,,$$

where $\mu_n|_B$ and $\mu|_B$ denote restriction of measures μ_n and μ, respectively, to the set B. Conversely, to show that $\mu_n \xrightarrow{v} \mu$, it is sufficient to check convergence of points in sets B from any base $\mathcal{C}_b$ of $\mathcal{B}_b$.

1.2.3 A comment on metrizability of the vague topology

In order to consider convergence in distribution in $\mathcal{M}(\mathbb{X})$ and $\mathcal{M}_p(\mathbb{X})$ related to the notion of vague convergence, one needs to impose a suitable topology on $\mathcal{M}(\mathbb{X})$. One such topology on $\mathcal{M}(\mathbb{X})$ is the smallest topology under which the maps $\mu \mapsto \mu(f)$ are continuous for all $f \in CB_b^+(\mathbb{X})$. Equivalently, this is the topology obtained by taking sets of the form

$$\{\nu \in \mathcal{M}(\mathbb{X}) : |\mu(f_i) - \nu(f_i)| < \epsilon \text{ for all } i = 1, \dots, k\} \tag{1.4}$$

for $\epsilon > 0, k \in \mathbb{N}$ and $f_1, \dots, f_k \in CB_b^+(\mathbb{X})$, to be the neighborhood base of $\mu \in \mathcal{M}(\mathbb{X})$. We call this topology the *vague* topology. Notice that, by definition, measures μ_n converge to a measure μ with respect to the vague topology if and only if $\mu_n \xrightarrow{v} \mu$.

It is shown in [Kal17, Theorem 4.2] that there exists a metric ρ on $\mathcal{M}(\mathbb{X})$ with the property that $\mu_n \xrightarrow{v} \mu$ in $\mathcal{M}(\mathbb{X})$ if and only if $\rho(\mu_n, \mu) \to 0$. Moreover, the metric space $(\mathcal{M}(\mathbb{X}), \rho)$ is complete and separable.

It is now tempting to immediately conclude that the topology generated by the metric ρ coincides with the vague topology and consequently that the vague topology is Polish. However, one can not deduce this without knowing a priori that the vague topology, i.e. the topology generated by the sets (1.4), is *sequential*, i.e. completely determined by its converging sequences, see [Fra65] (cf. also [Dud64]). For example, any first countable and hence any metrizable space is sequential. Since we have not been able to find such an argument in [Kal17] or anywhere else in the literature, we provide one here but omit some technical details.

Proposition 1.2.9. *The space $\mathcal{M}(\mathbb{X})$ equipped with the vague topology is metrizable and hence Polish.*

Sketch of the proof. Consider the space $\widehat{\mathcal{M}}(\mathbb{X})$ of all finite Borel measures on $\mathbb{X}$, i.e. Borel measures μ on $\mathbb{X}$ such that $\mu(\mathbb{X}) < \infty$. Equip $\widehat{\mathcal{M}}(\mathbb{X})$ with the smallest topology under which the maps $\mu \mapsto \mu(f)$ are continuous for all nonnegative and bounded continuous functions f on $\mathbb{X}$. This topology is usually called the *weak* topology. Also, let $\widehat{\mathcal{M}}_1(\mathbb{X}) \subseteq \widehat{\mathcal{M}}(\mathbb{X})$ be the subset of all probability measures equipped with the relative topology.

By [Bil68, Appendix III, Theorem 5], there exists a metric $\widehat{\rho}_1$ on $\widehat{\mathcal{M}}_1(\mathbb{X})$ which generates the weak topology and we can assume that $\widehat{\rho}_1$ is bounded by 1. It is then possible, but painful, to show that the function ρ on $\widehat{\mathcal{M}}(\mathbb{X}) \times \widehat{\mathcal{M}}(\mathbb{X})$ given by

$$\widehat{\rho}(\mu, \nu) = |\mu(\mathbb{X}) - \nu(\mathbb{X})| + (\mu(\mathbb{X}) \wedge \nu(\mathbb{X})) \cdot \widehat{\rho}_1(\mu(\,\cdot\,)/\mu(\mathbb{X}), \nu(\,\cdot\,)/\nu(\mathbb{X}))$$

is a proper metric which generates the weak topology on $\widehat{\mathcal{M}}(\mathbb{X})$.

Further, take a proper localizing sequence $(K_m)_{m\in\mathbb{N}}$. In particular (1.2) holds, and since $\mathbb{X}$ is a metric space, for every $m \in \mathbb{N}$ one can find a continuous function g_m on $\mathbb{X}$ such that $\mathbb{1}_{\overline{K}_m} \le g_m \le \mathbb{1}_{K_{m+1}}$. Clearly, $(g_m) \subseteq CB_b^+(\mathbb{X})$.

For every $m \in \mathbb{N}$ define a mapping $T_m : \mathcal{M}(\mathbb{X}) \to \widehat{\mathcal{M}}(\mathbb{X})$ such that $T_m(\mu)$ is the (unique) measure satisfying $T_m(\mu)(f) = \mu(f \cdot g_m)$ for all nonnegative and bounded functions f on $\mathbb{X}$. By the definitions of vague and weak topologies on the spaces $\mathcal{M}(\mathbb{X})$ and $\widehat{\mathcal{M}}(\mathbb{X})$, respectively, vague topology on $\mathcal{M}(\mathbb{X})$ coincides with the smallest topology under which the maps T_m, $m \in \mathbb{N}$, are continuous. Moreover, if $\mu \neq \nu$ for $\mu, \nu \in \mathcal{M}(\mathbb{X})$ then necessarily $T_m(\mu) \neq T_m(\nu)$ for some $m \in \mathbb{N}$. The last to two facts imply that the map $\mu \mapsto (T_m(\mu))_{m\in\mathbb{N}}$ is a homeomorphism between $\mathcal{M}(\mathbb{X})$ equipped with the vague topology and a subset of the space $\widehat{\mathcal{M}}(\mathbb{X})^{\mathbb{N}}$ equipped with the product weak topology, see e.g. [Wil04, Theorem 8.12]. Since the latter space is metrizable, it follows that the vague topology on $\mathcal{M}(\mathbb{X})$ is metrizable and in particular sequential, hence Polish.

$\square$

1.3 Convergence in distribution of random measures and point processes

A *random measure* on $\mathbb{X}$ is a random element in $\mathcal{M}(\mathbb{X})$ w.r.t. the smallest σ–algebra under which the maps $\mu \mapsto \mu(B)$ are measurable for all $B \in \mathcal{B}_b$. By [Kal17, Lemma 4.7], this σ–algebra equals the Borel σ–algebra on $\mathcal{M}(\mathbb{X})$ arising from the vague topology. By a *point process* on $\mathbb{X}$ we mean a random measure which is almost surely a point measure, i.e. element of the space $\mathcal{M}_p(\mathbb{X})$. Convergence in distribution in $\mathcal{M}(\mathbb{X})$ and $\mathcal{M}_p(\mathbb{X})$ is considered w.r.t. the vague topology and is denoted by "$\xrightarrow{d}$".

Even though our interest is only in convergence in distribution of point processes, in this section we treat general random measures since the proofs (and therefore the results) require no extra effort.

For random measures $N, N_1, N_2, \dots$ on $\mathbb{X}$, it is fundamental that $N_n \xrightarrow{d} N$ in $\mathcal{M}(\mathbb{X})$ if and only if $N_n(f) \xrightarrow{d} N(f)$ in $\mathbb{R}$ for all $f \in CB_b^+(\mathbb{X})$ which is further equivalent to convergence of Laplace functionals $\mathbb{E}[e^{-N_n(f)}] \to \mathbb{E}[e^{-N(f)}]$ for all $f \in CB_b^+(\mathbb{X})$, see [Kal17, Theorem 4.11]. We show that in latter two convergences it is sufficient to consider only functions which are Lipschitz continuous with respect to any suitable metric.

1.3.1 Lipschitz functions determine convergence in distribution

For any metric d on $\mathbb{X}$ denote by $LB_b^+(\mathbb{X}, d)$ the family of all bounded nonnegative functions f on $\mathbb{X}$ which have bounded support and are Lipschitz continuous with respect to d, i.e. for some constant $C > 0$

$$|f(x) - f(y)| \leq Cd(x, y) \ , \quad \text{for all } x, y \in \mathbb{X} \, .$$

Further, for a set $B \subseteq \mathbb{X}$ and $\epsilon > 0$ denote

$$B^\epsilon = B^\epsilon(d) = \{x \in \mathbb{X} : \ d(x, B) \leq \epsilon\} \, .$$

Proposition 1.3.1. *Assume that d is a metric on $\mathbb{X}$ which generates the corresponding topology and such that for any $B \in \mathcal{B}_b$ there exists an $\epsilon > 0$ such that $B^\epsilon \in \mathcal{B}_b$. Then $N_n \xrightarrow{d} N$ in $\mathcal{M}(\mathbb{X})$ if and only if $\mathbb{E}[e^{-N_n(f)}] \to \mathbb{E}[e^{-N(f)}]$ for every $f \in LB_b^+(\mathbb{X}, d)$.*

Remark 1.3.2. Every metric on $\mathbb{X}$ which generates the topology and the family of bounded sets (as in Theorem 1.2.1) satisfies the assumptions of the previous proposition.

Example 1.3.3. Consider the case from Example 1.2.5. Recall, $(\mathbb{X}', d')$ is assumed to be a complete and separable metric space and $\mathbb{C} \subseteq \mathbb{X}'$ a closed subset of $\mathbb{X}'$. The space $\mathbb{X} = \mathbb{X}' \setminus \mathbb{C}$ is equipped with the subspace topology (i.e. generated by d') and with $B \subseteq \mathbb{X}$ being bounded if and only if B is contained in $\{x \in \mathbb{X} : d'(x, \mathbb{C}) > 1/m\}$ for some $m \in \mathbb{N}$.

In this case, the metric d' generates the topology but the corresponding class of metrically bounded sets does not coincide with $\mathcal{B}_b$. Still, d' obviously satisfies the assumptions of the previous proposition.

Proof of Proposition 1.3.1. We only need to prove sufficiency. Take arbitrary $k \in \mathbb{N}$, $\lambda_1, \dots, \lambda_k \geq 0$ and $B_1, \dots, B_k \in \mathcal{B}_b$ such that $\mathbb{P}(N(\partial B) = 0) = 1$ for all $i = 1, \dots, k$. By [Kal17, Theorem 4.11], the result will follow if we show that

$$\lim_{n\to\infty} \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(B_i)}] = \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(B_i)}] \, . \tag{1.5}$$

For all $m \in \mathbb{N}, i = 1, \dots k$ and $x \in \mathbb{X}$ set

$$f^+_{m,i}(x) = 1 - (md(x, \overline{B}_i) \wedge 1) \ , \ f^-_{m,i}(x) = md(x, (B_i^\circ)^c) \wedge 1 \, , \tag{1.6}$$

B° denotes the interior of B. Using the elementary fact that for a closed set $C \subseteq \mathbb{X}$ and $x \in \mathbb{X}$, $x \in C$ if and only if $d(x, C) = 0$, it is straightforward to show that for all $i = 1, \dots, k$ as $m \to \infty$,

$$f^+_{m,i} \searrow \mathbb{1}_{\overline{B}_i} \text{ and } f^-_{m,i} \nearrow \mathbb{1}_{B_i^\circ} \, . \tag{1.7}$$

Functions $f^-_{m,i}$ obviously have bounded support for all i and m, and since $f^+_{m,i} \leq \mathbb{1}_{\{x\in\mathbb{X} \,:\, d(x,\overline{B}_i)\leq 1/m\}}$, by assumption on the metric d, $f^+_{m,i}$ has bounded support for m large enough. Assume without loss of generality that this is true for all $m \in \mathbb{N}$. Further, it is not difficult to show that for all i, m and all $x, y \in \mathbb{X}$

$$|f^+_{m,i}(x) - f^+_{m,i}(y)| \vee |f^-_{m,i}(x) - f^-_{m,i}(y)| \leq md(x, y) \, .$$

Hence, $f^+_{m,i}$ and $f^-_{m,i}$ are elements of $LB_b^+(\mathbb{X}, d)$.

By the monotone and the dominated convergence theorem,

$$\lim_{m\to\infty} \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(f^-_{m,i})}] = \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(B_i^\circ)}] = \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(B_i)}] \, ,$$

where the last equality follows since we assumed $N(\partial B_i) = 0$ a.s. for all i. Since $\sum_{i=1}^k \lambda_i f^-_{m,i}$ is again in $LB_b^+(\mathbb{X}, d)$, convergence in (1.5) will follow if we prove that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left| \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(B_i)}] - \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(f^-_{m,i})}] \right| = 0 \, . \tag{1.8}$$

Since for all i and m, $f^-_{m,i} \leq \mathbb{1}_{B_i} \leq f^+_{m,i}$,

$$0 \leq \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(f^-_{m,i})}] - \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(B_i)}] \leq \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(f^-_{m,i})}] - \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(f^+_{m,i})}] \, .$$

Notice that for for fixed i and all $m \in \mathbb{N}$, the support of $f^+_{m,i}$ is contained in the support of

$f_{1,i}^+$, which we assumed is a bounded set. Since N a.s. puts finite measure on such sets, applying the dominated convergence theorem twice yields that $\lim_{m\to\infty} \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(f_{m,i}^+)}] = \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(\overline{B}_i)}]$. Therefore,

$$\begin{aligned}\lim_{m\to\infty}\limsup_{n\to\infty}&\left|\mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(B_i)}] - \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N_n(f_{m,i}^-)}]\right| \\ &\le \lim_{m\to\infty} \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(f_{m,i}^-)}] - \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(f_{m,i}^+)}] \\ &= \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(B_i^\circ)}] - \mathbb{E}[e^{-\sum_{i=1}^k \lambda_i N(\overline{B}_i)}] = 0\,,\end{aligned}$$

where the last equality holds since $N(\partial B_i) = 0$ a.s. Hence, (1.8) holds and this finishes the proof. □

1.4 Poisson approximation for point processes on Polish spaces

Let $(I_n)_{n\in\mathbb{N}}$ be a sequence of finite index sets but such that $\lim_{n\to\infty} |I_n| = \infty$, where $|I_n|$ denotes the number of elements in I_n. For each $n \in \mathbb{N}$, let $(X_{n,i} : i \in I_n)$ be a family of random elements in a topological space $\mathbb{X}'$. Assume that there exists a Polish subset $\mathbb{X}$ of $\mathbb{X}'$ with a family of bounded Borel sets $\mathcal{B}_b = \mathcal{B}_b(\mathbb{X})$ such that, as $n \to \infty$,

$$\sup_{i\in I_n} \mathbb{P}(X_{n,i} \in B) \to 0\,,\ B \in \mathcal{B}_b\,. \tag{1.9}$$

The central theme of this section is convergence in distribution in $\mathcal{M}_p(\mathbb{X})$ (w.r.t. the vague topology) of point processes

$$N_n = \sum_{i\in I_n} \delta_{X_{n,i}}\,,\ n \in \mathbb{N}\,, \tag{1.10}$$

restricted to the space $\mathbb{X}$ (in words, N_n does not count $X_{n,i}$'s which fall outside $\mathbb{X}$). As usual, the key role is played by the Poisson point process. Recall, a point process N on $\mathbb{X}$ is a *Poisson (point) process* with intensity λ for some $\lambda \in \mathcal{M}(\mathbb{X})$ if (i) $N(B)$ is Poisson distributed with mean $\lambda(B)$ for each $B \in \mathcal{B}_b$; (ii) $N(B_1), \dots, N(B_k)$ are independent for all $B_1, \dots, B_k \in \mathcal{B}_b$ disjoint. The distribution of such Poisson process will be denoted by $\mathrm{PPP}(\lambda)$.

1.4.1 General Poisson approximation

Observe, if for each $n \in \mathbb{N}$, $(X_{n,i} : i \in I_n)$ were independent, (1.9) would imply that measures $\delta_{X_{n,i}}$ on $\mathbb{X}$, $n \in \mathbb{N}, i \in I_n$ form a null–array (see [Kal17, p. 129]) and by the so–called Grigelionis theorem (see [Kal17, Corollary 4.25]), for $\lambda \in \mathcal{M}(\mathbb{X})$, convergence $N_n \xrightarrow{d} N \overset{d}{=} \mathrm{PPP}(\lambda)$ holds in $\mathcal{M}_p(\mathbb{X})$ if and only if $\mathbb{E}[N_n(\,\cdot\,)] = \sum_{i\in I_n} \mathbb{P}(X_{n,i} \in \cdot) \xrightarrow{v} \lambda$

in $\mathcal{M}(\mathbb{X})$.

For general (i.e. dependent) $X_{n,i}$'s, one can still obtain the same Poisson limit if the asymptotic distributional behavior of N_n's is indistinguishable from its independent version. More precisely, let for each $n \in \mathbb{N}$, $(X^*_{n,i} : i \in I_n)$ be independent random elements such that for all $i \in I_n$, $X^*_{n,i}$ is distributed as $X_{n,i}$, and denote by $N^*_n = \sum_{i \in I_n} \delta_{X^*_{n,i}}$ the corresponding point process on $\mathbb{X}$.

Let $\mathcal{F}$ be a class of measurable and nonnegative functions on $\mathbb{X}$ with bounded support. We say that the family $(X_{n,i} : n \in \mathbb{N},\ i \in I_n)$ is *asymptotically $\mathcal{F}$–independent* ($AI(\mathcal{F})$) if

$$\left|\mathbb{E}\left[e^{-N_n(f)}\right] - \mathbb{E}\left[e^{-N^*_n(f)}\right]\right| = \left|\mathbb{E}\left[e^{-\sum_{i \in I_n} f(X_{n,i})}\right] - \prod_{i \in I_n} \mathbb{E}\left[e^{-f(X_{n,i})}\right]\right| \to 0, \text{ as } n \to \infty,$$

for all $f \in \mathcal{F}$, where we set $f(x) = 0$ for all $x \in \mathbb{X}' \setminus \mathbb{X}$. To obtain meaningful results we will require that functions from $\mathcal{F}$ determine convergence in distribution in $\mathcal{M}_p(\mathbb{X})$. More precisely, we say that a family $\mathcal{F} \subseteq CB^+_b(\mathbb{X})$ is *(point process) convergence determining* if, for point processes $N, N_1, N_2, \dots$ on $\mathbb{X}$, convergence $\mathbb{E}[e^{-N_n(f)}] \to \mathbb{E}[e^{-N(f)}]$ for all $f \in \mathcal{F}$ implies that $N_n \xrightarrow{d} N$ in $\mathcal{M}_p(\mathbb{X})$. For example, one can take the subfamily $\mathcal{F} \subseteq CB^+_b(\mathbb{X})$ of functions which are Lipschitz continuous as in Proposition 1.3.1. The following result is now immediate.

Theorem 1.4.1. *Assume that (1.9) holds and that there exists a measure $\lambda \in \mathcal{M}(\mathbb{X})$ such that, as $n \to \infty$,*

$$\sum_{i \in I_n} \mathbb{P}(X_{n,i} \in \cdot) \xrightarrow{v} \lambda\,. \tag{1.11}$$

Then for any convergence determining family $\mathcal{F}$, $N_n \xrightarrow{d} N \stackrel{d}{=} \mathrm{PPP}(\lambda)$ in $\mathcal{M}_p(\mathbb{X})$ if and only if $(X_{n,i} : n \in \mathbb{N},\ i \in I_n)$ is $AI(\mathcal{F})$.

Remark 1.4.2. One can allow $\mathcal{F}$ to contain functions f which are not necessarily continuous but are such that $N(disc(f)) = 0$ almost surely, where $disc(f)$ denotes the set of all discontinuity points of f. In particular, $\mathcal{F}$ can consist of nonnegative simple functions with bounded support, see [Kal17, Theorem 4.11].

Remark 1.4.3. Assume that (1.9) holds and that $X_{n,i}$'s are $AI(\mathcal{F})$ for some convergence determining family $\mathcal{F}$. In this case, if N_n converge in distribution to some limit, N say, then N is necessarily a Poisson process. Indeed, since also $N^*_n \xrightarrow{d} N$, by [Kal17, Theorem 4.22] N is infinitely divisible and moreover, by construction of N^*_n, its so–called Lévy measure (see [Kal17, p. 89]) is concentrated on the set $\{\delta_x : x \in \mathbb{X}\}$ which implies that N is Poisson.

Observe that condition $AI(\mathcal{F})$ is, in general, much weaker than simply requiring that $X_{n,i}$, $i \in I_n$ asymptotically behave as if they were independent. The key fact here is that all functions in $\mathcal{F}$ have bounded support so for every fixed $f \in \mathcal{F}$, $N_n(f)$ is affected only by the behavior of $X_{n,i}$'s which fall into a fixed bounded set. Sufficient conditions for $AI(\mathcal{F})$ to hold are given in Proposition 1.4.5 below.

Before that, we state a stationary version of the previous result, cf. [Res87, Proposition 3.21]. For $d \in \mathbb{N}$ consider the space $[0,1]^d \times \mathbb{X}$ with respect to the product topology and with $B' \in \mathcal{B}([0,1]^d \times \mathbb{X})$ being bounded if $\{x \in \mathbb{X} : (\boldsymbol{t}, x) \in B'\}$ is bounded in $\mathbb{X}$.

Corollary 1.4.4. *Assume that $I_n = \{1, 2, \dots, k_n\}^d \subseteq \mathbb{Z}^d$ for some $d \in \mathbb{N}$ with $k_n \to \infty$ and that $(X_{n,\boldsymbol{i}} : \boldsymbol{i} \in I_n)$ are identically distributed for every $n \in \mathbb{N}$. If there exists a measure $\nu \in \mathcal{M}(\mathbb{X})$ such that, as $n \to \infty$,*

$$k_n^d \mathbb{P}(X_{n,\mathbf{1}} \in \cdot) \xrightarrow{v} \nu\,, \tag{1.12}$$

then for any convergence determining family $\mathcal{F}'$ on $[0,1]^d \times \mathbb{X}$,

$$N_n' = \sum_{\boldsymbol{i} \in I_n} \delta_{(\boldsymbol{i}/k_n, X_{n,\boldsymbol{i}})} \xrightarrow{d} N' \stackrel{d}{=} \mathrm{PPP}(Leb \times \nu)$$

in $\mathcal{M}_p([0,1]^d \times \mathbb{X})$ if and only if $((\boldsymbol{i}/k_n, X_{n,\boldsymbol{i}}) : n \in \mathbb{N}, \boldsymbol{i} \in I_n)$ is $AI(\mathcal{F}')$, where Leb denotes the Lebesgue measure on $[0,1]^d$.

Proof. We simply apply Theorem 1.4.1 to random elements $X'_{n,\boldsymbol{i}} := (\boldsymbol{i}/k_n, X_{n,\boldsymbol{i}})$, $n \in \mathbb{N}, \boldsymbol{i} \in I_n$. Take an arbitrary $B' \in \mathcal{B}_b([0,1]^d \times \mathbb{X})$ and define $B = \{x \in \mathbb{X} : (\boldsymbol{t}, x) \in B'\}$. Since $B \in \mathcal{B}_b(\mathbb{X})$, (1.12) and [Kal17, Lemma 4.1(iv)] imply that

$$\limsup_{n\to\infty} \sum_{\boldsymbol{i} \in I_n} \mathbb{P}(X'_{n,\boldsymbol{i}} \in B') = \limsup_{n\to\infty} k_n^d \mathbb{P}(X_{n,\mathbf{1}} \in B) \le \nu(\overline{B}) < +\infty\,.$$

Hence, (1.9) holds since $k_n \to \infty$.

Further, note that for arbitrary $\boldsymbol{a} = (a_1, \dots, a_d)$ and $\boldsymbol{b} = (b_1, \dots, b_d)$ in $[0,1]^d$ such that $a_j \le b_j$ for all $j = 1, \dots, d$ and a set $B \in \mathcal{B}_b$ such that $\nu(\partial B) = 0$, (1.12) implies that as $n \to \infty$,

$$\sum_{\boldsymbol{i} \in I_n} \mathbb{P}(X'_{n,\boldsymbol{i}} \in (\boldsymbol{a}, \boldsymbol{b}] \times B) = \frac{1}{k_n^d} \prod_{j=1}^d \lfloor k_n(b_j - a_j) \rfloor \cdot k_n^d \mathbb{P}(X_{n,\mathbf{1}} \in B) \to \prod_{j=1}^d (b_j - a_j) \cdot \nu(B)$$

By [Kal17, Lemma 4.1], this implies that $\sum_{\boldsymbol{i} \in I_n} \mathbb{P}(X'_{n,\boldsymbol{i}} \in \cdot) \xrightarrow{v} Leb \times \nu$ in $\mathcal{M}([0,1]^d \times \mathbb{X})$, i.e. (1.11) holds with $\lambda = Leb \times \nu$. □

1.4.2 Sufficient condition for asymptotic $\mathcal{F}$–independence

For each $i \in I_n$, choose a subset of the index set $B_n(i) \subseteq I_n$ containing i, and call it the *neighborhood of dependence* of i. Intuitively, it will be useful to choose $B_n(i)$ as small as possible but such that $X_{n,i}$ is (nearly) independent of all $X_{n,j}$ for $j \notin B_n(i)$.

Select an arbitrary ordering of the elements in I_n. Without loss of generality, we will assume that $I_n = \{1, 2, \dots, m_n\}$ where $m_n \to \infty$ as $n \to \infty$. For all $i \in I_n$ partition $\{i+1, \dots, m_n\}$ into $\tilde{B}_n(i) := \{j \in B_n(i) : j > i\}$ and $\tilde{B}_n^c(i) := \{j \notin B_n(i) : j > i\}$. Further, fix an arbitrary countable base $(K_m)_{m\in\mathbb{N}} \subseteq \mathcal{B}_b$ of $\mathcal{B}_b$, i.e. for every $B \in \mathcal{B}_b$, $B \subseteq K_m$ for some $m \in \mathbb{N}$.

For a given neighborhood structure $(B_n(i) : n \in \mathbb{N},\, i \in I_n)$ and for all $m, n \in \mathbb{N}$ define

$$b_{n,1}^m = \sum_{i\in I_n} \sum_{j\in \tilde{B}_n(i)} \mathbb{P}(X_{n,i} \in K_m) \cdot \mathbb{P}(X_{n,j} \in K_m)\,,$$
$$b_{n,2}^m = \sum_{i\in I_n} \sum_{j\in \tilde{B}_n(i)} \mathbb{P}(X_{n,i} \in K_m, X_{n,j} \in K_m)\,.$$

Furthermore, for all $n \in \mathbb{N}$ and an arbitrary nonnegative measurable function f on $\mathbb{X}$ define

$$b_{n,3}(f) = \sum_{i\in I_n} \Big|\mathbb{E}\Big[e^{-f(X_{n,i})} \prod_{j\in\tilde{B}_n^c(i)} e^{-f(X_{n,j})}\Big] - \mathbb{E}\Big[e^{-f(X_{n,i})}\Big] \cdot \mathbb{E}\Big[\prod_{j\in\tilde{B}_n^c(i)} e^{-f(X_{n,j})}\Big]\Big|\,.$$

Proposition 1.4.5. *Let f be a nonnegative measurable function on $\mathbb{X}$ with bounded support. If $m \in \mathbb{N}$ is such that the support of f is contained in K_m, then for all $n \in \mathbb{N}$,*

$$\left|\mathbb{E}\left[e^{-\sum_{i\in I_n} f(X_{n,i})}\right] - \prod_{i\in I_n} \mathbb{E}\left[e^{-f(X_{n,i})}\right]\right| \le b_{n,1}^m + b_{n,2}^m + b_{n,3}(f)\,.$$

In particular, if there exists a neighborhood structure $(B_n(i) : n \in \mathbb{N},\, i \in I_n)$ such that for all $m \in \mathbb{N}$ and every $f \in \mathcal{F}$

$$\lim_{n\to\infty} b_{n,1}^m = \lim_{n\to\infty} b_{n,2}^m = \lim_{n\to\infty} b_{n,3}(f) = 0\,,$$

then the family $(X_{n,i} : n \in \mathbb{N},\, i \in I_n)$ is AI($\mathcal{F}$).

Proof. The proof is an adaptation of argument in Nakhapetyan [Nak88, Lemma 3], though the main idea goes back to [Ban80, Theorem 4]. Since e^{-f} is positive and bounded by 1 it follows that

$$\begin{aligned}&\left|\mathbb{E}\left[e^{-\sum_{i\in I_n} f(X_{n,i})}\right] - \prod_{i\in I_n} \mathbb{E}\left[e^{-f(X_{n,i})}\right]\right| \\ &\le \sum_{i=1}^{m_n-1} \Big|\mathbb{E}\Big[e^{-f(X_{n,i})} \prod_{j=i+1}^{m_n} e^{-f(X_{n,j})}\Big] - \mathbb{E}\Big[e^{-f(X_{n,i})}\Big] \cdot \mathbb{E}\Big[\prod_{j=i+1}^{m_n} e^{-f(X_{n,j})}\Big]\Big| =: \sum_{i=1}^{m_n-1} \varepsilon_i\,.\end{aligned}$$

Fix now an arbitrary $i \in \{1, \dots, m_n - 1\}$. After writing

$$\prod_{j=i+1}^{m_n} e^{-f(X_{n,j})} = \prod_{j\in\tilde{B}_n(i)} e^{-f(X_{n,j})} \prod_{j\in\tilde{B}_n^c(i)} e^{-f(X_{n,j})}\,,$$

one can easily check that

$$\begin{aligned}\varepsilon_i \leq &\Big|\mathbb{E}\Big[e^{-f(X_{n,i})} \cdot \Big(\prod_{j\in\tilde{B}_n(i)} e^{-f(X_{n,j})} - 1\Big) \prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big] \\ &- \mathbb{E}\Big[e^{-f(X_{n,i})}\Big] \cdot \mathbb{E}\Big[\Big(\prod_{j\in\tilde{B}_n(i)} e^{-f(X_{n,j})} - 1\Big) \prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big]\Big| \\ &+ \Big|\mathbb{E}\Big[e^{-f(X_{n,i})} \prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big] - \mathbb{E}\Big[e^{-f(X_{n,i})}\Big] \cdot \mathbb{E}\Big[\prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big]\Big| \,.\end{aligned}$$

Note that the first summand on the right hand side of the previous inequality equals

$$\begin{aligned}&\Big|\mathbb{E}\Big[\Big(e^{-f(X_{n,i})} - 1\Big) \cdot \Big(\prod_{j\in\tilde{B}_n(i)} e^{-f(X_{n,j})} - 1\Big) \prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big] \\ &- \mathbb{E}\Big[\Big(e^{-f(X_{n,i})} - 1\Big)\Big] \cdot \mathbb{E}\Big[\Big(\prod_{j\in\tilde{B}_n(i)} e^{-f(X_{n,j})} - 1\Big) \prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big]\Big| \,,\end{aligned}$$

and since $e^{-\sum_k f(x_k)} - 1 \neq 0$ implies that $f(x_k) > 0$, and hence $x_k \in K_m$, for at least one k, we obtain that

$$\begin{aligned}\varepsilon_i \leq\, &\mathbb{P}\Big(X_{n,i} \in K_m, \bigcup_{j\in\tilde{B}_n(i)} \{X_{n,j} \in K_m\}\Big) + \mathbb{P}\Big(X_{n,i} \in K_m\Big) \cdot \mathbb{P}\Big(\bigcup_{j\in\tilde{B}_n(i)} \{X_{n,j} \in K_m\}\Big) \\ &+ \Big|\mathbb{E}\Big[e^{-f(X_{n,i})} \prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big] - \mathbb{E}\Big[e^{-f(X_{n,i})}\Big] \cdot \mathbb{E}\Big[\prod_{j\in\tilde{B}^c_n(i)} e^{-f(X_{n,j})}\Big]\Big| \,.\end{aligned}$$

Hence,

$$\Big|\mathbb{E}\left[e^{-\sum_{i\in I_n} f(X_{n,i})}\right] - \prod_{i\in I_n} \mathbb{E}\left[e^{-f(X_{n,i})}\right]\Big| \leq \sum_{i=1}^{m_n-1} \varepsilon_i \leq b^m_{n,1} + b^m_{n,2} + b_{n,3}(f) \,.$$

□

Remark 1.4.6. Recall, $(X^*_{n,i} : i \in I_n)$ are independent random elements such that for all $i \in I_n$, $X^*_{n,i}$ is distributed as $X_{n,i}$. Further, let $(X^*_{n,i} : i \in I_n)$ and $(X_{n,i} : i \in I_n)$ be defined on the same probability space and independent. We can then bound $b_{n,3}(f)$ by

$$\begin{aligned}b_{n,3}(f) &\leq \sum_{i\in I_n} \mathbb{E}\Big|\mathbb{E}\Big[e^{-f(X_{n,i})} - e^{-f(X^*_{n,i})} \mid \sigma(X_{n,j} : j \in \tilde{B}^c_n(j))\Big]\Big| \\ &= \sum_{i\in I_n} \mathbb{E}\Big|\mathbb{E}\Big[e^{-f(X_{n,i})} \mid \sigma(X_{n,j} : j \in \tilde{B}^c_n(i))\Big] - \mathbb{E}\Big[e^{-f(X_{n,i})}\Big]\Big| \,.\end{aligned}$$

Since for any $f \in CB^+_b(\mathbb{X})$ the function $1 - e^{-f}$ is also an element $CB^+_b(\mathbb{X})$ and further bounded by 1, it follows that

$$\sum_{i\in I_n} \mathbb{E}\Big|\mathbb{E}\Big[f(X_{n,i}) \mid \sigma(X_{n,j} : j \in \tilde{B}^c_n(i))\Big] - \mathbb{E}\Big[f(X_{n,i})\Big]\Big| \to 0$$

for all $f \in CB_b^+(\mathbb{X})$ which are bounded by 1 implies that $b_{n,3}(f) \to 0$ for all $f \in CB_b^+(\mathbb{X})$.

Remark 1.4.7. The concept of neighborhoods implicitly appears already in Banys [Ban80, Theorem 4]. There, essentially the same sufficient conditions for convergence of N_n to a Poisson point process are given but with, in our notation, neighborhoods of the form $\tilde{B}_n(i) = \{i+1, \ldots, i+r_n\}$ and $\tilde{B}_n^c(i) = \{i+r_n+1, \ldots, m_n\}$ for all $i \in I_n$ where $(r_n)_{n\in\mathbb{N}}$ is a sequence of nonnegative integers. The proof is similar to ours and even though it is stated only for the case when $\mathbb{X}$ is locally compact, it transfers directly to the case of a general Polish space.

Remark 1.4.8. Similar results were also obtained by Schuhmacher [Sch05, Theorem 2.1], but with a completely different approach, using Stein's method. As a consequence, Schuhmacher even provides bounds on the convergence in the so–called Barbour-Brown distance d_2. However, this result does not directly imply our results, see [Sch05, Remark 2.4(b)] for the comparison to the result of Banys [Ban80] which is also relevant to our case.

Example 1.4.9. For Bernoulli random variables $X_{n,i}$ such that $\lim_{n\to\infty} \sup_{i\in I_n} \mathbb{P}(X_{n,i} = 1) = 0$ and $\lim_{n\to\infty} \sum_{i\in I_n} \mathbb{P}(X_{n,i} = 1) = \lambda \in (0,\infty)$, one can set $\mathbb{X}' = \{0,1\}$ and $\mathbb{X} = \{1\}$ with $\{1\}$ being the only non–empty bounded set in $\mathbb{X}$. Using Theorem 1.4.1 together with Proposition 1.4.5 and Remark 1.4.6, we recover the result of Arratia et al. [AGG89, Theorem 1] on convergence in distribution of $\sum_{i\in I_n} \mathbb{1}_{\{X_{n,i}=1\}}$ to a Poisson random variable with intensity λ, but without the bound on the distance in total variation.

Chapter 2

Regularly varying time series and random fields

2.1 Introduction

Over the years, a lot of progress has been made in the study of stationary time series under suitable regular variation and (weak) dependence assumptions. Intuitively, these assumptions imply that, over large time windows, observations exceeding a suitably chosen large threshold form a Poisson number of uniformly spaced i.i.d. clusters; after scaling, these clusters are distributed as a product of (i) a Pareto random variable representing the magnitude of the largest value in the cluster, and (ii) an independent random vector describing the serial dependence structure within the cluster; see also Janssen [Jan18, Introduction].

A way to make this formal is through the language of point processes. Let $(X_i)_{i\in\mathbb{Z}}$ be such a time series and a_n the $(1-1/n)$–quantile of the distribution of $|X_0|$. Building on the results of Mori [Mor77], Davis and Resnick [DR85], Davis and Hsing [DH95], Basrak and Segers [BS09], Basrak et al. [BKS12] and others, Basrak and Tafro [BT16] showed that

$$N_n = \sum_{i=1}^{n} \delta_{(i/n, X_i/a_n)} \xrightarrow{d} N = \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{Z}} \delta_{(T_i, P_i Q_j^i)} \,, \text{ as } n \to \infty \,, \tag{2.1}$$

where, in particular,

(i) $\sum_{i\in\mathbb{N}} \delta_{(T_i,P_i)}$ is a Poisson process on $[0,1]\times(0,\infty)$ which is homogeneous in time and with Pareto–like intensity in space;

(ii) $(Q_j^i)_{j\in\mathbb{Z}}$, $i \in \mathbb{N}$, are i.i.d. sequences of random variables (normalized *clusters*) independent of $\sum_{i\in\mathbb{N}} \delta_{(T_i,P_i)}$.

Thus, N is a Poisson cluster process. Note that this convergence takes place in the space of locally finite point measures on $[0,1]\times(\mathbb{R}\setminus\{0\})$ with sets of interest, i.e. bounded sets, being those which are bounded away from $[0,1] \times \{0\}$. In words, convergence in (2.1) controls only the asymptotic behavior of X_i's which exceed $a_n\epsilon$ in absolute value for fixed $\epsilon > 0$.

Once this so–called complete convergence result (or a variant of it including only the time or space component) is established, various (functional) limit theorems concerning (X_i) follow by applications of the continuous mapping theorem, see e.g. the references given above and also the books by Resnick [Res87, Res07] where the i.i.d. case is treated exhaustively and very intuitively.

Observe, however, that due to scaling of time to $[0,1]$ in (2.1), the information about the serial dependence of X_i's belonging to the same cluster is lost in the limit since they all collapse to the same time instance. More precisely, the limiting point process from (2.1) does not contain information on the order of Q_j^i's for fixed i. This information is important if e.g. one is interested in studying record times of (X_i).

The first goal of this chapter is to present a new type of point process convergence which will preserve the information about this order. In Chapter 3 this result is used to study sums and record times of stationary regularly varying time series.

The main idea is to break the dependent sample $(X_1, \dots, X_n)$ into smaller blocks whose size still tends to ∞, and to consider them as points of a point process on the (infinite–dimensional) space of real-valued double–sided sequences vanishing to zero in all directions. Besides preserving the order, the enlargement of the state space also transforms the problem of obtaining a Poisson cluster limit into obtaining a suitable Poisson approximation.

As in (2.1), the key role in our considerations is played by the so–called *tail process* introduced by [BS09]. A stationary $\mathbb{R}$–valued time series $(X_i)_{i\in\mathbb{Z}}$ has a tail process $(Y_i)_{i\in\mathbb{Z}}$ if for all $s \le t \in \mathbb{Z}$,

$$(u^{-1}X_s, \dots, u^{-1}X_t) \,\Big|\, |X_0| > u \xrightarrow{d} (Y_s, \dots, Y_t) \text{ in } \mathbb{R}^{t-s+1}, \text{ as } u \to \infty. \tag{2.2}$$

Here and in what follows, $A(u) \mid B(u) \xrightarrow{d} C$ as $u \to \infty$ for a family of random elements $A(u), C$ and events $B(u)$, $u > 0$, means that the law of $A(u)$ conditionally on $B(u)$ converges weakly as $u \to \infty$ to the law of C.

The second goal of this chapter is to extend the notion of the tail process and the corresponding point process convergence theory to $\mathbb{R}$–valued random fields indexed over the d–dimensional integer lattice. Using this framework, in Chapter 4 we revisit the classical problem of local alignment of (biological) sequences.

In what follows we consider only general random fields, i.e. we do not treat the time series case separately (except in some of the examples). In fact, extending the theory to higher dimensions yields some new insight on the existing results for time series, see in particular the concept of *anchoring* in Section 2.3.3.

Remark 2.1.1. Note, here we treat only the case of $\mathbb{R}$–valued time series and random fields. One can obtain multivariate analogues of all results by simply replacing the absolute value with an arbitrary norm. One reason for restricting to the univariate case is because all of our examples and applications are univariate. Secondly, in applications concerning multivariate data, instead of conditioning on the norm of the whole vector being large, it is often more natural to condition e.g. on the value of only one component being extreme.

The rest of the chapter is organized as follows. In Section 2.2 we extend the notion of the tail process to general stationary random fields and discuss some of its main properties. As in the time series case, it is not surprising that the existence of a tail process is equivalent to all finite–dimensional distributions of the field being multivariate regularly varying (hence the name *regularly varying* fields), see Theorem 2.2.1 (i). However, a nontrivial extension was to obtain sufficient conditions for existence of the tail process for general random fields, see Theorem 2.2.1 (iii).

In Section 2.3 we present a new type of point process convergence result for stationary regularly varying time series and random fields which preserves the order (or rather the shape) within the cluster in the limit, see Theorem 2.3.14. The main novelty is the order preserving infinite–dimensional space $\tilde{l}_0$ introduced in 2.3.1 which serves as a state space for point processes of blocks. The proof of Theorem 2.3.14 is then based on the Poisson approximation theory of Section 1.4. The key technical result, also of independent interest, is Proposition 2.3.10 which provides a way to check intensity measure convergence (1.12) of Corollary 1.4.4. It relies on the familiar *anti–clustering*[1] condition, see (2.18). Furthermore, the asymptotic $\mathcal{F}'$–independence condition applied to blocks, which is the remaining assumption of Corollary 1.4.4, provides a mixing condition which is also typical in this extremal context. Finally, we wish to emphasize the idea of *anchoring* described in Section 2.3.3 which has the intention of clarifying the link between the tail process and the components of the limiting Poisson (cluster) process of Theorem 2.3.14.

We finish the chapter by discussing techniques for verifying conditions of Theorem 2.3.14, or rather obtaining the desired point process convergence result. In particular, all of the assumptions are satisfied by m–dependent random fields, and in Section 2.4.1, the conclusion of Theorem 2.3.14 is extendend to fields which are, in a suitable sense, approximable by m–dependent ones (cf. [RS98]). Infinite order linear processes (or moving averages) constitute the most prominent examples of such fields. Further, we recall in Section 2.4.2 that, for stationary Markov chains, checking geometric ergodicity offers a practical way to verify the asymptotic independence condition. Also, note that the neighborhood approach described in Section 1.4.2 provides yet another way for checking the asymptotic independence condition, particularly useful for general random fields. This is illustrated in Chapter 4 in the study of the local alignment problem, see in particular

[1] Note that, even though catchy, this is not a good name for this condition since it does not eliminate the possibility of clustering.

Corollary 4.3.4.

This chapter is based on the papers [BPS18] and [BP18b].

2.2 The tail field

Consider a (strictly) stationary $\mathbb{R}$–valued random field $\boldsymbol{X} = (X_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ with $d \in \mathbb{N}$. For every finite and non–empty subset of indices $I \subseteq \mathbb{Z}^d$, denote by $\boldsymbol{X}_I$ the $\mathbb{R}^{|I|}$–valued random vector $(X_{\boldsymbol{i}} : \boldsymbol{i} \in I)$, i.e. $\boldsymbol{X}_I$'s represent finite–dimensional distributions of $\boldsymbol{X}$.

We say that a random field $\boldsymbol{Y} = (Y_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ is the *tail field* (or *tail process*) of $\boldsymbol{X}$, if for all finite and non–empty $I \subseteq \mathbb{Z}^d$,

$$u^{-1}\boldsymbol{X}_I \,\Big|\, |X_{\boldsymbol{0}}| > u \xrightarrow{d} \boldsymbol{Y}_I\,, \text{ as } u \to \infty\,, \tag{2.3}$$

where $\boldsymbol{0} = (0, \ldots, 0) \in \mathbb{Z}^d$. Note that we implicitly assume that $\mathbb{P}(|X_{\boldsymbol{0}}| > u) > 0$ for all $u > 0$. Observe, taking $I = \{\boldsymbol{0}\}$ in (2.3) yields that $\lim_{u\to\infty} \mathbb{P}(|X_{\boldsymbol{0}}| > uy)/\mathbb{P}(|X_{\boldsymbol{0}}| > u) = \mathbb{P}(|Y_{\boldsymbol{0}}| > y)$ for all except at most countably many $y \in [1, \infty)$. By standard arguments (see [BGT87, Theorem 1.4.1] and the discussion before it), this implies that $u \mapsto \mathbb{P}(|X_{\boldsymbol{0}}| > u)$ is a regularly varying function with index $-\alpha$ for some $\alpha > 0$, i.e.

$$\lim_{u\to\infty} \frac{\mathbb{P}(|X_{\boldsymbol{0}}| > uy)}{\mathbb{P}(|X_{\boldsymbol{0}}| > u)} = y^{-\alpha}\,,\; y > 0\,. \tag{2.4}$$

In particular, $\mathbb{P}(|Y_{\boldsymbol{0}}| > y) = y^{-\alpha}$ for all $y \geq 1$, i.e. $|Y_{\boldsymbol{0}}|$ is Pareto distributed with index α.

2.2.1 Existence of the tail field

A family of indices $\mathcal{I} \subseteq \mathbb{Z}^d$ is said to be *encompassing* if for every finite and non–empty $I \subseteq \mathbb{Z}^d$ there exists at least one $\boldsymbol{i}^* \in I$ such that $I - \boldsymbol{i}^* \subseteq \mathcal{I}$. Note that necessarily $\boldsymbol{0} \in \mathcal{I}$.

If $d = 1$, the set of nonnegative (or nonpositive) integers is an example of such family. More generally, assume that $\preceq$ is an arbitrary total order on $\mathbb{Z}^d$ which is translation–invariant in the sense that for all $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ in $\mathbb{Z}^d$, $\boldsymbol{i} \preceq \boldsymbol{j}$ implies $\boldsymbol{i} + \boldsymbol{k} \preceq \boldsymbol{j} + \boldsymbol{k}$. Then the set $\mathbb{Z}^d_{\succeq} = \{\boldsymbol{i} \in \mathbb{Z}^d : \boldsymbol{i} \succeq \boldsymbol{0}\}$ is clearly encompassing. Indeed, simply set $\boldsymbol{i}^* \in I$ to be the (unique) minimal element of the finite set I with respect to $\preceq$. We refer to such orders as *group orders* on $\mathbb{Z}^d$.

In particular, the lexicographic order on $\mathbb{Z}^d$, denoted by $\preceq_l$, is a group order. Recall, for indices $\boldsymbol{i} = (i_1, \ldots, i_d), \boldsymbol{j} = (j_1, \ldots, j_d) \in \mathbb{Z}^d$, $\boldsymbol{i} \prec_l \boldsymbol{j}$ if $i_k < j_k$ for the first k where i_k and j_k differ, and $\boldsymbol{i} \preceq_l \boldsymbol{j}$ if $\boldsymbol{i} \prec_l \boldsymbol{j}$ or $\boldsymbol{i} = \boldsymbol{j}$.

The following result extends [BS09, Theorem 2.1] which treats the case $d = 1$; the proof is postponed to Section 2.5.

Theorem 2.2.1. *For a stationary random field $\boldsymbol{X} = (X_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ and $\alpha > 0$, the following three statements are equivalent:*

(i) All finite–dimensional distributions of $\boldsymbol{X}$ are multivariate regularly varying with index α;

(ii) The field $\boldsymbol{X}$ has a tail field $\boldsymbol{Y} = (Y_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ with $\mathbb{P}(|Y_{\boldsymbol{0}}| \geq y) = y^{-\alpha}$ for $y \geq 1$.

(iii) There exists an encompassing $\mathcal{I} \subseteq \mathbb{Z}^d$ and a family of random variables $(Y_{\boldsymbol{i}} : \boldsymbol{i} \in \mathcal{I})$ with $\mathbb{P}(|Y_{\boldsymbol{0}}| \geq y) = y^{-\alpha}$ for $y \geq 1$, such that for all finite and non–empty $I \subseteq \mathcal{I}$,

$$u^{-1}\boldsymbol{X}_I \,\Big|\, |X_{\boldsymbol{0}}| > u \xrightarrow{d} (Y_{\boldsymbol{i}})_{\boldsymbol{i} \in I}\,, \text{ as } u \to \infty\,. \tag{2.5}$$

Recall that for finite $I \subseteq \mathbb{Z}^d$, $\boldsymbol{X}_I$ is multivariate regularly varying with index $\alpha > 0$ if for some norm $\|\cdot\|$ on $\mathbb{R}^{|I|}$ there exists a random vector on $\mathbb{R}^{|I|}$, say $\boldsymbol{\Theta}^{(I)}$, such that $\|\boldsymbol{\Theta}^{(I)}\| = 1$ and

$$\left(u^{-1}\|\boldsymbol{X}_I\|, \|\boldsymbol{X}_I\|^{-1}\boldsymbol{X}_I\right) \Big| \, \|\boldsymbol{X}_I\| > u \xrightarrow{d} (Y, \boldsymbol{\Theta}^{(I)})\,, \text{ as } u \to \infty\,,$$

where Y is independent of $\boldsymbol{\Theta}^{(I)}$ and satisfies $\mathbb{P}(Y > y) = y^{-\alpha}$ for $y \geq 1$.

The equivalence between (i) and (ii) (cf. (2.8) below) explains why fields admitting a tail process will simply be called *regularly varying*. We refer to the corresponding α as the *(tail) index* of the field.

Remark 2.2.2. While writing the thesis, we learned of a parallel study by Wu and Samorodnitsky [WS18] who also consider regularly varying fields with emphasis on various notions of extremal index in this context. They show by an example that for $d \geq 2$ existence of the limit of $u^{-1}\boldsymbol{X}_I \,\big|\, |X_{\boldsymbol{0}}| > u$ for all finite $I \subseteq \mathcal{I}$ when $\mathcal{I}$ is an orthant in $\mathbb{Z}^d$, is not sufficient for regular variation of $\boldsymbol{X}$ and hence existence of the tail field. This made us reconsider an earlier (incorrect) version of Theorem 2.2.1 which eventually led to a proper extension of [BS09, Theorem 2.1(ii)].

2.2.2 The spectral tail field

Consider now the space $\mathbb{R}^{\mathbb{Z}^d}$ equipped with the product topology and corresponding Borel σ–algebra. One can then rephrase (2.3) simply as

$$u^{-1}\boldsymbol{X} \,\Big|\, |X_{\boldsymbol{0}}| > u \xrightarrow{d} \boldsymbol{Y} \text{ in } \mathbb{R}^{\mathbb{Z}^d}\,, \tag{2.6}$$

see e.g. [Bil68, p. 19]. The *spectral tail field* $\boldsymbol{\Theta} = (\Theta_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ of $\boldsymbol{X}$ is defined by $\Theta_{\boldsymbol{i}} = Y_{\boldsymbol{i}}/|Y_{\boldsymbol{0}}|\,, \boldsymbol{i} \in \mathbb{Z}^d$. In particular, $|\Theta_{\boldsymbol{0}}| = 1$.

Proposition 2.2.3. *The spectral field* $\boldsymbol{\Theta}$ *is independent of* $|Y_0|$ *and satisfies*

$$|X_{\mathbf{0}}|^{-1}\boldsymbol{X} \;\Big|\; |X_{\mathbf{0}}| > u \xrightarrow{d} \boldsymbol{\Theta} \;\; in\, \mathbb{R}^{\mathbb{Z}^d} . \tag{2.7}$$

Proof. The continuous mapping theorem (see e.g. [Bil68, Corollary 1, p. 31]) applied to (2.6) implies that

$$(u^{-1}|X_{\mathbf{0}}|, |X_{\mathbf{0}}|^{-1}\boldsymbol{X}) \;\Big|\; |X_{\mathbf{0}}| > u \xrightarrow{d} (|Y_{\mathbf{0}}|, \boldsymbol{\Theta}) \;\text{ in } [1,\infty) \times \mathbb{R}^{\mathbb{Z}^d} . \tag{2.8}$$

In particular, (2.7) holds and moreover, using (2.4) yields that

$$\begin{aligned}
\mathbb{P}(|Y_{\mathbf{0}}| > y, \boldsymbol{\Theta} \in B) &= \lim_{u\to\infty} \mathbb{P}(|X_{\mathbf{0}}| > uy, \boldsymbol{X}/|X_{\mathbf{0}}| \in B \mid |X_{\mathbf{0}}| > u) \\
&= \lim_{u\to\infty} \frac{\mathbb{P}(|X_{\mathbf{0}}| > uy)}{\mathbb{P}(|X_{\mathbf{0}}| > u)} \cdot \mathbb{P}(\boldsymbol{X}/|X_{\mathbf{0}}| \in B \mid |X_{\mathbf{0}}| > uy) \\
&= y^{-\alpha}\mathbb{P}(\boldsymbol{\Theta} \in B) = \mathbb{P}(|Y_{\mathbf{0}}| > y)\mathbb{P}(\boldsymbol{\Theta} \in B) ,
\end{aligned} \tag{2.9}$$

for all $y > 0$ and all measurable $B \subseteq \mathbb{R}^{\mathbb{Z}^d}$ such that $\mathbb{P}(\boldsymbol{\Theta} \in \partial B) = 0$. Denote by $\mathcal{S}$ the family of all such B's. Note that $\mathcal{S}$ is closed under finite intersections and that every open set in $\mathbb{R}^{\mathbb{Z}^d}$ can be represented as a countable union of elements in $\mathcal{S}$ (fix a metric and use open balls). In particular, $\mathcal{S}$ generates the Borel σ–algebra on $\mathbb{R}^{\mathbb{Z}^d}$. Thus, (2.9) implies that $\boldsymbol{\Theta}$ and $|Y_{\mathbf{0}}|$ are independent. □

Even though the tail field is typically not stationary, regular variation and stationarity of the underlying random field $\boldsymbol{X}$ yield specific distributional properties of $\boldsymbol{\Theta}$ (and hence of $\boldsymbol{Y}$) summarized by the so–called time–change formula: for every integrable (in the sense that one of the expectations below exists) or nonnegative measurable function $h : \mathbb{R}^{\mathbb{Z}^d} \to \mathbb{R}$ and all $\boldsymbol{j} \in \mathbb{Z}^d$,

$$\mathbb{E}\left[h\left((\Theta_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}\right)\mathbb{1}\{\Theta_{-\boldsymbol{j}} \neq 0\}\right] = \mathbb{E}\left[h\left((\Theta_{\boldsymbol{i}+\boldsymbol{j}}/|\Theta_{\boldsymbol{j}}|)_{\boldsymbol{i}\in\mathbb{Z}^d}\right)|\Theta_{\boldsymbol{j}}|^{\alpha}\mathbb{1}\{\Theta_{\boldsymbol{j}} \neq 0\}\right] . \tag{2.10}$$

In the case of time series, (2.10) appears in [BS09] and the proof is easily extended to the case of random fields, see [WS18, Theorem 3.2]. Alternatively, one can arrive at (2.10) following the approach of [PS18] who use the so–called tail measure of $\boldsymbol{X}$ introduced in [SO12], see also [DHS17].

Remark 2.2.4. Let $\boldsymbol{X}$ be a stationary random field and $\alpha > 0$. If $\lim_{u\to\infty}\mathbb{P}(|X_{\mathbf{0}}| > uy)/\mathbb{P}(|X_{\mathbf{0}}| > u) \to y^{-\alpha}$ for all $y > 0$ and for some encompassing $\mathcal{I} \subseteq \mathbb{Z}^d$ there exist random variables $(\Theta_{\boldsymbol{i}} : \boldsymbol{i} \in \mathcal{I})$ such that for all finite and non–empty $I \subseteq \mathcal{I}$, $|X_{\mathbf{0}}|^{-1}\boldsymbol{X}_I \;\Big|\; |X_{\mathbf{0}}| > u \xrightarrow{d} (\Theta_i)_{i\in I}$, then $\boldsymbol{X}$ is regularly varying with index α; combine the proof of [BS09, Corollary 3.2] and Theorem 2.2.1.

If $\mathcal{I} \neq \mathbb{Z}^d$, the distribution of the whole spectral process $\boldsymbol{\Theta}$ is then determined by (2.10) and the tail field of $\boldsymbol{X}$ is given by $\boldsymbol{Y} = Y\boldsymbol{\Theta}$ where Y is independent of $\boldsymbol{\Theta}$ and satisfies

$\mathbb{P}(Y \geq y) = y^{-\alpha}$ for $y \geq 1$.

2.3 Point process convergence

Denote by $\leq$ the component–wise order on $\mathbb{Z}^d$, thus for $\boldsymbol{i} = (i_1, \dots, i_d), \boldsymbol{j} = (j_1, \dots, j_d) \in \mathbb{Z}^d$, $\boldsymbol{i} \leq \boldsymbol{j}$ if $i_k \leq j_k$ for all $k = 1, \dots, d$. Take a sequence of positive integers (r_n) such that $\lim_{n\to\infty} r_n = \lim_{n\to\infty} n/r_n = \infty$ and let $k_n = \lfloor n/r_n \rfloor$. For each $n \in \mathbb{N}$, decompose $\{1, \dots, n\}^d$ into blocks $J_{n,\boldsymbol{i}}$, $\boldsymbol{i} \in I_n := \{1, \dots, k_n\}^d$, of size r_n^d by

$$J_{n,\boldsymbol{i}} = (\boldsymbol{j} \in \mathbb{Z}^d : (\boldsymbol{i} - \boldsymbol{1}) \cdot r_n + \boldsymbol{1} \leq \boldsymbol{j} \leq \boldsymbol{i} \cdot r_n) \,. \tag{2.11}$$

The goal of this section is to study convergence in the distribution of point processes based on (increasing) blocks

$$\boldsymbol{X}_{n,\boldsymbol{i}} := \boldsymbol{X}_{J_{n,\boldsymbol{i}}}, \; \boldsymbol{i} \in I_n \,, \tag{2.12}$$

see (2.15) below. We start by introducing a suitable space for the $\boldsymbol{X}_{n,\boldsymbol{i}}$'s.

2.3.1 A space for blocks – $\tilde{l}_0$

Let l_0 be the space of all $\mathbb{R}$-valued arrays on $\mathbb{Z}^d$ converging to zero in all directions, i.e. $l_0 = \{(x_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d} : \lim_{|\boldsymbol{i}|\to\infty} |x_i| = 0\}$, where $|\boldsymbol{i}| = \max_{k=1,\dots,d} |i_k|$ for $\boldsymbol{i} = (i_1, \dots, i_d) \in \mathbb{Z}^d$. On l_0 consider the uniform norm

$$\|\boldsymbol{x}\|_\infty = \sup_{\boldsymbol{i}\in\mathbb{Z}^d} |x_{\boldsymbol{i}}| \,, \; \boldsymbol{x} = (x_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d} \,,$$

which makes l_0 into a separable Banach space. Indeed, l_0 is the closure of all rational–valued arrays on $\mathbb{Z}^d$ with at most finitely many non–zero terms in the Banach space of all bounded $\mathbb{R}$–valued arrays on $\mathbb{Z}^d$.

Define the family of shift operators $B^{\boldsymbol{k}}$, $\boldsymbol{k} \in \mathbb{Z}^d$, on $\mathbb{R}^{\mathbb{Z}^d}$ by $B^{\boldsymbol{k}}(x_{\boldsymbol{i}})_{\boldsymbol{i}} = (x_{\boldsymbol{i}+\boldsymbol{k}})_{\boldsymbol{i}}$ and introduce an equivalence relation $\sim$ on l_0 by letting $\boldsymbol{x} \sim \boldsymbol{y}$ for $\boldsymbol{x}, \boldsymbol{y} \in l_0$ if $\boldsymbol{y} = B^{\boldsymbol{k}}\boldsymbol{x}$ for some $\boldsymbol{k} \in \mathbb{Z}^d$. In the sequel, we consider the quotient space $\tilde{l}_0 = l_0/\sim$ of shift–equivalent arrays. Observe, for $\tilde{\boldsymbol{x}} \in \tilde{l}_0$ and an arbitrary $\tilde{\boldsymbol{x}} \in \tilde{\boldsymbol{x}}$, $\tilde{\boldsymbol{x}} = \{B^{\boldsymbol{k}}\boldsymbol{x} : \boldsymbol{k} \in \mathbb{Z}^d\}$. Further, define a function $\tilde{d} : \tilde{l}_0 \times \tilde{l}_0 \longrightarrow [0, \infty)$ by

$$\tilde{d}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = \inf\{\|\boldsymbol{x} - \boldsymbol{y}\|_\infty : \boldsymbol{x} \in \tilde{\boldsymbol{x}}, \boldsymbol{y} \in \tilde{\boldsymbol{y}}\} \,, \; \tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \in \tilde{l}_0 \,. \tag{2.13}$$

Since $\|B^{\boldsymbol{k}}\boldsymbol{x} - B^{\boldsymbol{k}'}\boldsymbol{y}\|_\infty = \|B^{\boldsymbol{k}-\boldsymbol{k}'}\boldsymbol{x} - \boldsymbol{y}\|_\infty$ for all $\boldsymbol{x}, \boldsymbol{y} \in l_0$ and $\boldsymbol{k}, \boldsymbol{k}' \in \mathbb{Z}^d$, it follows that for all $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \in \tilde{l}_0$ and arbitrary $\boldsymbol{y} \in \tilde{\boldsymbol{y}}$,

$$\tilde{d}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = \inf_{\boldsymbol{x}\in\tilde{\boldsymbol{x}}} \|\boldsymbol{x} - \boldsymbol{y}\|_\infty \,. \tag{2.14}$$

This fact leads to the following technical result the proof of which is postponed to Section 2.5.

Lemma 2.3.1. *The function $\tilde{d}$ is a metric on $\tilde{l}_0$ and moreover, $(\tilde{l}_0, \tilde{d})$ is a separable and complete metric space.*

Observe that for $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_2, \dots \in \tilde{l}_0$, $\tilde{d}(\tilde{\boldsymbol{x}}_n, \tilde{\boldsymbol{x}}) \to 0$ as $n \to \infty$ if and only if for some, and then for every, $\boldsymbol{x} \in \tilde{\boldsymbol{x}}$ there exists $\boldsymbol{x}_n \in \tilde{\boldsymbol{x}}_n$, $n \in \mathbb{N}$, such that $\|\boldsymbol{x}_n - \boldsymbol{x}\|_\infty \to 0$. Further, for any finite $I \subseteq \mathbb{Z}^d$ one can naturally consider finite arrays $\boldsymbol{x} \in \mathbb{R}^I$ as elements of $\tilde{l}_0$ by simply adding infinitely many zeros around $\boldsymbol{x}$ and then mapping this element of l_0 into its equivalence class.

In what follows, on l_0 and $\tilde{l}_0$ consider their respective Borel σ–algebras $\mathcal{B}(l_0)$ and $\mathcal{B}(\tilde{l}_0)$. Note that $\mathcal{B}(l_0)$ coincides with trace σ–algebra of l_0 in $\mathbb{R}^{\mathbb{Z}^d}$ considered with respect to its cylindrical σ-algebra. Further, call a set $B \subseteq l_0$ *shift–invariant* if $\boldsymbol{x} \in B$ implies that $B^{\boldsymbol{k}}\boldsymbol{x} \in B$ for all $\boldsymbol{k} \in \mathbb{Z}^d$. Also, a function h on l_0 is *shift–invariant* if $h(B^{\boldsymbol{k}}\boldsymbol{x}) = h(\boldsymbol{x})$ for all $\boldsymbol{x} \in l_0, \boldsymbol{k} \in \mathbb{Z}^d$.

Since topologies of l_0 and $\tilde{l}_0$ are Polish and the corresponding quotient map $\pi : l_0 \to \tilde{l}_0$ is continuous, [Ber88, Corollary A.2.5] implies that for every $\tilde{B} \subseteq \tilde{l}_0$, $\tilde{B} \in \mathcal{B}(\tilde{l}_0)$ if and only if $\pi^{-1}(\tilde{B}) \in \mathcal{B}(l_0)$. In other words, $\mathcal{B}(\tilde{l}_0)$ coincides with the family of sets $\pi(B)$, for $B \in \mathcal{B}(l_0)$ which are shift–invariant. Moreover, a function $\tilde{h}$ on $\tilde{l}_0$ is measurable if and only if $\tilde{h} = h \circ \pi$ for some shift–invariant measurable function h on l_0.

2.3.2 Point process of blocks

Consider now the space $\tilde{l}_{0,0} := \tilde{l}_0 \setminus \{\boldsymbol{0}\}$ with bounded sets being those which are bounded away from $\boldsymbol{0}$ w.r.t. $\tilde{d}$, see Example 1.2.5. Define point processes of blocks

$$N_n' = \sum_{\boldsymbol{i} \in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\,,\; n \in \mathbb{N}\,, \tag{2.15}$$

in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$, where the sequence (a_n) is chosen such that

$$\lim_{n\to\infty} n^d \mathbb{P}(|X_{\boldsymbol{0}}| > a_n) = 1\,. \tag{2.16}$$

To obtain convergence of N_n' one can apply Corollary 1.4.4. For each $n \in \mathbb{N}$, denote $J_{r_n} := \{1, \dots, r_n\}^d = J_{n,\boldsymbol{1}}$ and let $\boldsymbol{X}_{r_n} := \boldsymbol{X}_{J_{r_n}} = \boldsymbol{X}_{n,\boldsymbol{1}}$ represent the common distribution of blocks $\boldsymbol{X}_{n,\boldsymbol{i}}$, $\boldsymbol{i} \in I_n$. Under this notation, condition (1.12) reduces to existence of a measure ν in $\mathcal{M}(\tilde{l}_{0,0})$ such that

$$k_n^d \mathbb{P}(a_n^{-1}\boldsymbol{X}_{r_n} \in \cdot) \xrightarrow{v} \nu\,. \tag{2.17}$$

Assumption 2.3.2. *There exists a sequence of positive integers* $(r_n)_n$ *such that* $\lim_{n\to\infty} r_n = \lim_{n\to\infty} n/r_n = \infty$ *and for every* $u > 0$*,*

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{m<|\boldsymbol{i}|\le r_n} |X_{\boldsymbol{i}}| > a_n u \;\middle|\; |X_{\boldsymbol{0}}| > a_n u \right) = 0\,. \tag{2.18}$$

As we show in Proposition 2.3.10 below, for a sequence (r_n) satisfying (2.18), convergence in (2.17) holds with the limiting measure ν of the form

$$\nu(\,\cdot\,) = \vartheta \int_0^\infty \mathbb{P}(y\boldsymbol{Q} \in \cdot)\alpha y^{-\alpha-1} dy\,,$$

for some $\vartheta \in (0,1]$ and $\boldsymbol{Q}$ being a random element in $\tilde{l}_0$ satisfying $\|\boldsymbol{Q}\|_\infty = 1$ almost surely. In the following we first describe ϑ and $\boldsymbol{Q}$ in terms of the tail field of $\boldsymbol{X}$.

2.3.3 Anchoring the tail process – ϑ and $\boldsymbol{Q}$

Let $\boldsymbol{Y} = (Y_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}$ be the tail field of $\boldsymbol{X} = (X_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}$. It follows easily (cf. [BS09, Proposition 4.2]) that Assumption 2.3.2 implies that $\mathbb{P}(\lim_{|\boldsymbol{i}|\to\infty} |Y_{\boldsymbol{i}}| = 0) = 1$, i.e. that $\mathbb{P}(\boldsymbol{Y} \in l_0) = 1$. Recall that $|Y_{\boldsymbol{0}}| > 1$ so in particular $\|\boldsymbol{Y}\|_\infty > 1$.

We say that a measurable function $A : \{\boldsymbol{x} \in l_0 : \|\boldsymbol{x}\|_\infty > 1\} \to \mathbb{Z}^d$ is an *anchoring function* if

(i) $A((x_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}) = \boldsymbol{j}$ for some $\boldsymbol{j} \in \mathbb{Z}^d$ implies that $|x_{\boldsymbol{j}}| > 1$;

(ii) For each $\boldsymbol{j} \in \mathbb{Z}^d$, $A((x_{\boldsymbol{i}-\boldsymbol{j}})_{\boldsymbol{i}}) = A((x_{\boldsymbol{i}})_{\boldsymbol{i}}) + \boldsymbol{j}$.

In words, A picks one of finitely many $x_{\boldsymbol{i}}$'s which are larger than one in absolute value in a way which is insensitive to translations. Observe, for an arbitrary group order on $\mathbb{Z}^d$, the following are examples of anchoring functions.

– *first exceedance*: $A^{fe}((x_{\boldsymbol{i}})_{\boldsymbol{i}}) = \min\{\boldsymbol{j} \in \mathbb{Z}^d : |x_{\boldsymbol{j}}| > 1\}$,

– *last exceedance*: $A^{le}((x_{\boldsymbol{i}})_{\boldsymbol{i}}) = \max\{\boldsymbol{j} \in \mathbb{Z}^d : |x_{\boldsymbol{j}}| > 1\}$,

– *first maximum*: $A^{fm}((x_{\boldsymbol{i}})_{\boldsymbol{i}}) = \min\{\boldsymbol{j} \in \mathbb{Z}^d : |x_{\boldsymbol{j}}| = \|(x_{\boldsymbol{i}})_{\boldsymbol{i}}\|\}$.

We will exploit the following property of the tail field which is implied solely by stationarity of $\boldsymbol{X}$.

Lemma 2.3.3. *For every bounded measurable function* $h : \mathbb{R}^{\mathbb{Z}^d} \to \mathbb{R}$ *and all* $\boldsymbol{j} \in \mathbb{Z}^d$*,*

$$\mathbb{E}\left[h\left((Y_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}\right) \mathbb{1}\{|Y_{\boldsymbol{j}}| > 1\}\right] = \mathbb{E}\left[h\left((Y_{\boldsymbol{i}-\boldsymbol{j}})_{\boldsymbol{i}\in\mathbb{Z}^d}\right) \mathbb{1}\{|Y_{-\boldsymbol{j}}| > 1\}\right]\,. \tag{2.19}$$

Proof. Assume in addition that h is continuous with respect to the product topology on $\mathbb{R}^{\mathbb{Z}^d}$. Then, since $\mathbb{P}(Y_{\boldsymbol{j}} = 1) = \mathbb{P}(|Y_{\mathbf{0}}| \cdot |\Theta_{\boldsymbol{j}}| = 1) = 0$ for all $\boldsymbol{j} \in \mathbb{Z}^d$, the definition of the tail process and stationarity of $(X_{\boldsymbol{i}})$ imply

$$\begin{aligned}
\mathbb{E}\left[h\left((Y_{\boldsymbol{i}})_{\boldsymbol{i}}\right)\mathbb{1}\{|Y_{\boldsymbol{j}}|>1\}\right] &= \lim_{u\to\infty}\mathbb{E}\left[h\left((u^{-1}X_{\boldsymbol{i}})_{\boldsymbol{i}}\right)\mathbb{1}\{|X_{\boldsymbol{j}}|>u\}\mid |X_{\mathbf{0}}|>u\right] \\
&= \lim_{u\to\infty}\frac{\mathbb{E}\left[h\left((u^{-1}X_{\boldsymbol{i}})_{\boldsymbol{i}}\right)\mathbb{1}\{|X_{\boldsymbol{j}}|>u, |X_{\mathbf{0}}|>u\}\right]}{\mathbb{P}(|X_{\mathbf{0}}|>u)} \\
&= \lim_{u\to\infty}\frac{\mathbb{E}\left[h\left((u^{-1}X_{\boldsymbol{i}-\boldsymbol{j}})_{\boldsymbol{i}}\right)\mathbb{1}\{|X_{\mathbf{0}}|>u, |X_{-\boldsymbol{j}}|>u\}\right]}{\mathbb{P}(|X_{\mathbf{0}}|>u)} \\
&= \mathbb{E}\left[h\left((Y_{\boldsymbol{i}-\boldsymbol{j}})_{\boldsymbol{i}}\right)\mathbb{1}\{|Y_{-\boldsymbol{j}}|>1\}\right].
\end{aligned}$$

Since finite Borel measures on a metric space are determined by integrals of continuous and bounded functions, this yields (2.19). □

Remark 2.3.4. Using the already mentioned tail measure of $\boldsymbol{X}$, one can give a one–line proof of the previous result, see [PS18, Lemma 2.2].

Lemma 2.3.5. *Assume that $\mathbb{P}(\boldsymbol{Y} \in l_0) = 1$. Then for every anchoring function A*

$$\mathbb{P}(A(\boldsymbol{Y}) = \mathbf{0}) > 0\,.$$

Proof. Assume that $\mathbb{P}(A(\boldsymbol{Y}) = \mathbf{0}) = 0$. Applying (2.19) yields

$$\begin{aligned}
1 &= \sum_{\boldsymbol{j}\in\mathbb{Z}^d}\mathbb{P}(A(\boldsymbol{Y}) = \boldsymbol{j}) = \sum_{\boldsymbol{j}\in\mathbb{Z}^d}\mathbb{P}(A(\boldsymbol{Y}) = \boldsymbol{j}, |Y_{\boldsymbol{j}}|>1) \\
&= \sum_{\boldsymbol{j}\in\mathbb{Z}^d}\mathbb{P}(A((Y_{\boldsymbol{i}-\boldsymbol{j}})_{\boldsymbol{i}}) = \boldsymbol{j}, |Y_{-\boldsymbol{j}}|>1) = \sum_{\boldsymbol{j}\in\mathbb{Z}^d}\mathbb{P}(A(\boldsymbol{Y}) = \mathbf{0}, |Y_{-\boldsymbol{j}}|>1) = 0\,.
\end{aligned}$$

Hence, $\mathbb{P}(A(\boldsymbol{Y}) = \mathbf{0}) > 0$. □

Whenever $\mathbb{P}(A(\boldsymbol{Y}) = \mathbf{0}) > 0$, one can consider the *anchored tail process* $\boldsymbol{Z}^A = (Z^A_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ which has the same distribution as $\boldsymbol{Y} = (Y_{\boldsymbol{i}})_{\boldsymbol{i}}$ but conditionally on $A(\boldsymbol{Y}) = 0$. Also, define $\boldsymbol{Q}^A = (Q^A_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ by $Q^A_{\boldsymbol{i}} = Z^A_{\boldsymbol{i}}/\|\boldsymbol{Z}^A\|_\infty$.

Lemma 2.3.6. *Assume that $\mathbb{P}(\boldsymbol{Y} \in l_0) = 1$ and let A, A' be two anchoring functions. Then*

$$\mathbb{P}(A(\boldsymbol{Y}) = \mathbf{0}) = \mathbb{P}(A'(\boldsymbol{Y}) = \mathbf{0})\,,$$

and for every measurable and bounded function $h : l_0 \to \mathbb{R}$ which is shift–invariant,

$$\mathbb{E}[h(\boldsymbol{Z}^A)] = \mathbb{E}[h(\boldsymbol{Z}^{A'})]\,.$$

Proof. Using (2.19) and shift–invariance of h we obtain

$$\begin{aligned}\mathbb{E}[h(\boldsymbol{Y})\mathbb{1}\{A(\boldsymbol{Y})=\mathbf{0}\}] &= \sum_{\boldsymbol{j}\in\mathbb{Z}^d}\mathbb{E}[h(\boldsymbol{Y})\mathbb{1}\{A(\boldsymbol{Y})=\mathbf{0},A'(\boldsymbol{Y})=\boldsymbol{j},|Y_{\boldsymbol{j}}|>1\}]\\ &= \sum_{\boldsymbol{j}\in\mathbb{Z}^d}\mathbb{E}[h(\boldsymbol{Y})\mathbb{1}\{A(\boldsymbol{Y})=-\boldsymbol{j},A'(\boldsymbol{Y})=\mathbf{0}\}]\\ &= \mathbb{E}[h(\boldsymbol{Y})\mathbb{1}\{A'(\boldsymbol{Y})=\mathbf{0}\}]\,.\end{aligned}$$

Taking $h\equiv 1$ yields the first statement, and then the second one follows immediately. □

If $\mathbb{P}(\boldsymbol{Y}\in l_0)=1$, denote by ϑ the common value of $\mathbb{P}(A(\boldsymbol{Y})=\mathbf{0})$, i.e. for an arbitrary anchoring function A set

$$\vartheta=\mathbb{P}(A(\boldsymbol{Y})=\mathbf{0})\,. \tag{2.20}$$

In particular, for any group order $\preceq$ on $\mathbb{Z}^d$, using the first/last exceedance as anchor yields,

$$\vartheta=\mathbb{P}(\sup_{\boldsymbol{j}\prec\mathbf{0}}|Y_{\boldsymbol{j}}|\le 1)=\mathbb{P}(\sup_{\boldsymbol{j}\succ\mathbf{0}}|Y_{\boldsymbol{j}}|\le 1)\,. \tag{2.21}$$

Also,

$$\vartheta=\mathbb{P}(A^{fm}(\boldsymbol{Y})=0)=\mathbb{P}(A^{fm}(\boldsymbol{\Theta})=0)\,. \tag{2.22}$$

Observe here that the function A^{fm} remains well defined on the whole set l_0 without $\mathbf{0}$. As it turns out, under suitable dependence conditions, ϑ represents the extremal index of the field $(|X_{\boldsymbol{j}}|)_{\boldsymbol{j}}$, see Remark 2.3.16 below (cf. also Remark 2.3.12).

Furthermore, the second part of the previous result implies that the distribution of $\boldsymbol{Z}^A$ (and hence of $\boldsymbol{Q}^A$), when viewed as an element in $\tilde{l}_0$, does not depend on the anchoring function A; see the end of Section 2.3.1. Therefore, we can denote by $\boldsymbol{Z}$ and $\boldsymbol{Q}$ random elements in $\tilde{l}_0$ so that

$$\boldsymbol{Z}\stackrel{d}{=}\boldsymbol{Z}^A \quad\text{and}\quad \boldsymbol{Q}\stackrel{d}{=}\boldsymbol{Q}^A\,, \tag{2.23}$$

for an arbitrary anchoring function A. Note, conditionally on $A^{fm}(\boldsymbol{Y})=\mathbf{0}$ (or equivalently on $A^{fm}(\boldsymbol{\Theta})=\mathbf{0}$), $\|\boldsymbol{Y}\|_\infty$ is equal to $|Y_{\mathbf{0}}|$ and therefore $\boldsymbol{Y}/\|\boldsymbol{Y}\|_\infty=\boldsymbol{\Theta}$. Consequently,

- $\mathbb{P}(\|\boldsymbol{Z}\|_\infty\ge y)=y^{-\alpha}$ for all $y\ge 1$,
- $\|\boldsymbol{Z}\|_\infty$ and $\boldsymbol{Q}$ are independent.

Moreover,

$$\boldsymbol{Q}\stackrel{d}{=}\boldsymbol{\Theta}\,\Big|\,A^{fm}(\boldsymbol{\Theta})=\mathbf{0}\quad\text{in }\tilde{l}_0\,. \tag{2.24}$$

As discussed in Remark 2.3.12 below, under Assumption 2.3.2, $\boldsymbol{Z}$ (i.e $\boldsymbol{Z}^A$ for any A) describes the asymptotic distribution of a cluster of extremes of $\boldsymbol{X}$, i.e. (normalized) block $\boldsymbol{X}_{r_n}$ conditioned on having at least one extreme observation.

Before proving the intensity convergence (2.17) we discuss couple of examples.

2.3.4 Examples

Below we present two classes of regularly varying models which frequently arise in applications. As discussed in Remark 2.3.9, these classes represent two essentially different mechanisms which induce the regularly varying structure.

Note that the second example concerns only the time series case. A random field version of the corresponding model arises in the local sequence alignment problem studied in Chapter 4 below.

Example 2.3.7 (**Moving averages**). Let $(\xi_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ be i.i.d. random variables with regularly varying distribution with index $\alpha > 0$, i.e.

$$\lim_{u\to\infty} \frac{\mathbb{P}(|\xi_{\boldsymbol{0}}| > uy)}{\mathbb{P}(|\xi_{\boldsymbol{0}}| > u)} = y^{-\alpha}\,,\; y > 0\,,$$

and for some $p \in [0, 1]$,

$$\lim_{u\to\infty} \mathbb{P}(\xi_{\boldsymbol{0}} > 0 \mid |\xi_{\boldsymbol{0}}| > u) = p\,,\; \lim_{u\to\infty} \mathbb{P}(\xi_{\boldsymbol{0}} < 0 \mid |\xi_{\boldsymbol{0}}| > u) = 1 - p\,.$$

Consider the infinite order moving average process $\boldsymbol{X} = (X_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ defined by

$$X_{\boldsymbol{i}} = \sum_{\boldsymbol{j}\in\mathbb{Z}^d} c_{\boldsymbol{j}}\xi_{\boldsymbol{i}-\boldsymbol{j}}\,, \tag{2.25}$$

where $(c_{\boldsymbol{j}} : \boldsymbol{j} \in \mathbb{Z}^d)$ is a field of real numbers satisfying

$$0 < \sum_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^{\delta} < \infty\,, \tag{2.26}$$

for some $\delta > 0$ such that $\delta < \alpha$ and $\delta \le 1$. It is easily shown (see e.g. [Res87, Section 4.5]) that this condition ensures that the series above is absolutely convergent. Note also that $\sum_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^{\alpha} < \infty$. Furthermore, it can be proved as in [Res87, Lemma 4.24] that

$$\lim_{u\to\infty} \frac{\mathbb{P}(|X_{\boldsymbol{0}}| > u)}{\mathbb{P}(|\xi_{\boldsymbol{0}}| > u)} = \sum_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^{\alpha}\,. \tag{2.27}$$

Moreover, extending the arguments of Meinguet and Segers [MS10, Example 9.2], one can show that the stationary field $\boldsymbol{X}$ is jointly regularly varying with index α and spectral tail field given by

$$(\Theta_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d} \stackrel{d}{=} (Kc_{\boldsymbol{i}+J}/|c_J|)_{\boldsymbol{i}\in\mathbb{Z}^d} \tag{2.28}$$

where K is a $\{-1,1\}$–valued random variable with $\mathbb{P}(K=1)=p$, and J an $\mathbb{Z}^d$–valued random variable, independent of K, such that $\mathbb{P}(J=\boldsymbol{j}) = |c_{\boldsymbol{j}}|^\alpha / \sum_{\boldsymbol{i}\in\mathbb{Z}^d} |c_{\boldsymbol{i}}|^\alpha$ for all $\boldsymbol{j}\in\mathbb{Z}^d$.

In particular, $\mathbb{P}(\boldsymbol{\Theta}\in l_0) = \mathbb{P}(\boldsymbol{Y}\in l_0) = 1$. Choosing A^{fm} as the anchoring function (see (2.22) and (2.24)) yields that

$$\vartheta = \frac{\max_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^\alpha}{\sum_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^\alpha}\,, \quad \boldsymbol{Q} \stackrel{d}{=} \left(\frac{Kc_{\boldsymbol{j}}}{\max_{\boldsymbol{i}\in\mathbb{Z}^d}|c_{\boldsymbol{i}}|}\right)_{\boldsymbol{j}\in\mathbb{Z}^d} \text{ in } \tilde{l}_0\,. \tag{2.29}$$

Example 2.3.8 (**Solutions to stochastic recurrence equations**). Assume that a (non-negative) time series $(X_i)_{i\in\mathbb{Z}}$ is the stationary solution to the stochastic recurrence equation

$$X_i = A_i X_{i-1} + B_i\,,\ i\in\mathbb{Z}\,, \tag{2.30}$$

where $((A_i,B_i))_{i\in\mathbb{Z}}$ are i.i.d. $[0,\infty)^2$–random vectors such that for some $\alpha>0$ conditions of [BDM16, Theorem 2.4.4] hold, in particular $\mathbb{E}[A_0^\alpha]=1$, $\mathbb{E}[\log A_0]<0$ and $\mathbb{E}[B_0^\alpha]<\infty$. Such a solution exists and moreover the marginal distribution of (X_i) is regularly varying with index α. By [Seg07, Theorem 2.3], the *forward* spectral tail process $(\Theta_i : i\geq 0)$ exists (hence, (X_i) is jointly regularly varying) with distribution given by $\Theta_0=1$ and

$$\Theta_i = \prod_{k=1}^{i} A_k\,,\ i\geq 1\,. \tag{2.31}$$

With the aid of the time–change formula (2.10) one can verify that the *backward* spectral tail process $(\Theta_i : i\leq 0)$ also has a multiplicative structure

$$\Theta_{-i} = \prod_{k=1}^{i} \frac{1}{A_k^*}\,,\ i\geq 1\,, \tag{2.32}$$

where $(A_k^*)_{k\in\mathbb{N}}$ are i.i.d. (positive) random variables independent of $(A_k)_{k\in\mathbb{N}}$ and such that $\mathbb{P}(A_1^*\in\cdot\,) = \mathbb{E}[\mathbb{1}_{\{A_1\in\cdot\,\}}A_1^\alpha]$; see [Seg07, Theorem 5.2 and equations (4.3)–(4.5)] for details and also the proof of Proposition 4.2.1 below. The *tilted* random variable A_1^* satisfies $\mathbb{E}[\log A_1^*]>0$ (see the comments related to Theorem 2.4.4 in [BDM16, p. 48], cf. also the discussion after Proposition 4.2.1) and since $\mathbb{E}[\log A_1]<0$, the strong law of large numbers implies that $\mathbb{P}(\lim_{|i|\to\infty}\log\Theta_i=-\infty)=1$, i.e. $\mathbb{P}(\boldsymbol{\Theta}\in l_0)=\mathbb{P}(\boldsymbol{Y}\in l_0)=1$. By (2.21),

$$\vartheta = \mathbb{P}(\sup_{i\geq 1} Y_i\leq 1) = \mathbb{P}(\sup_{i\geq 1}\prod_{k=1}^{i} A_k \leq 1/Y_0)\,. \tag{2.33}$$

Further, using (2.24) yields that

$$\boldsymbol{Q} \stackrel{d}{=} \Big(\Theta_i, i \in \mathbb{Z} \,\Big|\, \sup_{i\geq 1} \prod_{k=1}^{i} (A_k^*)^{-1} < 1, \sup_{i\geq 1} \prod_{k=1}^{i} A_k \leq 1\Big) . \tag{2.34}$$

Remark 2.3.9. To compare these two classes of examples it is instructive to consider the special case of the process from (2.25) with $d = 1$, $c_j = 0$, $j < 0$, and $c_j = c^j$, $j \geq 0$, for some constant $c \in (0, 1)$. Assume also that the innovations (ξ_i) are nonnegative. By (2.28), the spectral tail process $(\Theta_i)_{i\in\mathbb{Z}}$ of the resulting time series $(X_i)_{i\in\mathbb{Z}}$ is given by

$$\Theta_i = \begin{cases} c^i, & i \geq -J\,, \\ 0, & i < -J\,, \end{cases}$$

where J is a geometric random variable with parameter $\vartheta = 1 - c^\alpha$, i.e. $\mathbb{P}(J = j) = (1 - c^\alpha)c^{\alpha j}$ for $j \geq 0$. On the other hand, observe that $(X_i)_{i\in\mathbb{Z}}$ satisfies the stochastic recurrence equation (2.30) with $A_i = c$ and $B_i = \xi_i$ (note that in this case $\mathbb{E}[A_0^\beta] < 1$ for all $\beta > 0$), i.e.

$$X_i = cX_{i-1} + \xi_i\,,\ i \in \mathbb{Z}\,, \tag{2.35}$$

hence (X_i) is a causal autoregressive process of order one. Moreover, the spectral tail process of (X_i) also has a similar multiplicative structure, in fact, the forward spectral tail process of (X_i) has the same form as the one in (2.31). However, unlike in Example 2.3.8, here always exists a first $i \in \mathbb{Z}$ for which $\Theta_i > 0$. Informally speaking, this reflects the fact that a cluster of exceedances over a large threshold of the process in (2.35) occurs due to the effect of one single large $B_i = \xi_i$, while for processes from Example 2.3.8, exceedances occur as a result of multiplying a large (in fact, asymptotically infinite) number of (tilted) A_k's.

2.3.5 Intensity convergence

Recall that, for each $n \in \mathbb{N}$, $J_{r_n} = \{1, \ldots, r_n\}^d$ and $\boldsymbol{X}_{r_n} = \boldsymbol{X}_{J_{r_n}}$.

Proposition 2.3.10. *If $(r_n)_{n\in\mathbb{N}}$ is a sequence of positive integers satisfying $r_n \to \infty$, $r_n/n \to 0$, and such that (2.18) holds, then as $n \to \infty$,*

$$k_n^d \mathbb{P}(a_n^{-1}\boldsymbol{X}_{r_n} \in \cdot) \xrightarrow{v} \nu(\,\cdot\,) = \vartheta \int_0^\infty \mathbb{P}(y\boldsymbol{Q} \in \cdot)\alpha y^{-\alpha-1} dy \ \text{ in } \mathcal{M}(\tilde{l}_{0,0})\,. \tag{2.36}$$

Remark 2.3.11. Note that ν is a proper element of $\mathcal{M}(\tilde{l}_{0,0})$. Indeed, since $\|\boldsymbol{Q}\|_\infty = 1$, $\nu(\{\boldsymbol{x} \in \tilde{l}_{0,0} : \|\boldsymbol{x}\|_\infty > \epsilon\}) = \vartheta\epsilon^{-\alpha} < \infty$ for all $\epsilon > 0$.

Remark 2.3.12. Observe, since $\nu(\{\boldsymbol{x} : \|\boldsymbol{x}\|_\infty = u\}) = 0$ for all $u > 0$, (2.36) implies that

$$k_n^d \mathbb{P}(M_{r_n} > a_n u) \to \vartheta u^{-\alpha}\,,\ u > 0\,, \tag{2.37}$$

as $n \to \infty$, where $M_{r_n} = \|\boldsymbol{X}_{r_n}\|_\infty$. Moreover, for every $u > 0$,

$$\begin{aligned}\mathbb{P}((a_n u)^{-1}\boldsymbol{X}_{r_n} \in \cdot \mid M_{r_n} > a_n u) &= \frac{k_n^d \mathbb{P}((a_n u)^{-1}\boldsymbol{X}_{r_n} \in \cdot\,,\, M_{r_n} > a_n u)}{k_n^d \mathbb{P}(M_{r_n} > a_n u)} \\ &\xrightarrow{w} \frac{u^\alpha}{\vartheta}\vartheta \int_u^\infty \mathbb{P}(u^{-1} y \boldsymbol{Q} \in \cdot\,)\alpha y^{-\alpha-1} dy \\ &= \int_1^\infty \mathbb{P}(y\boldsymbol{Q} \in \cdot\,)\alpha y^{-\alpha-1} dy = \mathbb{P}(\boldsymbol{Z} \in \cdot\,)\,,\end{aligned}$$

where $\xrightarrow{w}$ denotes weak convergence of finite measures. Hence, for any anchoring function A and all $u > 0$,

$$(a_n u)^{-1}\boldsymbol{X}_{r_n} \mid M_{r_n} > a_n u \xrightarrow{d} \boldsymbol{Z} \stackrel{d}{=} \boldsymbol{Z}^A \text{ in } \tilde{l}_0\,, \tag{2.38}$$

i.e. $\boldsymbol{Z}^A$ represents the asymptotic distribution of a cluster of extremes of $\boldsymbol{X}$. Also, we identify $\boldsymbol{Q}^A$ by

$$M_{r_n}^{-1}\boldsymbol{X}_{r_n} \mid M_{r_n} > a_n u \xrightarrow{d} \boldsymbol{Q} \stackrel{d}{=} \boldsymbol{Q}^A \text{ in } \tilde{l}_0\,. \tag{2.39}$$

In fact, it can be shown that (2.37) and (2.38) for some $\vartheta > 0$ and $\boldsymbol{Z}$ imply (2.36) for $\boldsymbol{Q} := \boldsymbol{Z}/\|\boldsymbol{Z}\|_\infty$, this is actually the approach of [BPS18, Theorem 2.2, Lemma 3.3].

Remark 2.3.13 (Convergence determining family in $\mathcal{M}(\tilde{l}_{0,0})$). To prove Proposition 2.3.10 we will use the following family of functions on $\tilde{l}_0$. For an element $\boldsymbol{x} \in \tilde{l}_0$ and any $\delta > 0$ denote by $\boldsymbol{x}^\delta \in \tilde{l}_0$ the equivalence class of the sequence $(x_{\boldsymbol{i}}\mathbb{1}\{|x_{\boldsymbol{i}}| > \delta\})_{\boldsymbol{i}}$, where $(x_{\boldsymbol{i}})_{\boldsymbol{i}} \in l_0$ is an arbitrary representative of $\boldsymbol{x}$. Let $\mathcal{F}_0$ be the family of all functions $f \in CB_b^+(\tilde{l}_{0,0})$ such that for some $\delta > 0$, $f(\boldsymbol{x}) = f(\boldsymbol{x}^\delta)$ for all $\boldsymbol{x} \in \tilde{l}_0$, where we set $f(\boldsymbol{0}) = 0$, i.e. f depends only on coordinates greater than δ in absolute value.

Since functions from $\mathcal{F}_0$ approximate well functions in $LB_b^+(\tilde{l}_{0,0}, \tilde{d})$, an application of Proposition 1.3.1 yields that for random measures $N, N_1, N_2, \dots$ on $\tilde{l}_{0,0}$, $N_n(f) \xrightarrow{d} N(f)$ for all $f \in \mathcal{F}_0$ implies that $N_n \xrightarrow{d} N$ in $\mathcal{M}(\tilde{l}_{0,0})$, see Lemma 2.5.2 below. In particular, for deterministic measures $\mu, \mu_1, \mu_2, \dots \in \mathcal{M}(\tilde{l}_{0,0})$, $\mu_n(f) \to \mu(f)$ for all $f \in \mathcal{F}_0$ implies that $\mu_n \xrightarrow{v} \mu$ in $\mathcal{M}(\tilde{l}_{0,0})$.

Proof of Proposition 2.3.10. By the Remark 2.3.13, it suffices to show that

$$\lim_{n\to\infty} k_n^d \mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})] = \nu(f)$$

for all $f \in \mathcal{F}_0$. Observe, since $k_n^d = \lfloor n/r_n \rfloor^d \sim (r_n^d \mathbb{P}(|X_{\boldsymbol{0}}| > a_n))^{-1}$ as $n \to \infty$, it is equivalent to prove that

$$\lim_{n\to\infty} \frac{\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})]}{r_n^d \mathbb{P}(|X_{\mathbf{0}}| > a_n)} = \nu(f)\,.$$

First, fix an arbitrary group order $\preceq$ on $\mathbb{Z}^d$ (think of $\preceq$ being the lexicographic order and $d = 2$). Take now an arbitrary $f \in \mathcal{F}_0$ and choose $\epsilon > 0$ small enough such that $f(\boldsymbol{x}) = f(\boldsymbol{x}^\epsilon)$ for all $\boldsymbol{x} \in \tilde{l}_0$. In particular, $\|\boldsymbol{x}\|_\infty \le \epsilon$ implies that $f(\boldsymbol{x}) = 0$, hence,

$$\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})] = \mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})\mathbb{1}\{M_{r_n} > a_n\epsilon\}]\,,$$

and decomposing on the first $\boldsymbol{j} \in J_{r_n}$ (w.r.t. $\preceq$) for which $|X_{\boldsymbol{j}}| > a_n\epsilon$ we get

$$\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})] = \sum_{\boldsymbol{j}\in J_{r_n}} \mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})\mathbb{1}\{\max_{\boldsymbol{j}'\in J_{r_n},\,\boldsymbol{j}'\prec\boldsymbol{j}} |X_{\boldsymbol{j}'}| \le a_n\epsilon, |X_{\boldsymbol{j}}| > a_n\epsilon\}]\,. \tag{2.40}$$

Fix now an $m \in \mathbb{N}$ and take n big enough so that $r_n \ge 2m+1$. Intuitively, for every $\boldsymbol{j} \in J_{r_n}$ such that $\{\boldsymbol{j}' \in \mathbb{Z}^d : |\boldsymbol{j}' - \boldsymbol{j}| \le m\} \subseteq J_{r_n}$, by Assumption 2.3.2, when $|X_{\boldsymbol{j}}| > a_n\epsilon$ we can assume that

$$\max\{|X_{\boldsymbol{j}'}| : \boldsymbol{j}' \in J_{r_n},\, |\boldsymbol{j} - \boldsymbol{j}'| > m\} \le a_n\epsilon\,,$$

and in this case, by the properties of f,

$$f(a_n^{-1}\boldsymbol{X}_{r_n}) = f(a_n^{-1}\boldsymbol{X}_{\{\boldsymbol{j}'\in\mathbb{Z}^d:|\boldsymbol{j}'-\boldsymbol{j}|\le m\}})\,.$$

More precisely, for all such $\boldsymbol{j}$'s, using stationarity, boundedness of f (assume w.l.o.g. that $0 \le f \le 1$) and the fact that $\{\boldsymbol{j}' \in J_{r_n} : |\boldsymbol{j} - \boldsymbol{j}'| > m\} \subseteq \{\boldsymbol{j}' \in \mathbb{Z}^d : m < |\boldsymbol{j} - \boldsymbol{j}'| \le r_n\}$,

$$\begin{aligned}
\Big|\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})\mathbb{1}\{&\max_{\boldsymbol{j}'\in J_{r_n},\,\boldsymbol{j}'\prec\boldsymbol{j}} |X_{\boldsymbol{j}'}| \le a_n\epsilon, |X_{\boldsymbol{j}}| > a_n\epsilon\}] \\
&- \mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{\{|\boldsymbol{j}'|\le m\}})\mathbb{1}\{\max_{|\boldsymbol{j}'|\le m,\,\boldsymbol{j}'\prec\mathbf{0}} |X_{\boldsymbol{j}'}| \le a_n\epsilon, |X_{\mathbf{0}}| > a_n\epsilon\}]\Big| \\
&\qquad\qquad \le \mathbb{P}\left(\max_{m<|\boldsymbol{j}'|\le r_n} |X_{\boldsymbol{j}'}| > a_n\epsilon,\, |X_{\mathbf{0}}| > a_n\epsilon\right)\,.
\end{aligned}$$

For remaining $\boldsymbol{j} \in J_{r_n}$, simply bound the term on the left hand side above by $\mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon)$. Now, going back to (2.40) we can conclude that

$$\begin{aligned}
\Delta_{n,m} := \Big|&\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})]/\{r_n^d\mathbb{P}(|X_{\mathbf{0}}| > a_n)\} \\
&- \mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{\{|\boldsymbol{j}'|\le m\}})\mathbb{1}\{\max_{|\boldsymbol{j}'|\le m,\,\boldsymbol{j}'\prec\mathbf{0}} |X_{\boldsymbol{j}'}| \le a_n\epsilon,\, |X_{\mathbf{0}}| > a_n\epsilon\}]/\mathbb{P}(|X_{\mathbf{0}}| > a_n)\Big| \\
&\le \frac{\mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon)}{\mathbb{P}(|X_{\mathbf{0}}| > a_n)} \left\{\frac{r_n^d - (r_n - 2m)^d}{r_n^d} + \mathbb{P}\left(\max_{m<|\boldsymbol{j}'|\le r_n} |X_{\boldsymbol{j}'}| > a_n\epsilon \,\middle|\, |X_{\mathbf{0}}| > a_n\epsilon\right)\right\}\,.
\end{aligned}$$

Observe, by regular variation $\mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon) \sim \epsilon^{-\alpha}\mathbb{P}(|X_{\mathbf{0}}| > a_n)$ as $n \to \infty$, so together with $r_n \to \infty$ and Assumption 2.3.2 we get that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \Delta_{n,m} = 0\,. \tag{2.41}$$

Next, since f, when viewed as a function on $\mathbb{R}^{d(2m+1)}$, is bounded and continuous, and since $\mathbb{P}(\max_{|\boldsymbol{j}'|\le m,\, \boldsymbol{j}'\prec\mathbf{0}} |Y_{\boldsymbol{j}'}| = |Y_{\mathbf{0}}| \max_{|\boldsymbol{j}'|\le m,\, \boldsymbol{j}'\prec\mathbf{0}} |\Theta_{\boldsymbol{j}'}| = 1) = 0$, by definition of the tail process $\boldsymbol{Y}$ and the continuous mapping theorem,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{\{|\boldsymbol{j}'|\le m\}})\mathbb{1}\{\max_{|\boldsymbol{j}'|\le m,\, \boldsymbol{j}'\prec\mathbf{0}} |X_{\boldsymbol{j}'}| \le a_n\epsilon\} \,\Big|\, |X_{\mathbf{0}}| > a_n\epsilon] \\ = \mathbb{E}[f(\epsilon\boldsymbol{Y}_{\{|\boldsymbol{j}'|\le m\}})\mathbb{1}\{\max_{|\boldsymbol{j}'|\le m,\, \boldsymbol{j}'\prec\mathbf{0}} |Y_{\boldsymbol{j}'}| \le 1\}]\,.\end{aligned}$$

Now since $\mathbb{P}(\lim_{|\boldsymbol{j}'|\to\infty} |Y_{\boldsymbol{j}'}| = 0) = 1$, (2.41) and application of the bounded convergence theorem yield

$$\begin{aligned}\lim_{n\to\infty} \frac{\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})]}{r_n^d\mathbb{P}(|X_{\mathbf{0}}| > a_n)} &= \lim_{m\to\infty} \epsilon^{-\alpha}\mathbb{E}[f(\epsilon\boldsymbol{Y}_{\{|\boldsymbol{j}'|\le m\}})\mathbb{1}\{\max_{|\boldsymbol{j}'|\le m,\, \boldsymbol{j}'\prec\mathbf{0}} |Y_{\boldsymbol{j}'}| \le 1\}] \\ &= \epsilon^{-\alpha}\mathbb{E}[f(\epsilon\boldsymbol{Y})\mathbb{1}\{\max_{\boldsymbol{j}'\prec\mathbf{0}} |Y_{\boldsymbol{j}'}| \le 1\}] \\ &= \epsilon^{-\alpha}\vartheta\mathbb{E}[f(\epsilon\boldsymbol{Z}^{Afe})] = \epsilon^{-\alpha}\vartheta\mathbb{E}[f(\epsilon\boldsymbol{Z})]\,,\end{aligned}$$

where the last equality follows since f is a function on $\tilde{l}_{0,0}$. Since $\|\boldsymbol{Z}\|_\infty$ is Pareto distributed and independent of $\boldsymbol{Q} = \boldsymbol{Z}/\|\boldsymbol{Z}\|_\infty$,

$$\epsilon^{-\alpha}\vartheta\mathbb{E}[f(\epsilon\boldsymbol{Z})] = \epsilon^{-\alpha}\vartheta\int_1^\infty \mathbb{E}[f(\epsilon y\boldsymbol{Q})]\alpha y^{-\alpha-1}dy = \vartheta\int_\epsilon^\infty \mathbb{E}[f(y\boldsymbol{Q})]\alpha y^{-\alpha-1}dy\,.$$

Now since $\|\boldsymbol{Q}\|_\infty = 1$ and since $f(\boldsymbol{x}) = 0$ whenever $\|\boldsymbol{x}\|_\infty \le \epsilon$, we finally obtain that

$$\lim_{n\to\infty} \frac{\mathbb{E}[f(a_n^{-1}\boldsymbol{X}_{r_n})]}{r_n^d\mathbb{P}(|X_{\mathbf{0}}| > a_n)} = \vartheta\int_0^\infty \mathbb{E}[f(y\boldsymbol{Q})]\alpha y^{-\alpha-1}dy = \nu(f)\,.$$

□

2.3.6 Convergence to a (compound) Poisson process

In view of Proposition 2.3.10, our main point process convergence result given below is now a simple application of Corollary 1.4.4. For that purpose, we introduce a convergence determining family in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$ related the family $\mathcal{F}_0$ from Remark 2.3.13.

Let $\mathcal{F}_0'$ be the family of all functions $f \in CB_b^+([0,1]^d \times \tilde{l}_{0,0})$ such that for some $\delta > 0$, $f(\boldsymbol{t},\boldsymbol{x}) = f(\boldsymbol{t},\boldsymbol{x}^\delta)$ for all $\boldsymbol{t} \in [0,1]^d$ and $\boldsymbol{x} \in \tilde{l}_0$, where we set $f(\boldsymbol{t},\mathbf{0}) = 0$ (the mapping $\boldsymbol{x} \mapsto \boldsymbol{x}^\delta$ was defined in Remark 2.3.13). As a consequence of Lemma 2.5.2 below, $\mathcal{F}_0'$ is point process convergence determining, see Remark 2.5.4.

Theorem 2.3.14. *Let $\boldsymbol{X}$ be a stationary regularly varying random field with tail index $\alpha > 0$ and $(r_n)_{n\in\mathbb{N}}$ a sequence of positive integers satisfying $r_n \to \infty$, $r_n/n \to 0$. If (2.18) holds and the family $((\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n) : n \in \mathbb{N},\ \boldsymbol{i} \in I_n)$ is AI$(\mathcal{F}_0')$, then*

$$N_n' = \sum_{\boldsymbol{i}\in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)} \xrightarrow{d} N' = \sum_{i\in\mathbb{N}} \delta_{(\boldsymbol{T}_i, P_i \boldsymbol{Q}^i)} \tag{2.42}$$

in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$, where $N' \stackrel{d}{=} \mathrm{PPP}(Leb \times \nu)$ and

(i) $\sum_{i\in\mathbb{N}} \delta_{(\boldsymbol{T}_i, P_i)}$ is a Poisson point process on $[0,1]^d \times (0,\infty)$ with intensity measure $Leb \times d(-\vartheta y^{-\alpha})$;

(ii) $\boldsymbol{Q}^i = (Q^i_{\boldsymbol{j}})_{\boldsymbol{j}\in\mathbb{Z}^d}$, $i \in \mathbb{N}$ is a sequence of i.i.d. elements in $\tilde{l}_0$, independent of $\sum_{i\in\mathbb{N}} \delta_{(\boldsymbol{T}_i, P_i)}$ and with common distribution equal to the distribution of $\boldsymbol{Q}$.

Proof. An application of Corollary 1.4.4 together with Proposition 2.3.10 yields convergence in distribution of N_n' to a Poisson process with intensity measure $Leb \times \nu$. Hence, it only remains to show that the point process in N' in (2.42) is indeed Poisson with intensity measure $Leb \times \nu$.

Define the subset $\mathbb{S}$ of $\tilde{l}_0$ by $\mathbb{S} = \{\boldsymbol{x} \in \tilde{l}_0 : \|\boldsymbol{x}\|_\infty = 1\}$. By [Res87, Proposition 3.8], point process $\sum_{i\in\mathbb{N}} \delta_{(\boldsymbol{T}_i, P_i, \boldsymbol{Q}^i)}$ is a Poisson point process on $[0,1]^d \times (0,\infty) \times \mathbb{S}$ with intensity measure $Leb \times d(-\vartheta y^{-\alpha}) \times \mathbb{P}_{\boldsymbol{Q}}$, where $\mathbb{P}_{\boldsymbol{Q}}$ is the distribution of $\boldsymbol{Q}$. Note that [Res87, Proposition 3.8] applies to point processes on locally compact spaces, but its proof is easily extended to a more general state space. Further, define the mapping $T : [0,1]^d \times (0,\infty) \times \mathbb{S} \to [0,1]^d \times \tilde{l}_{0,0}$ by $T(\boldsymbol{t}, p, \boldsymbol{q}) = (\boldsymbol{t}, p\boldsymbol{q})$. Now an application of [Res87, Proposition 3.8] yields that the transformed point process $N_n' = \sum_{i\in\mathbb{N}} \delta_{(\boldsymbol{T}_i, P_i \boldsymbol{Q}^i)}$ is again Poisson with intensity measure $(Leb \times d(-\vartheta y^{-\alpha}) \times \mathbb{P}_{\boldsymbol{Q}}) \circ T^{-1} = Leb \times \nu$. □

Corollary 2.3.15. *Under notation of Theorem 2.3.14, if there exists a sequence $r_n \to \infty$, $r_n/n \to 0$ for which (2.42) holds, then, with $J_n = \{1,\dots,n\}^d$,*

$$\sum_{\boldsymbol{j}\in J_n} \delta_{(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)} \xrightarrow{d} \sum_{i\in\mathbb{N}} \sum_{\boldsymbol{j}\in\mathbb{Z}^d} \delta_{(\boldsymbol{T}_i, P_i Q^i_{\boldsymbol{j}})} \tag{2.43}$$

in $\mathcal{M}_p([0,1]^d \times (\mathbb{R}\setminus\{0\}))$ with bounded sets being those which are bounded away from $[0,1]^d \times \{0\}$.

The proof of the previous result is in Section 2.5; it first applies continuous mapping theorem to convergence (2.42) and then uses the fact that (assume for simplicity that $d = 1$) time instances j/n and i/k_n for $i \in I_n = \{1,\dots,k_n\}$, $j \in J_{n,i} = \{(i-1)r_n + 1,\dots,ir_n\}$, differ by at most $2r_n/n$ which tends to zero as $n \to \infty$. Note, for $d = 1$, the previous result corresponds to [BT16, Theorem 3.1].

Remark 2.3.16 (**Extremal index**). As already noticed by [BS09, Remark 4.7], when convergence in (2.43) holds, the quantity ϑ is the *extremal index* of the field $(|X_{\boldsymbol{j}}|)_{\boldsymbol{j}\in\mathbb{Z}^d}$ since $n^d\mathbb{P}(|X_{\boldsymbol{0}}| > a_n u) \to u^{-\alpha}$ and

$$\mathbb{P}(\max_{\boldsymbol{j}\in J_n} |X_{\boldsymbol{j}}| \leq a_n u) \to \mathbb{P}\left(\sum_{i\in\mathbb{N}} \mathbb{1}_{\{P_i > u\}} = 0\right) = e^{-\vartheta u^{-\alpha}} ,$$

as $n \to \infty$, for all $u > 0$.

In the rest of the chapter we focus on techniques for verifying assumptions of Theorem 2.3.14 or obtaining the convergence (2.42).

First note that these assumptions hold in the case of m–dependent stationary fields, see Lemma 2.4.1. Moreover, one can extend convergence in (2.42) to fields which can be approximated by m–dependent fields, such as infinite order moving average processes from Example 2.3.7. This is the content of Section 2.4.1.

In the time series case, the concept of strong mixing offers a way to check the $AI(\mathcal{F}_0')$ condition. Corresponding conditions are satisfied for a wide family of geometrically ergodic Markov chains which (under mild additional conditions) include solutions of stochastic recurrence equations from Example 2.3.8. This is discussed in Section 2.4.2.

On the other hand, for proper random fields, i.e. when $d > 1$, the approach of Section 1.4.2 using the concept of neighborhoods can be used to check the asymptotic $\mathcal{F}_0'$–independence condition. This is illustrated on the local sequence alignment problem studied in Chapter 4, see in particular Corollary 4.3.4.

2.4 Checking assumptions of Theorem 2.3.14

2.4.1 Fields admitting m–dependent approximation

A random field $\boldsymbol{X} = (X_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ is said to be m–dependent for some $m \in \mathbb{N}$ if for all finite $I, J \subseteq \mathbb{Z}^d$ such that $\inf\{|\boldsymbol{i} - \boldsymbol{j}| : \boldsymbol{i} \in I, \boldsymbol{j} \in J\} > m$, σ-algebras $\sigma(X_{\boldsymbol{i}} : \boldsymbol{i} \in I)$ and $\sigma(X_{\boldsymbol{j}} : \boldsymbol{j} \in J)$ are independent. Observe, if $(a_n)_n$ is chosen such that $n^d\mathbb{P}(|X_{\boldsymbol{0}}| > a_n) \to 1$, m–dependence implies that condition 2.18 is satisfied for any $(r_n)_n$ such that $r_n \to \infty$ and $r_n/n \to 0$. Moreover, using the properties of the family $\mathcal{F}_0'$, it is easy to show that the asymptotic $\mathcal{F}_0'$–independence assumption of Theorem 2.3.14 is also satisfied; cf. the proof of [Bas00, Lemma 2.3.9], see also Section 4.3 below.

Lemma 2.4.1. *If a stationary regularly varying random field $\boldsymbol{X}$ is m–dependent for some $m \in \mathbb{N}$, then $\boldsymbol{X}$ satisfies assumptions of Theorem 2.3.14 for any $(r_n)_n$ such that $r_n \to \infty$ and $r_n/n \to 0$.*

Let $\boldsymbol{X} = (X_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{Z}^d)$ now be a general stationary random field (not necessarily regularly varying). Assume that there exists a sequence of stationary regularly varying

m–dependent fields $\boldsymbol{X}^{(m)} = (X_{\boldsymbol{i}}^{(m)} : \boldsymbol{i} \in \mathbb{Z}^d)$, $m \in \mathbb{N}$, and a sequence of real numbers $(b_n)_{n\in\mathbb{N}}$ such that for all $m \in \mathbb{N}$

$$n^d \mathbb{P}(|X_{\boldsymbol{0}}^{(m)}| > b_n) \to d^{(m)} > 0\,, \text{ as } n \to \infty\,. \tag{2.44}$$

Observe that we keep the same normalizing sequence (b_n) for all fields. In particular, the tail index of $\boldsymbol{X}^{(m)}$ is the same for all m. Denote it by $\alpha > 0$. Further, for each $\boldsymbol{X}^{(m)}$ denote by $\vartheta^{(m)}$ the quantity defined in (2.20) and by $\boldsymbol{Q}^{(m)}$ the random element in $\tilde{l}_0$ defined in (2.23).

Assumption 2.4.2. *(i) There exists $\sigma > 0$ and a random element $\boldsymbol{Q}$ in $\tilde{l}_0$ such that, as $m \to \infty$, $\vartheta^{(m)} d^{(m)} \to \sigma$ and $\boldsymbol{Q}^{(m)} \xrightarrow{d} \boldsymbol{Q}$ in $\tilde{l}_0$.*

(ii) For any $u > 0$

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(\max_{\boldsymbol{1}\le\boldsymbol{i}\le\boldsymbol{1}\cdot n} |X_{\boldsymbol{i}}^{(m)} - X_{\boldsymbol{i}}| > b_n u) = 0\,.$$

Note that, since $\|\boldsymbol{Q}^{(m)}\| = 1$ for all $m \in \mathbb{N}$, the same holds for $\boldsymbol{Q}$. Further, let (r_n) be any sequence of positive integers satisfying $r_n \to \infty$ and $k_n := \lfloor n/r_n \rfloor \to \infty$. Recall the blocks $\boldsymbol{X}_{n,\boldsymbol{i}}$, $\boldsymbol{i} \in I_n = \{1, \ldots, k_n\}^d$ defined in (2.12).

Theorem 2.4.3. *Assume that $(b_n)_{n\in\mathbb{N}}$ is a sequence of real numbers satisfying (2.44) and that Assumption 2.4.2 holds for some $\sigma > 0$ and random element $\boldsymbol{Q}$ in $\tilde{l}_0$. Then*

$$\sum_{\boldsymbol{i}\in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/b_n)} \xrightarrow{d} \sum_{i\in\mathbb{N}} \delta_{(\boldsymbol{T}_i, P_i \boldsymbol{Q}^i)} \tag{2.45}$$

in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$, where $\sum_{i=1}^{\infty} \delta_{(T_i,P_i)}$ is a Poisson point process on $[0,1]^d \times (0,\infty)$ with intensity measure $\sigma Leb \times d(-y^{-\alpha})$, independent of the i.i.d. sequence $(\boldsymbol{Q}^i)_{i\in\mathbb{N}}$ with common distribution equal to the distribution of $\boldsymbol{Q}$.

Remark 2.4.4. The fields $\boldsymbol{X}^{(m)}$ can in general be n_m–dependent for a sequence $n_m \to \infty$ as $m \to \infty$.

Example 2.4.5 (**Moving averages**). Consider again the regularly varying field $\boldsymbol{X} = (X_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}$ from Example 2.3.7 defined by $X_{\boldsymbol{i}} = \sum_{\boldsymbol{j}\in\mathbb{Z}^d} c_{\boldsymbol{j}} \xi_{\boldsymbol{i}-\boldsymbol{j}}$ for a field of real numbers $(c_{\boldsymbol{j}})$ and field of i.i.d. random variables $(\xi_{\boldsymbol{j}})$ which are regularly varying with index $\alpha > 0$.

Instead of checking assumptions of Theorem 2.3.14 for the field $\boldsymbol{X}$, we show that Theorem 2.4.3 can be elegantly applied. For each $m \in \mathbb{N}$ and $\boldsymbol{i} \in \mathbb{Z}^d$ define

$$X_{\boldsymbol{i}}^{(m)} = \sum_{|\boldsymbol{j}|\le m} c_{\boldsymbol{j}} \xi_{\boldsymbol{i}-\boldsymbol{j}}\,.$$

By assumption, the field $\boldsymbol{X}^{(m)} = (X_{\boldsymbol{i}}^{(m)} : \boldsymbol{i} \in \mathbb{Z}^d)$ is in general $(2m+1)$–dependent. Take a sequence of real numbers (b_n) such that $\lim_{n\to\infty} n^d \mathbb{P}(|\xi_{\boldsymbol{0}}| > b_n) = 1$. It now follows from

(2.27) that (2.44) is satisfied with $d^{(m)} = \sum_{|\boldsymbol{j}|\le m} |c_{\boldsymbol{j}}|^\alpha$. Further, for each $m \in \mathbb{N}$, special case of (2.29) for $c_{\boldsymbol{j}} = 0$, $|\boldsymbol{j}| > m$ yields

$$\vartheta^{(m)} = \frac{\max_{|\boldsymbol{j}|\le m} |c_{\boldsymbol{j}}|^\alpha}{\sum_{|\boldsymbol{j}|\le m} |c_{\boldsymbol{j}}|^\alpha}, \quad \boldsymbol{Q}^{(m)} \stackrel{d}{=} \left(\frac{K c_{\boldsymbol{j}} \mathbb{1}\{|\boldsymbol{j}| \le m\}}{\max_{|\boldsymbol{i}|\le m} |c_{\boldsymbol{i}}|} \right)_{\boldsymbol{j}\in\mathbb{Z}^d} \text{ in } \tilde{l}_0 .$$

Observe, since $\lim_{|\boldsymbol{j}|\to\infty} |c_{\boldsymbol{j}}| = 0$, Assumption 2.4.2 (i) is satisfied for $\sigma = \max_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^\alpha$ and $\boldsymbol{Q}$ from (2.29). Finally, (2.27) implies that Assumption 2.4.2 (ii) holds; see [DR85, Lemma 2.3].

Thus, all of the conditions of Theorem 2.4.3 are met and hence convergence in (2.45) holds. Moreover, the limiting point process can be represented as $\sum_{i=1}^\infty \delta_{(\boldsymbol{T}_i, P_i K_i (c_{\boldsymbol{j}})_{\boldsymbol{j}})}$ where $\sum_{i=1}^\infty \delta_{(\boldsymbol{T}_i, P_i)}$ is a Poisson point process on $[0,1]^d \times (0,\infty)$ with intensity measure $Leb \times d(-y^{-\alpha})$ and (K_i) is an i.i.d. sequence of random variables distributed as K and independent of $\sum_{i=1}^\infty \delta_{(\boldsymbol{T}_i, P_i)}$.

Note, for $a_n = b_n(\sum_{\boldsymbol{j}\in\mathbb{Z}^d} |c_{\boldsymbol{j}}|^\alpha)^{1/\alpha}$, (2.27) implies that $n^d \mathbb{P}(|X_{\boldsymbol{0}}| > a_n) \to 1$, so by applying continuous mapping theorem to (2.45) one obtains that the conclusion of Theorem 2.3.14 holds for the regularly varying field $(X_{\boldsymbol{i}})$, i.e. that the convergence (2.42) holds for any $r_n \to \infty, r_n/n \to 0$, with the corresponding ϑ and $\boldsymbol{Q}$ from (2.29).

Proof of Theorem 2.4.3. For every $\boldsymbol{X}^{(m)}$ denote by $\boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}$ the corresponding blocks from (2.12) and define point processes $N_n^{(m)}$ on $[0,1]^d \times \tilde{l}_{0,0}$ by

$$N_n^{(m)} = \sum_{\boldsymbol{i}\in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}/b_n)}, \; n \in \mathbb{N} .$$

Note, for each $m \in \mathbb{N}$, by (2.44) and regular variation of $|X_{\boldsymbol{0}}^{(m)}|$, the sequence $a_n^{(m)} := b_n (d^{(m)})^{\frac{1}{\alpha}}$, $n \in \mathbb{N}$, satisfies (2.16), i.e. $\lim_{n\to\infty} n^d \mathbb{P}(|X_{\boldsymbol{0}}^{(m)}| > a_n^{(m)}) = 1$. Since $\boldsymbol{X}^{(m)}$ is m–dependent, by Lemma 2.4.1 we can apply Theorem 2.3.14 which, together with an an application of the continuous mapping theorem, implies that for each $m \in \mathbb{N}$,

$$N_n^{(m)} \xrightarrow{d} N^{(m)} \stackrel{d}{=} \mathrm{PPP}(Leb \times \nu^{(m)}) , \quad \text{as } n \to \infty ,$$

in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$, where

$$\nu^{(m)}(\,\cdot\,) = \vartheta^{(m)} d^{(m)} \int_0^\infty \mathbb{P}(y\boldsymbol{Q}^{(m)} \in \cdot) \alpha y^{-\alpha-1} dy .$$

Further, by Assumption 2.4.2 (i) and the dominated convergence theorem, as $m \to \infty$,

$$\nu^{(m)}(f) \to \sigma \int_0^\infty \mathbb{E}[f(y\boldsymbol{Q})] \alpha y^{-\alpha-1} dy ,$$

for all $f \in CB_b^+(\tilde{l}_{0,0})$, i.e.

$$\nu^{(m)} \xrightarrow{v} \nu^{(\infty)}(\,\cdot\,) = \sigma \int_0^\infty \mathbb{P}(y\boldsymbol{Q} \in \cdot)\alpha y^{-\alpha-1} dy\,,$$

in $\mathcal{M}(\tilde{l}_{0,0})$. This implies that (see e.g. [Res07, Problem 5.3])

$$\mathrm{PPP}(Leb \times \nu^{(m)}) \stackrel{d}{=} N^{(m)} \xrightarrow{d} N^{(\infty)} \stackrel{d}{=} \mathrm{PPP}(Leb \times \nu^{(\infty)})\,.$$

Set $N_n'' = \sum_{\boldsymbol{i} \in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/b_n)}$ for all $n \in \mathbb{N}$. Since the distribution of $N^{(\infty)}$ coincides with the distribution of the limit in (2.45), to prove convergence in (2.45) it suffices to show that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(|N_n^{(m)}(f) - N_n''(f)| > \eta) = 0\,, \tag{2.46}$$

for all $\eta > 0$ and every f in some family of functions on $[0,1]^d \times \tilde{l}_{0,0}$ which is point process convergence determining. Indeed, by [Bil68, Theorem 4.2], this implies that $N_n''(f) \xrightarrow{d} N^{(\infty)}(f)$ in $\mathbb{R}$ for all such f, and hence $N_n'' \xrightarrow{d} N^{(\infty)}$ in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$.

Recall the metric $\tilde{d}$ on $\tilde{l}_0$ defined in (2.13). Adapting the lines of the proof of [DR85, Theorem 2.4, equation (2.11)], we show that Assumption 2.4.2 (ii) implies (2.46) for every $f \in LB_b^+([0,1]^d \times \tilde{l}_{0,0}, d')$ where d' is a metric on $[0,1]^d \times \tilde{l}_{0,0}$ defined by $d'((\boldsymbol{t},\boldsymbol{x}),(\boldsymbol{s},\boldsymbol{y})) = |\boldsymbol{t}-\boldsymbol{s}| \vee d(\boldsymbol{x},\boldsymbol{y})$ for all $\boldsymbol{t},\boldsymbol{s} \in [0,1]^d$, $\boldsymbol{x},\boldsymbol{y} \in \tilde{l}_{0,0}$, with $|\cdot|$ denoting the sup–norm on $[0,1]^d$. Since, by Proposition 1.3.1, the family $LB_b^+([0,1]^d \times \tilde{l}_{0,0}, d')$ is point process convergence determining, this will prove the result.

Fix an $f \in LB_b^+([0,1]^d \times \tilde{l}_{0,0}, d')$ and let $\epsilon > 0$ be such that $\|\boldsymbol{x}\|_\infty \le \epsilon$ implies that $f(\boldsymbol{t},\boldsymbol{x}) = 0$ for all $\boldsymbol{t} \in [0,1]^d$. Assume without loss of generality that $|f(\boldsymbol{t},\boldsymbol{x}) - f(\boldsymbol{s},\boldsymbol{y})| \le |\boldsymbol{t}-\boldsymbol{s}| \vee \tilde{d}(\boldsymbol{x},\boldsymbol{y})$ for all $(\boldsymbol{t},\boldsymbol{x}),(\boldsymbol{s},\boldsymbol{y}) \in [0,1]^d \times \tilde{l}_{0,0}$. In particular,

$$|f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}/b_n) - f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/b_n)| \le \frac{\tilde{d}(\boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}, \boldsymbol{X}_{n,\boldsymbol{i}})}{b_n}\,, \tag{2.47}$$

for all $n,m \in \mathbb{N}$ and $\boldsymbol{i} \in I_n$.

For all $r > 0$ set $B(r) = \{\boldsymbol{x} \in \tilde{l}_{0,0} : \|\boldsymbol{x}\|_\infty \le r\}$ and for an arbitrary $0 < u < \epsilon/2$ define the events $A_n^{(m)} = \{\max_{\boldsymbol{1} \le \boldsymbol{i} \le n \cdot \boldsymbol{1}} |X_{\boldsymbol{i}}^{(m)} - X_{\boldsymbol{i}}| \le b_n u\}$, $n,m \in \mathbb{N}$. Note that $A_n^{(m)} \subseteq \{\max_{\boldsymbol{i} \in I_n} \tilde{d}(\boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}, \boldsymbol{X}_{n,\boldsymbol{i}}) \le b_n u\}$. Hence, on the event $A_n^{(m)}$, if $\boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}/b_n \in B(\epsilon/2)$ for some $\boldsymbol{i} \in I_n$ then $f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}^{(m)}/b_n) = f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/b_n) = 0$. Together with Assumption 2.4.2 (ii) and (2.47) this yields that

$$\begin{aligned}\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(|N_n^{(m)}(f) - N_n''(f)| > \eta) &\le \lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(\{|N_n^{(m)}(f) - N_n''(f)| > \eta\} \cap A_n^{(m)}) \\ &\le \lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(u N_n^{(m)}([0,1]^d \times B(\epsilon/2)^c) > \eta)\,.\end{aligned}$$

Since $N_n^{(m)} \xrightarrow{d} N^{(m)}$ as $n \to \infty$ and $N^{(m)} \xrightarrow{d} N^{(\infty)}$ as $m \to \infty$, and since all of the limiting point processes a.s. put zero mass on the boundary of the bounded set $[0,1]^d \times B(\epsilon/2)^c$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}(N_n^{(m)}([0,1]^d \times B(\epsilon/2)^c) > u^{-1}\eta) = \mathbb{P}(N^{(\infty)}([0,1]^d \times B(\epsilon/2)^c) > u^{-1}\eta)$$

Finally, since $N^{(\infty)}$ is a.s. finite on the set $[0,1]^d \times B(\epsilon/2)^c$, letting $u \to 0$ yields (2.46) and since $f \in LB_b^+([0,1]^d \times \tilde{l}_{0,0}, d')$ was arbitrary this finishes the proof.

□

2.4.2 Strongly mixing time series

Let $\boldsymbol{X} = (X_i)_{i\in\mathbb{Z}}$ be a stationary regularly varying time series. Define the *strong mixing coefficients*

$$\alpha_l = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \sigma(X_i : i \le 0), A \in \sigma(X_i : i \ge l)\}\,,\; l \in \mathbb{N}\,.$$

As usual, let $(a_n)_n$ be such that $n\mathbb{P}(|X_0| > a_n) \to 1$.

Lemma 2.4.6. *Let $(r_n)_n$ be a sequence of positive integers satisfying $r_n \to \infty, k_n = \lfloor n/r_n \rfloor \to \infty$ as $n \to \infty$. If there exists a sequence $(l_n)_n$ of positive integers such that*

$$l_n/r_n \to 0\,,\; k_n\alpha_{l_n} \to 0\,,$$

then the family $((\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n) : n \in \mathbb{N},\, \boldsymbol{i} \in I_n)$ is $AI(\mathcal{F}_0')$.

Using the properties of the family $\mathcal{F}_0'$, the proof of the previous result is essentially the same as the proof of [Bas00, Lemma 2.3.9] and therefore omitted.

Even though intuitively clear, computing the mixing coefficients is in general a difficult task. Fortunately, if $\boldsymbol{X}$ is a stationary Markov chain, conditions of the previous lemma can be verified by showing that $\boldsymbol{X}$ is geometrically ergodic, see [Bas00, Section 2.2.2] for more details and a sufficient condition for geometric ergodicity. In this case, assumptions of Lemma 2.4.6 are satisfied for $r_n = \lfloor n^\epsilon \rfloor$ for every $\epsilon \in (0,1)$, see [Bas00, Remark 2.3.10].

Example 2.4.7 (**Solutions to stochastic recurrence equations**). As in Example 2.3.8, let $(X_i)_{i\in\mathbb{Z}}$ be the solution to the stochastic recurrence equation

$$X_i = A_i X_{i-1} + B_i\,,\; i \in \mathbb{Z}\,,$$

where $((A_i, B_i))_{i\in\mathbb{Z}}$ are i.i.d. $[0,\infty)^2$–random vectors such that for some $\alpha > 0$ conditions of [BDM16, Theorem 2.4.4] hold, in particular $\mathbb{E}[A^\alpha] = 1$, $\mathbb{E}[\log A] < 0$ and $\mathbb{E}[B^\alpha] < \infty$.

By [Bas00, Lemma 3.2.7 and Proposition 3.2.9], if either A_0 or B_0 has a density, the Markov chain (X_i) is geometrically ergodic, hence, for $r_n = \lfloor n^\epsilon \rfloor$ for any $\epsilon \in (0,1)$, the corresponding blocks $\boldsymbol{X}_{n,i}$ satisfy the asymptotic $\mathcal{F}'_0$–independence condition. Moreover, by [Bas00, Lemma 4.1.4], for all $\epsilon > 0$ small enough, the condition (2.18) is also satisfied. Hence, for such (r_n)'s all assumptions of Theorem 2.3.14 are satisfied.

Remark 2.4.8. One can define the strong mixing coefficients for general stationary random fields $(X_{\boldsymbol{i}})_{\boldsymbol{i}\in\mathbb{Z}^d}$, i.e. when $d > 1$ (see e.g. [Ros85, p. 73] and also [Bra93]), so that an analogue of Lemma 2.4.6 holds. However, investigation of practical techniques for checking its assumptions, such as geometric ergodicity in the case of time series, and corresponding class of examples is out of the scope of this thesis.

2.5 Postponed proofs

Proof of Theorem 2.2.1

We only prove (iii)⇒(i) since (i)⇒(ii) follows as in [BS09, Theorem 2.1] and (ii)⇒(iii) is obvious. Also, since we essentially adapt the arguments of [BS09, Theorem 2.1], some details are omitted.

Observe first that (2.5) with $I = \{\boldsymbol{0}\}$ implies that for all $\epsilon > 0$,

$$\lim_{u\to\infty} \frac{\mathbb{P}(|X_{\boldsymbol{0}}| > u\epsilon)}{\mathbb{P}(|X_{\boldsymbol{0}}| > u)} = \epsilon^{-\alpha}\,, \tag{2.48}$$

and moreover that $X_{\boldsymbol{0}}$ is a regularly varying random variable with index α, see [BS09, Theorem 2.1].

Take now an arbitrary finite $I \subseteq \mathbb{Z}^d$ such that $|I| \geq 2$ and consider the space $\mathbb{R}^{|I|} \setminus \{\boldsymbol{0}\}$ with bounded sets being those which are contained in sets $B_\epsilon := \{(x_{\boldsymbol{i}})_{\boldsymbol{i}\in I} \in \mathbb{R}^{|I|} : \sup_{\boldsymbol{i}\in I} |x_{\boldsymbol{i}}| > \epsilon\}$, $\epsilon > 0$. In view of (2.48), multivariate regular variation (with index α) of $\boldsymbol{X}_I$ is equivalent to the existence of a non–zero measure $\mu_I \in \mathcal{M}(\mathbb{R}^{|I|} \setminus \{\boldsymbol{0}\})$ such that

$$\mu_u^I(\,\cdot\,) := \frac{\mathbb{P}(u^{-1}\boldsymbol{X}_I \in \cdot)}{\mathbb{P}(|X_{\boldsymbol{0}}| > u)} \xrightarrow{v} \mu_I\,,\ \text{as}\ n \to \infty\,, \tag{2.49}$$

see [SZM17, Definition 3.1, Proposition 3.1] (cf. [BS09, Equation (1.3)]).

Arguing exactly as in [BS09, Theorem 2.1] it follows that the vague limit of μ_u^I, if it exists, is necessarily non–zero, and furthermore, that $\limsup_{u\to\infty} \mu_u^I(B_\epsilon) \leq |I|\epsilon^{-\alpha} < \infty$ for every $\epsilon > 0$. Since sets $\{(x_{\boldsymbol{i}})_{\boldsymbol{i}\in I} \in \mathbb{R}^{|I|} : \sup_{\boldsymbol{i}\in I} |x_{\boldsymbol{i}}| \in [\epsilon, M]\}$ are compact for every $\epsilon, M > 0$, by [Kal17, Theorem 4.2] it follows that the set $\{\mu_u^I : u > 0\}$ is relatively compact in the vague topology of $\mathcal{M}(\mathbb{R}^{|I|} \setminus \{\boldsymbol{0}\})$.

Since $\mathcal{I}$ is encompassing, we can take $\boldsymbol{i}^* \in I$ such that $I' := I - \boldsymbol{i}^* \subseteq \mathcal{I}$. By [BS09, Lemma 2.2], to show that measures μ_u^I vaguely converge as $u \to \infty$, it suffices to prove that $\lim_{u\to\infty} \mu_u^I(f)$ exists for all $f \in \mathcal{F}$ where $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \subseteq CB_b^+(\mathbb{R}^{|I|} \setminus \{\mathbf{0}\})$ with

$$\begin{aligned}
\mathcal{F}_1 &= \{f : \text{for some } \epsilon > 0, f((x_{\boldsymbol{i}})_{\boldsymbol{i}\in I}) = 0 \text{ if } |x_{\boldsymbol{i}^*}| \le \epsilon\}\,,\\
\mathcal{F}_2 &= \{f : f((x_{\boldsymbol{i}})_{\boldsymbol{i}\in I}) \text{ does not depend on } x_{\boldsymbol{i}^*}\}\,.
\end{aligned}$$

Note that families $\mathcal{F}_1$ and $\mathcal{F}_2$ depend on I but we omit this in the notation.

Since $I' \subseteq \mathcal{I}$, stationarity, (2.5) and (2.48) imply that for every $f \in \mathcal{F}_1$ and $\epsilon > 0$ as in the definition of $\mathcal{F}_1$,

$$\mu_u^I(f) = \frac{\mathbb{P}(|X_{\mathbf{0}}| > u\epsilon)}{\mathbb{P}(|X_{\mathbf{0}}| > u)} \cdot \mathbb{E}[f(u^{-1}\boldsymbol{X}_{I'}) \mid |X_{\mathbf{0}}| > u\epsilon] \to \epsilon^{-\alpha}\mathbb{E}[f(\epsilon(Y_{\boldsymbol{i}})_{\boldsymbol{i}\in I'})]\,, \text{ as } u \to \infty\,.$$

Further, every $f \in \mathcal{F}_2$ naturally induces a function $\tilde{f}$ in $CB_b^+(\mathbb{R}^{|I|-1} \setminus \{\mathbf{0}\})$ and by stationarity

$$\mu_u^I(f) = \frac{\mathbb{E}[\tilde{f}(u^{-1}\boldsymbol{X}_{I\setminus\{\boldsymbol{i}^*\}})]}{\mathbb{P}(|X_{\mathbf{0}}| > u)} = \mu_u^{I\setminus\{\boldsymbol{i}^*\}}(\tilde{f})\,.$$

Hence, $\lim_{u\to\infty} \mu_u^I(f)$ exists for all $f \in \mathcal{F}_2$ if $\boldsymbol{X}_{I\setminus\{\boldsymbol{i}^*\}}$ is multivariate regularly varying.

Observe, we have shown that for an arbitrary finite $I \subseteq \mathbb{Z}^d$ such that $|I| \ge 2$, $\boldsymbol{X}_I$ is multivariate regularly varying if $\boldsymbol{X}_{I\setminus\{\boldsymbol{i}^*\}}$ is, where $\boldsymbol{i}^* \in I$ is such that $I - \boldsymbol{i}^* \subseteq \mathcal{I}$. Therefore, (i) now follows by regular variation of $X_{\mathbf{0}}$ and since $\mathcal{I}$ is encompassing.

Metric on the space $\tilde{l}_0$

Let $(\mathbb{X}, d)$ be a metric space. Assume that $\sim$ is an equivalence relation on $\mathbb{X}$ and let $\tilde{\mathbb{X}}$ be the induced quotient space. Define a function $\tilde{d} : \tilde{\mathbb{X}} \times \tilde{\mathbb{X}} \to [0, \infty)$ by:

$$\tilde{d}(\tilde{x}, \tilde{y}) = \inf\{d(x, y) : x \in \tilde{x}, y \in \tilde{y}\}\,,\ \tilde{x}, \tilde{y} \in \tilde{\mathbb{X}}\,.$$

Lemma 2.5.1. *If for all $\tilde{x}, \tilde{y} \in \tilde{\mathbb{X}}$ and all $y \in \tilde{y}$,*

$$\tilde{d}(\tilde{x}, \tilde{y}) = \inf_{x\in\tilde{x}} d(x, y)\,, \tag{2.50}$$

then $\tilde{d}$ is a pseudo-metric on $\tilde{\mathbb{X}}$. If moreover $(\mathbb{X}, d)$ is separable and/or complete, then so is $(\tilde{\mathbb{X}}, \tilde{d})$.

Proof. To prove that $\tilde{d}$ is a pseudo-metric, the only nontrivial step is to show that $\tilde{d}$ satisfies the triangle inequality, but this is implied by (2.50). Indeed, take any $\tilde{x}, \tilde{y}, \tilde{z} \in \tilde{\mathbb{X}}$ and fix an arbitrary $z \in \tilde{z}$. Since (2.50), then for all $y \in \tilde{y}$

$$\tilde{d}(\tilde{x}, \tilde{y}) = \inf_{x\in\tilde{x}} d(x, y) \le \inf_{x\in\tilde{x}}(d(x, z) + d(z, y)) = \tilde{d}(\tilde{x}, \tilde{z}) + d(z, y)\,.$$

Now taking the infimum over all $y \in \tilde{y}$ and using (2.50) again yields the triangle inequality.

Further, it follows easily that $(\tilde{\mathbb{X}}, \tilde{d})$ is separable whenever (X, d) is. Assume now that $(\mathbb{X}, d)$ is complete and let $(\tilde{x}_n)_{n\in\mathbb{N}}$ be an arbitrary Cauchy sequence in $(\tilde{\mathbb{X}}, \tilde{d})$. Then we can find a strictly increasing sequence of nonnegative integers $(n_k)_{k\in\mathbb{N}}$ such that

$$\tilde{d}(\tilde{x}_m, \tilde{x}_n) < \frac{1}{2^{k+1}} ,$$

for all $m, n \geq n_k$ and every $k \geq 1$. We define a sequence $(y_k)_{k\in\mathbb{N}}$ in $\mathbb{X}$ inductively as follows:

- Let y_1 be an arbitrary element of $\tilde{x}_{n_1}$.
- For $k \geq 1$ let y_{k+1} be an element of $\tilde{x}_{n_{k+1}}$ such that $d(y_k, y_{k+1}) < \frac{1}{2^{k+1}}$. Such an y_{k+1} exists by (2.50).

Then (y_k) is a Cauchy sequence in $(\mathbb{X}, d)$. Indeed, for every $k \geq 1$ and for all $m \geq n \geq k$,

$$d(y_m, y_n) \leq \sum_{l=n}^{m-1} d(y_l, y_{l+1}) < \sum_{l=k}^{\infty} \frac{1}{2^{l+1}} = \frac{1}{2^k} .$$

Since $(\mathbb{X}, d)$ is complete, $\lim_{k\to\infty} d(y_k, x) = 0$ for some $x \in \mathbb{X}$. Let $\tilde{x} \in \tilde{\mathbb{X}}$ be the equivalence class of x. Since, by definition of $\tilde{d}$, $\tilde{d}(\tilde{x}_{n_k}, \tilde{x}) \leq d(y_k, x)$ for all $k \geq 1$, the subsequence $(\tilde{x}_{n_k})_k$ converges to $\tilde{x}$ in $(\tilde{\mathbb{X}}, \tilde{d})$. Finally, since $(\tilde{x}_n)_n$ is a Cauchy sequence, it follows easily that the whole sequence $(\tilde{x}_n)$ also converges to $\tilde{x}$, hence $(\tilde{\mathbb{X}}, \tilde{d})$ is complete. □

Proof of Lemma 2.3.1. In view of and (2.14) and Lemma 2.5.1 it only remains to show that $\tilde{d}$ is a metric, rather than just a pseudo-metric.

Assume that $\tilde{d}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = 0$ for some $\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}} \in \tilde{l}_0$. Then, for arbitrary $\boldsymbol{x} \in \tilde{\boldsymbol{x}}, \boldsymbol{y} \in \tilde{\boldsymbol{y}}$, there exists a sequence $(\boldsymbol{k}_n)_{n\in\mathbb{N}} \subseteq \mathbb{Z}^d$ such that $\|B^{\boldsymbol{k}_n}\boldsymbol{x} - \boldsymbol{y}\|_\infty \to 0$ as $n \to \infty$. It suffices to show that the sequence $|\boldsymbol{k}_n|_n$ is bounded. Indeed, in that case there exists a $\boldsymbol{k} \in \mathbb{Z}^d$ such that $\boldsymbol{k}_n = \boldsymbol{k}$ for infinitely many $n \in \mathbb{N}$ which implies that $\boldsymbol{y} = B^{\boldsymbol{k}}\boldsymbol{x}$, and hence that $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{y}}$.

Suppose now that the sequence $|\boldsymbol{k}_n|_n$ is unbounded and that $\boldsymbol{y} \neq \boldsymbol{0}$ (the case $\boldsymbol{y} = \boldsymbol{0}$ is straightforward). Since $\lim_{|\boldsymbol{i}|\to\infty} |y_{\boldsymbol{i}}| = 0$, we can find $\boldsymbol{i}_0 \in \mathbb{Z}^d$ and $N \in \mathbb{N}$ such that $|y_{\boldsymbol{i}_0}| = \|\boldsymbol{y}\|_\infty > 0$ and $|y_{\boldsymbol{i}}| < \|\boldsymbol{y}\|_\infty/4$ for all $|\boldsymbol{i}| \geq N$. By shifting $\boldsymbol{y}$ (and $\boldsymbol{k}_n$'s), we can assume that $\boldsymbol{i}_0 = \boldsymbol{0}$.

Take now an integer $n_0 > 0$ such that $\|B^{\boldsymbol{k}_n}\boldsymbol{x} - \boldsymbol{y}\|_\infty < \|\boldsymbol{y}\|_\infty/4$ for all $n \geq n_0$. Since $|\boldsymbol{k}_n|_n$ is unbounded we can also find an integer $n_1 \geq n_0$ such that $|\boldsymbol{k}_{n_1} - \boldsymbol{k}_{n_0}| \geq N$. It now follows that

$$\frac{3}{4}\|\boldsymbol{y}\|_\infty < |(B^{\boldsymbol{k}_{n_1}}\boldsymbol{x})_{\boldsymbol{0}}| = |x_{\boldsymbol{k}_{n_1}}| = |(B^{\boldsymbol{k}_{n_0}}\boldsymbol{x})_{\boldsymbol{k}_{n_1}-\boldsymbol{k}_{n_0}}| < \frac{1}{2}\|\boldsymbol{y}\|_\infty ,$$

which is a contradiction. Hence, the sequence $|k_n|_n$ is bounded. □

Convergence determining families for random measures on $\tilde{l}_{0,0}$

Recall the family $\mathcal{F}_0 \subseteq CB_b^+(\tilde{l}_{0,0})$ defined in Remark 2.3.13 and let $N, N_1, N_2, \dots$ be random measures on $\tilde{l}_{0,0}$.

Lemma 2.5.2. *If $\mathbb{E}[e^{-N_n(f)}] \to \mathbb{E}[e^{-N(f)}]$ for all $f \in \mathcal{F}_0$, then $N_n \xrightarrow{d} N$ in $\mathcal{M}(\tilde{l}_{0,0})$.*

Remark 2.5.3. Since $f \in \mathcal{F}_0$ implies that $\lambda f \in \mathcal{F}_0$ for all $\lambda \geq 0$, $\mathbb{E}[e^{-N_n(f)}] \to \mathbb{E}[e^{-N(f)}]$ for all $f \in \mathcal{F}_0$ is equivalent to $N_n(f) \xrightarrow{d} N(f)$ in $\mathbb{R}$ for all $f \in \mathcal{F}_0$.

Remark 2.5.4. Essentially the same proof as the one given below shows that for random measures $N', N'_1, N'_2, \dots$ on $[0,1]^d \times \tilde{l}_{0,0}$, $\mathbb{E}[e^{-N'_n(f)}] \to \mathbb{E}[e^{-N'(f)}]$ for all $f \in \mathcal{F}'_0$, where $\mathcal{F}'_0 \subseteq CB_b^+([0,1]^d \times \tilde{l}_{0,0})$ is defined in the beginning of Section 2.3.6 implies that $N'_n \xrightarrow{d} N'$ in $\mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$.

Proof. Recall that a Borel set $B \subseteq \tilde{l}_{0,0}$ is bounded if for some $\epsilon > 0$, $\tilde{d}(\mathbf{0}, \boldsymbol{x}) = \|\boldsymbol{x}\|_\infty > \epsilon$ for all $\boldsymbol{x} \in B$. By Proposition 1.3.1 (see Example 1.3.3), to show that $N_n \xrightarrow{d} N$ it suffices to prove that $\mathbb{E}[e^{-N_n(g)}] \to \mathbb{E}[e^{-N(g)}]$ for all $g \in LB_b^+(\tilde{l}_{0,0}, \tilde{d})$.

For $\delta > 0$ let $\phi^\delta : [0, \infty) \to [0,1]$ be a (uniformly continuous) function defined by (i) $\phi^\delta(x) = 0$ for $x \leq \delta$; (ii) $\phi^\delta(x) = 1$ for $x \geq 2\delta$; (iii) $\phi^\delta(x) = x/\delta - 1$ for $x \in [\delta, 2\delta]$. Further, for any $\boldsymbol{x} \in \tilde{l}_0$, by slight abuse of notation, denote by $\phi^\delta(\boldsymbol{x})$ the equivalence class of the sequence $(x_{\boldsymbol{i}}\phi^\delta(|x_{\boldsymbol{i}}|))_{\boldsymbol{i}}$, where $(x_{\boldsymbol{i}})_{\boldsymbol{i}} \in l_0$ is an arbitrary representative of $\boldsymbol{x}$. Note, for every $\boldsymbol{x} \in \tilde{l}_{0,0}$, $\tilde{d}(\boldsymbol{x}, \phi^\delta(\boldsymbol{x})) \leq \delta$.

Take now an arbitrary $g \in LB_b^+(\tilde{l}_{0,0}, \tilde{d})$ and let $\epsilon > 0$ be such that support of f is contained in $B := \{\boldsymbol{x} \in \tilde{l}_{0,0} : \|\boldsymbol{x}\|_\infty \geq \epsilon\} \in \mathcal{B}_b(\tilde{l}_{0,0})$. Assume without loss of generality that $|g(\boldsymbol{x}) - g(\boldsymbol{y})| \leq \tilde{d}(\boldsymbol{x}, \boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \tilde{l}_{0,0}$. For each $\delta > 0$ define a function g^δ on $\tilde{l}_{0,0}$ by $g^\delta(\boldsymbol{x}) = g(\phi^\delta(\boldsymbol{x}))$ with convention that $g(\mathbf{0}) = 0$. By construction, g^δ is an element of $\mathcal{F}_0$ and moreover support of g^δ is also contained in B for each $\delta > 0$. Further, since g is Lipschitz, for all $\delta > 0$ and all $\boldsymbol{x} \in \tilde{l}_{0,0}$

$$|g(\boldsymbol{x}) - g^\delta(\boldsymbol{x})| \leq \tilde{d}(\boldsymbol{x}, \phi^\delta(\boldsymbol{x}))\mathbb{1}_B(\boldsymbol{x}) \leq \delta \mathbb{1}_B(\boldsymbol{x})\,. \tag{2.51}$$

Observe now that for all $\delta > 0$, since $g^\delta \in \mathcal{F}_0$, convergence $\mathbb{E}[e^{-N_n(g^\delta)}] \to \mathbb{E}[e^{-N(g^\delta)}]$ holds and hence

$$\limsup_{n\to\infty} \left|\mathbb{E}[e^{-N_n(g)}] - \mathbb{E}[e^{-N(g)}]\right| \leq \limsup_{n\to\infty} \left|\mathbb{E}[e^{-N_n(g)}] - \mathbb{E}[e^{-N_n(g^\delta)}]\right| + \left|\mathbb{E}[e^{-N(g^\delta)}] - \mathbb{E}[e^{-N(g)}]\right|. \tag{2.52}$$

Using (2.51) and the simple bound $|e^{-x} - e^{-y}| \leq |x - y| \wedge 1$, $x, y \geq 0$, we obtain that for any random measure M in $\mathcal{M}(\tilde{l}_{0,0})$

$$\left|\mathbb{E}[e^{-M(g)}] - \mathbb{E}[e^{-M(g^\delta)}]\right| \leq \mathbb{E}[|M(g) - M(g^\delta)| \wedge 1] \leq \mathbb{P}(M(B) > C) + \delta C\,,$$

for all $\delta > 0$ and $C > 0$. Using this bound for N_n and N, (2.52) yields that for all $\delta > 0$ and $C > 0$

$$\limsup_{n\to\infty} \left|\mathbb{E}[e^{-N_n(g)}] - \mathbb{E}[e^{-N(g)}]\right| \leq \limsup_{n\to\infty} \mathbb{P}(N_n(B) > C) + \mathbb{P}(N(B) > C) + 2\delta C\,. \tag{2.53}$$

It remains to notice that, since one can find a function $g_0 \in LB_b^+(\tilde{l}_{0,0}, \tilde{d})$ such that $g_0 \geq \mathbb{1}_B$ (see (1.6)), there exist a function $f_0 \in \mathcal{F}_0$ such that $f_0 \geq \mathbb{1}_B$ (with notation as above, simply take $f_0 = 2g_0^{1/2}$). Since by assumption $N_n(f_0) \xrightarrow{d} N(f_0)$ in $\mathbb{R}$, the sequence $(N_n(f))_n$, and therefore the sequence $(N_n(B))_n$, is tight. Thus, letting $\delta \to 0$ and then $C \to \infty$ in (2.53) yields that $\left|\mathbb{E}[e^{-N_n(g)}] - \mathbb{E}[e^{-N(g)}]\right| \to 0$ as $n \to \infty$. Since $g \in LB_b^+(\tilde{l}_{0,0}, \tilde{d})$ was arbitrary, this proves the claim. □

Proof of Corollary 2.3.15

Let $P : \mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0}) \to \mathcal{M}_p([0,1]^d \times (\mathbb{R} \setminus \{0\}))$ be the (projection) mapping given by

$$P\Big(\sum_i \delta_{(\boldsymbol{t}_i,(\boldsymbol{x}_j^i)_j)}\Big) = \sum_i \sum_j \delta_{(\boldsymbol{t}_i,\boldsymbol{x}_j^i)}\,.$$

Note that P is well–defined (i.e. for every $\mu \in \mathcal{M}_p([0,1]^d \times \tilde{l}_{0,0})$, $P(\mu)$ is a locally finite point measure on $[0,1]^d \times (\mathbb{R} \setminus \{0\})$ and moreover, using Proposition 1.2.8 it easy to show that P is continuous.

Recall, for each $\boldsymbol{i} \in I_n$ we denote $J_{n,\boldsymbol{i}} = (\boldsymbol{j} \in \mathbb{Z}^d : (\boldsymbol{i} - \boldsymbol{1}) \cdot r_n + \boldsymbol{1} \leq \boldsymbol{j} \leq \boldsymbol{i} \cdot r_n)$ and $\boldsymbol{X}_{n,\boldsymbol{i}} = \boldsymbol{X}_{J_{n,\boldsymbol{i}}}$. Therefore, (2.42) and an application of the continuous mapping theorem (see e.g. [Bil68, Corollary 1, p. 31]) imply that

$$\sum_{\boldsymbol{i}\in I_n} \sum_{\boldsymbol{j}\in J_{n,\boldsymbol{i}}} \delta_{(\boldsymbol{i}/k_n, X_{\boldsymbol{j}}/a_n)} = P(N_n') \xrightarrow{d} P(N') = \sum_{i\in\mathbb{N}} \sum_{\boldsymbol{j}\in\mathbb{Z}^d} \delta_{(\boldsymbol{T}_i, P_i Q_{\boldsymbol{j}}^i)} \tag{2.54}$$

in $\mathcal{M}_p([0,1]^d \times (\mathbb{R} \setminus \{0\}))$.

Define a metric ρ on $[0,1]^d \times \mathbb{R}$ by $\rho((\boldsymbol{t}, x), (\boldsymbol{s}, y)) = |\boldsymbol{t} - \boldsymbol{s}| \vee |x - y|$, where for $\boldsymbol{t} = (t_1, \ldots, t_d) \in \mathbb{R}^d$, $|\boldsymbol{t}| = \max_{i=1,\ldots,d} |t_i|$. By Proposition 1.3.1, the family $LB_b^+([0,1]^d \times (\mathbb{R} \setminus \{0\}), \rho)$ is (point process) convergence determining and hence, in view of (2.54), it suffices to show that

$$\left|\mathbb{E}\Big[e^{-\sum_{\boldsymbol{i}\in I_n}\sum_{\boldsymbol{j}\in J_{n,\boldsymbol{i}}} f(\boldsymbol{i}/k_n, X_{\boldsymbol{j}}/a_n)}\Big] - \mathbb{E}\Big[e^{-\sum_{\boldsymbol{j}\in J_n} f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)}\Big]\right| \to 0\,, \tag{2.55}$$

for all $f \in LB_b^+([0,1]^d \times (\mathbb{R} \setminus \{0\}), \rho)$. For that purpose, we use similar arguments as in [Kri10, Proposition 1.34].

Assume for simplicity that $d = 2$. Let $f \in LB_b^+([0,1]^2 \times (\mathbb{R} \setminus \{0\}), \rho)$ be arbitrary and take an $\epsilon > 0$ such that $|x| \leq \epsilon$ implies that $f(\boldsymbol{t}, x) = 0$. Recall, $k_n = \lfloor n/r_n \rfloor$ and $I_n = \{1, \ldots, k_n\}^2$. Furthermore, blocks $J_{n,\boldsymbol{i}} \subseteq J_n = \{1, \ldots, n\}^2$, $\boldsymbol{i} \in I_n$, are disjoint and each has r_n^2 indices. Hence, the set of indices $J_n' = \cup_{\boldsymbol{i}\in I_n} J_{n,\boldsymbol{i}}$ is contained in J_n and furthermore, J_n has at most $2r_n n$ indices which are not in J_n'. Therefore, by the properties of the function f and stationarity of $(X_{\boldsymbol{j}})$,

$$\Big|\mathbb{E}\Big[e^{-\sum_{\boldsymbol{j}\in J_n'} f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)}\Big] - \mathbb{E}\Big[e^{-\sum_{\boldsymbol{j}\in J_n} f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)}\Big]\Big| \leq 2r_n n\mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon) \to 0 \tag{2.56}$$

since by (2.16) and regular variation of $|X_{\mathbf{0}}|$, $\mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon) \sim \epsilon^{-\alpha} n^{-2}$ and since $r_n/n \to 0$. Furthermore, using the simple inequality $|e^{-x} - e^{-y}| \leq |x - y|$ valid for all $x, y \geq 0$ yields that

$$\begin{aligned}\Big|\mathbb{E}\Big[e^{-\sum_{\boldsymbol{i}\in I_n}\sum_{\boldsymbol{j}\in J_{n,\boldsymbol{i}}} f(\boldsymbol{i}/k_n, X_{\boldsymbol{j}}/a_n)}\Big] - \mathbb{E}\Big[e^{-\sum_{\boldsymbol{j}\in J_n'} f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)}\Big]\Big| \\ \leq \sum_{\boldsymbol{i}\in I_n}\sum_{\boldsymbol{j}\in J_{n,\boldsymbol{i}}} \mathbb{E}\Big|f(\boldsymbol{i}/k_n, X_{\boldsymbol{j}}/a_n) - f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)\Big| .\end{aligned}$$

Since for some $L > 0$, $|f(\boldsymbol{t}, x) - f(\boldsymbol{s}, y)| \leq L(|\boldsymbol{t} - \boldsymbol{s}| \vee |x - y|)$ for all $\boldsymbol{t}, \boldsymbol{s} \in [0, 1]^d$ and $x, y \in \mathbb{R} \setminus \{0\}$, and since $|x| \leq \epsilon$ implies that $f(\boldsymbol{t}, x) = 0$,

$$\mathbb{E}\Big|f(\boldsymbol{i}/k_n, X_{\boldsymbol{j}}/a_n) - f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)\Big| \leq L|\boldsymbol{i}/k_n - \boldsymbol{j}/n| \cdot \mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon) ,$$

for all $\boldsymbol{i} \in I_n$ and $\boldsymbol{j} \in J_{n,\boldsymbol{i}}$. Now since $|\boldsymbol{i}r_n - \boldsymbol{j}| \leq r_n$ for all such $\boldsymbol{i}$ and $\boldsymbol{j}$ and since $k_n = \lfloor n/r_n \rfloor$ one can show that

$$|\boldsymbol{i}/k_n - \boldsymbol{j}/n| \leq \frac{2r_n}{n} ,$$

for all $\boldsymbol{i} \in I_n$ and $\boldsymbol{j} \in J_{n,\boldsymbol{i}}$. This implies that

$$\Big|\mathbb{E}\Big[e^{-\sum_{\boldsymbol{i}\in I_n}\sum_{\boldsymbol{j}\in J_{n,\boldsymbol{i}}} f(\boldsymbol{i}/k_n, X_{\boldsymbol{j}}/a_n)}\Big] - \mathbb{E}\Big[e^{-\sum_{\boldsymbol{j}\in J_n'} f(\boldsymbol{j}/n, X_{\boldsymbol{j}}/a_n)}\Big]\Big| \leq k_n^2 r_n^2 L \frac{2r_n}{n}\mathbb{P}(|X_{\mathbf{0}}| > a_n\epsilon) .$$

Note that $k_n r_n \sim n$, so again by (2.16) and since $r_n/n \to 0$, the right hand side above tends to 0 as $n \to \infty$. In view of (2.56), this proves (2.55) and finishes the proof.

Chapter 3

An invariance principle for sums and record times of regularly varying stationary time series

3.1 Introduction

In this chapter, we put to use the order preserving point process convergence theory developed in Chapter 2. In particular, we study partial sums and record times of a stationary regularly varying $\mathbb{R}$–valued time series. The running assumption of the chapter will be that the conclusion of Theorem 2.3.14, i.e. point process convergence (2.42), holds. In particular, this includes infinite order moving average processes from Examples 2.3.7 and 2.4.5. In both problems, the information about the temporal ordering of extreme observations within the same cluster will be essential. In fact, the study of record times in a dependent sequence was the main motivation for development of results in Chapter 2.

The rest of the chapter is divided in two sections on record times and partial sums, respectively, and in the beginning of each we briefly motivate the corresponding problem and state our main results.

The results of this chapter are based on the paper [BPS18].

3.2 Record times

3.2.1 Introduction

In this section we study record times in a stationary regularly varying time series $(X_i)_{i\in\mathbb{N}}$. For convenience we restrict our attention to the case when X_i's are nonnegative and say that $i \in \mathbb{N}$ is a *record time* and X_i the corresponding *record (value)* if

$$X_i > \max_{j=1,\dots,i-1} X_j \,,$$

with $\max \emptyset = 0$. It is well known that, if X_i's are i.i.d. regularly varying,

$$R_n = \sum_{j=1}^{\infty} \delta_{j/n} \mathbb{1}_{\{X_j \text{ is a record}\}} \xrightarrow{d} R = \sum_{i\in\mathbb{Z}} \delta_{\tau_i} \,, \tag{3.1}$$

as $n \to \infty$ in $\mathcal{M}_p((0,\infty))$ (with bounded sets being those contained in $[\epsilon, \epsilon^{-1}]$, $\epsilon > 0$), where $\sum_{i\in\mathbb{Z}} \delta_{\tau_i}$ is a Poisson process on $(0,\infty)$ with intensity measure $\mu(dx) = x^{-1}dx$, i.e. $\mu((a,b]) = \log(b/a)$ for $0 < a < b$; see [Res87, Corollary 4.22]. In particular, if T_n is the last record time among $X_1, \dots, X_n$, (3.1) yields that, for $x \in [0,1]$,

$$\mathbb{P}(T_n/n \le x) = \mathbb{P}(R_n((x,1]) = 0) \to \mathbb{P}(R((x,1]) = 0) = e^{-\mu((x,1])} = x \,,$$

i.e. the distribution of T_n/n is asymptotically uniform on $[0,1]$. For further implications of (3.1) see [Res87, pp. 219–220]. Note, the limiting process R is *scale–invariant* in the sense that, for any $c > 0$, $\sum_{i\in\mathbb{Z}} \delta_{c\tau_i} \stackrel{d}{=} \sum_{i\in\mathbb{Z}} \delta_{\tau_i}$, see [Arr98] for an interesting discussion on such processes.

In this section, we extend convergence in (3.1) to stationary regularly varying time series. Under appropriate assumptions, we show that the process of record times R_n converges in distribution to a rather simple scale invariant *compound* Poisson process, see Theorem 3.2.3 for details. Note, since record times remain unaltered after a strictly increasing transformation, the main result below holds for stationary sequences with a general marginal distribution as long as they can be monotonically transformed into a regularly varying sequence.

3.2.2 Continuity of the record times functional

We start by introducing the notion of records for sequences in $\tilde{l}_0$. For $y \ge 0$ and $\boldsymbol{x} = (x_j)_{j\in\mathbb{Z}} \in \tilde{l}_0$ define

$$R^{\boldsymbol{x}}(y) = \sum_{j=-\infty}^{\infty} \mathbb{1}_{\{x_j > y \vee \sup_{i<j} x_i\}} \,,$$

where $a \vee b := \max\{a, b\}$, representing the number of records in the sequence $\boldsymbol{x}$ larger than y. Observe, this number is finite for every $\boldsymbol{x} \in \tilde{l}_0$ and every $y > 0$ since at most finite number of coordinates of $\boldsymbol{x}$ are larger than y. Note, $\boldsymbol{x} = (x_j)_{j\in\mathbb{Z}} \in \tilde{l}_0$ means that $\boldsymbol{x}$ is an element of $\tilde{l}_0$ and (x_j) is an arbitrary representative of $\boldsymbol{x}$. Write also $\|\boldsymbol{x}\|_\infty^+ = \sup_j x_j \vee 0$ for $\boldsymbol{x} = (x_j)_{j\in\mathbb{Z}} \in \tilde{l}_0$.

Let $\gamma = \sum_{i=1}^{\infty} \delta_{t_i, \boldsymbol{x}^i} \in \mathcal{M}_p([0,\infty) \times \tilde{l}_{0,0})$, where $\boldsymbol{x}^i = (x^i_j)_{j\in\mathbb{Z}} \in \tilde{l}_0$, $i \in \mathbb{N}$. Bounded sets in $[0,\infty) \times \tilde{l}_{0,0}$ are those contained in sets $[0, \epsilon^{-1}] \times \{\boldsymbol{x} \in \tilde{l}_{0,0} : \|\boldsymbol{x}\|_\infty > \epsilon\}$, $\epsilon > 0$. Define, for $t > 0$,

$$M^\gamma(t) = \sup_{t_i \le t} \|\boldsymbol{x}^i\|_\infty^+ \,, \quad M^\gamma(t-) = \sup_{t_i < t} \|\boldsymbol{x}^i\|_\infty^+ \,,$$

where we set $\sup \emptyset = 0$ for convenience. Next, let R_γ be the (counting) point process on $(0,\infty)$ defined by

$$R_\gamma = \sum_{i=1}^\infty \delta_{t_i} R^{\boldsymbol{x}^i}(M^\gamma(t_i-)) \,,$$

hence for arbitrary $0 < a < b$,

$$R_\gamma(a,b] = \sum_{a < t_i \le b} \sum_{j=-\infty}^\infty \mathbb{1}_{\{x_j^i > M^\gamma(t_i-) \vee \sup_{k<j} x_k^i\}} \,.$$

If all t_i's are mutually different (this will always hold in our case), we say that γ has $k \ge 1$ records at time $t \in [0,\infty)$ if $t \in \{t_i\}$ and $R^{\boldsymbol{x}^i}(M^\gamma(t_i-)) = k$. Observe, whenever γ satisfies $M^\gamma(t) > 0$ for all $t > 0$, $R_\gamma(a,b] < \infty$ for all $0 < a < b$, i.e. R_γ is a locally finite measure on $(0,\infty)$ with bounded sets being those contained in $(\epsilon, \epsilon^{-1})$, $\epsilon > 0$. Let $\mathcal{A} \subseteq \mathcal{M}_p([0,\infty) \times \tilde{l}_{0,0})$ denote the set of all such γ's. Note that $\mathcal{A}$ is an open subset of $\mathcal{M}_p([0,\infty) \times \tilde{l}_{0,0})$ with respect to the vague topology.

Lemma 3.2.1. *The mapping $\gamma \mapsto R_\gamma$ from $\mathcal{A}$ to $\mathcal{M}_p((0,\infty))$ is continuous at every $\gamma = \sum_{i=1}^\infty \delta_{t_i, \boldsymbol{x}^i} \in \mathcal{A}$ such that*

(i) $t_i \neq t_j$ for all $i \neq j \in \mathbb{N}$;

(ii) $x_j^i \neq x_{j'}^{i'}$ for all $i, i' \in \mathbb{N}$ and all $j \neq j' \in \mathbb{Z}$ such that $x_j^i, x_{j'}^{i'} > 0$.

Proof. Fix an arbitrary $\gamma = \sum_{i=1}^\infty \delta_{t_i, \boldsymbol{x}^i} \in \mathcal{A}$ satisfying the above assumptions, and assume that $(\gamma_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{A}$ satisfying $\gamma_n \xrightarrow{v} \gamma$. We must prove that $R_{\gamma_n} \xrightarrow{v} R_\gamma$ in $\mathcal{M}_p((0,\infty))$. By [Kal17, Lemma 4.1(iii)], it is sufficient to show that for all $0 < a < b \in \{t_i\}^c$

$$R_{\gamma_n}(a,b] \to R_\gamma(a,b] \,,$$

as $n \to \infty$. Fix arbitrary $0 < a < b \in \{t_i\}^c$. First, let $t_{i_0} \in [0,a)$ be such that $\|\boldsymbol{x}^{i_0}\|_\infty^+ = M^\gamma(a) > 0$. Further, there are finitely many time instances in $(a,b]$, say $t_{i_1}, \dots, t_{i_k}$, such that $\|\boldsymbol{x}^{i_l}\|_\infty^+ > M^\gamma(a) > 0$ for all $l = 1, \dots, k$. In particular,

$$R_\gamma(a,b] = \sum_{t_i \in (a,b]} R^{\boldsymbol{x}^i}(M^\gamma(t_i-)) = \sum_{l=1}^k R^{\boldsymbol{x}^{i_l}}(M^\gamma(t_{i_l}-)) \,. \tag{3.2}$$

Observe, by (ii) above there does not exist $t_i \in [0,b]$, $i \neq i_0$, with $\|\boldsymbol{x}^i\|_\infty^+ = M^\gamma(a)$, in particular, for some $\epsilon > 0$ small enough t_{i_0} is the only time instance in $[0,b]$ such that

$0 < M^\gamma(a) - \epsilon \le \|\boldsymbol{x}^i\|_\infty^+ \le M^\gamma(a)$. By Proposition 1.2.8, $\gamma_n \xrightarrow{v} \gamma$ implies that for all $\gamma_n = \sum_{i=1}^\infty \delta_{t_i^n, \boldsymbol{x}^{n,i}}$ with n large enough, there exists

- exactly 1 time instance $t_{i_0}^n \in [0, a)$ such that $\|\boldsymbol{x}^{n,i_0}\|_\infty^+ \ge M^\gamma(a) - \epsilon$;
- exactly k time instances $t_{i_1}^n, \dots, t_{i_k}^n \in (a, b]$ such that $\|\boldsymbol{x}^{n,i_l}\|_\infty^+ \ge M^\gamma(a) - \epsilon$ for all $l = 1, \dots, k$.

Moreover, they satisfy $\boldsymbol{x}^{n,i_l} \to \boldsymbol{x}^{i_l}$ and $t_{i_l}^n \to t_{i_l}$ for $l = 0, \dots k$ as $n \to \infty$. In particular, $M^{\gamma_n}(a) = \|\boldsymbol{x}^{n,i_0}\|_\infty^+ \to \|\boldsymbol{x}^{i_0}\|_\infty^+ = M^\gamma(a)$ as $n \to \infty$, and for all n large enough,

$$R_{\gamma_n}(a, b] = \sum_{t_i^n \in (a,b]} R^{\boldsymbol{x}^{n,i}}(M^{\gamma_n}(t_i^n -)) = \sum_{l=1}^k R^{\boldsymbol{x}^{n,i_l}}(M^{\gamma_n}(t_{i_l}^n -)) \,. \tag{3.3}$$

Assume that, as $n \to \infty$, $y_n \to y > 0$ and $\boldsymbol{x}^n \to \boldsymbol{x} = (x_j)_{j \in \mathbb{Z}} \in \tilde{l}_0$ where all positive x_j's are pairwise distinct and $x_j \neq y$ for all $j \in \mathbb{Z}$. It is straightforward to check that then, for all n large enough,

$$R^{\boldsymbol{x}^n}(y_n) = R^{\boldsymbol{x}}(y) \,.$$

Observe further that for the choice of t_{i_l}, $t_{i_l}^n$ we made above, it holds that $M^{\gamma_n}(t_{i_l}^n -) \to M^\gamma(t_{i_l} -)$ for all $l = 1, \dots, k$, since t_{i_l}'s are mutually different. Together with assumption (ii), (3.2) and (3.3) this yields that, for all n large enough,

$$R_{\gamma_n}(a, b] = R_\gamma(a, b] \,.$$

□

3.2.3 Limiting result

Assume that $(X_i)_{i \in \mathbb{Z}}$ is a nonnegative and stationary regularly varying time series (for general $\mathbb{R}$–valued (X_i) consider (X_i^+)). Let $(r_n)_n$ be such that $r_n \to \infty, r_n/n \to 0$ as $n \to \infty$, and for each $n \in \mathbb{N}$ define blocks

$$\boldsymbol{X}_{n,i} = (X_{(i-1)r_n+1}, \dots, X_{ir_n}) \,,\ i \in \mathbb{N} \,.$$

We assume that the conclusion of Theorem 2.3.14 holds. More precisely, that for $(a_n)_n$ such that $n\mathbb{P}(X_0 > a_n) \to 1$,

$$N_n' = \sum_{i=1}^\infty \delta_{(i/k_n, \boldsymbol{X}_{n,i}/a_n)} \xrightarrow{d} N' = \sum_{i=1}^\infty \delta_{(T_i, P_i \boldsymbol{Q}^i)} \,, \tag{3.4}$$

in $\mathcal{M}_p([0, \infty) \times \tilde{l}_{0,0})$, where $\sum_{i=1}^\infty \delta_{(T_i, P_i)}$ is a Poisson process in $\mathcal{M}_p([0, \infty) \times (0, \infty))$ with intensity measure $\vartheta Leb \times d(-y^{-\alpha})$, and $\boldsymbol{Q}^i = (Q_j^i)_{j \in \mathbb{Z}}$, $i \ge 1$, i.i.d. elements of $\tilde{l}_0$, distributed as $\boldsymbol{Q} = (Q_j)_{j\in\mathbb{Z}}$ and independent of $\sum_{i=1}^{\infty}\delta_{(T_i,P_i)}$. Here we implicitly assume that the tail process of (X_i) is in l_0 a.s., so in particular, the corresponding ϑ and $\boldsymbol{Q}$ from Section 2.3.3 are well defined. Note, since X_i's are assumed nonnegative, Q_j's are $[0,1]$–valued with at least one of them being equal to 1. This will be important in the proof of the main theorem below.

Remark 3.2.2. Observe, here we consider the whole time axis $[0,\infty)$, while in Theorem 2.3.14 we restricted to the segment $[0,1]$. Still, by minor modifications of assumptions, results of Chapter 2 are easily extended to the time interval $[0,T]$ for arbitrary $T\in\mathbb{N}$, and hence to whole $[0,\infty)$. In particular, (3.4) holds for m–dependent time series, and also for infinite order linear processes from Examples 2.3.7 and 2.4.5.

We will also need the point process N_n on $[0,\infty)\times(0,\infty)$ defined by

$$N_n = \sum_{j=1}^{\infty}\delta_{(j/n, X_j/a_n)}\,.$$

By identifying $x>0$ with a sequence with exactly one nonzero coordinate equal to x, in the sequel we treat N_n as a process on the space $[0,\infty)\times\tilde{l}_{0,0}$. Observe,

$$R_{N_n} = \sum_{j=1}^{\infty}\delta_{j/n}\mathbb{1}_{\{X_j \text{ is a record}\}}\,, n\in\mathbb{N}\,.$$

Recall, we say that random variable ζ is Pareto distributed with tail index $\alpha>0$ if $\mathbb{P}(\zeta\geq y) = y^{-\alpha}$ for $y\geq 1$.

Theorem 3.2.3. *Let $(X_i)_{i\in\mathbb{Z}}$ be a nonnegative stationary regularly varying sequence with tail index $\alpha>0$. Assume that the convergence in (3.4) holds and moreover that*

$$\mathbb{P}(\textit{all nonzero } Q_j\textit{'s are mutually different}) = 1\,.$$

Then

$$R_{N_n}\xrightarrow{d} R_{N_n'}\,,$$

as $n\to\infty$ in $\mathcal{M}_p((0,\infty))$. Moreover, the limiting process is a compound Poisson process with representation

$$R_{N'} \stackrel{d}{=} \sum_{i\in\mathbb{Z}}\delta_{\tau_i}\kappa_i\,,$$

where

(i) $\sum_{i\in\mathbb{Z}}\delta_{\tau_i}$ is a Poisson point process on $(0,\infty)$ with intensity measure $x^{-1}\mathrm{d}x$;

(ii) $(\kappa_i)_{i\in\mathbb{Z}}$ is a sequence of i.i.d. random variables independent of $\sum_{i\in\mathbb{Z}}\delta_{\tau_i}$ with the same distribution as the integer–valued random variable $R^{\boldsymbol{Q}}(1/\zeta)$ where ζ is a Pareto random variable with tail index α, independent of $\boldsymbol{Q}$.

Example 3.2.4. For an illustration of the previous theorem, consider the moving average process of order 1

$$X_i = \xi_i + c\xi_{i-1}\,,\; i \in \mathbb{Z}\,,$$

for a sequence of i.i.d. nonnegative regularly varying random variables $(\xi_i : i \in \mathbb{Z})$ with tail index $\alpha > 0$. Assume further that $c > 1$. By (2.29), the corresponding random element $\boldsymbol{Q}$ in $\tilde{l}_0$ equals the deterministic sequence $(\ldots, 0, 1/c, 1, 0, \ldots)$. Intuitively speaking, in each cluster of extremely large values, there are exactly two successive extreme values with the second one c times larger that the first. Therefore, each such cluster can give rise to at most 2 records, see Figure 3.1. By straightforward calculations, the random variables κ_i from Theorem 3.2.3 have the following distribution

$$\mathbb{P}(\kappa_i = 2) = \mathbb{P}(1/\zeta \leq 1/c) = \mathbb{P}(\zeta \geq c) = \frac{1}{c^\alpha} = 1 - \mathbb{P}(\kappa_i = 1)\;.$$

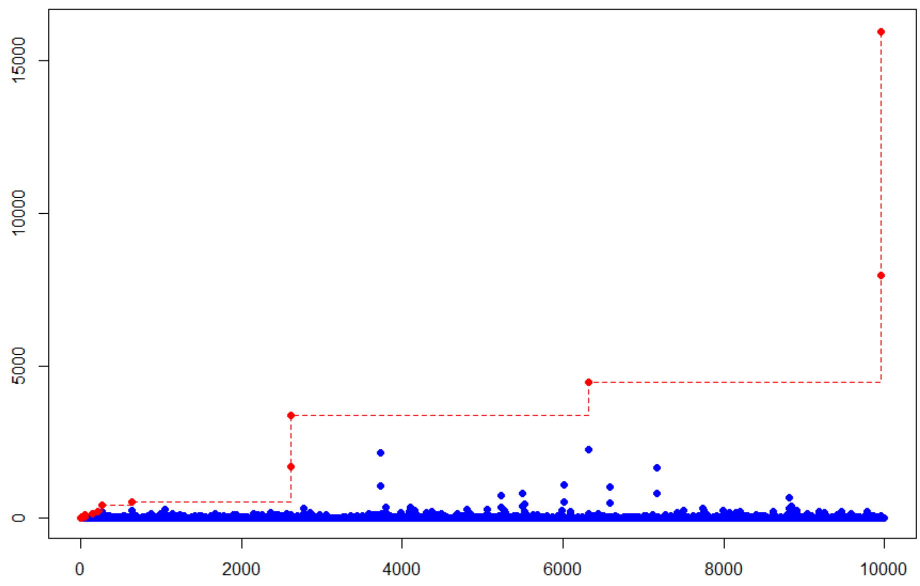

Figure 3.1: Simulation of 10000 observations of the process (X_i) from Example 3.2.4 with $c = 2$ and $\alpha = 1.2$. The dashed line represents the running maxima $M_n = \max_{i=1,\ldots,n} X_i$ of the process and red dots correspond to record X_i's.

Proof of Theorem 3.2.3. Step 1. We first show that N' almost surely satisfies all of the assumptions of Lemma 3.2.1. Observe that $\mathbb{P}(N'([0,\epsilon] \times \tilde{l}_{0,0}) = \infty) = 1$ for all $\epsilon > 0$, so in particular $\mathbb{P}(N \in \mathcal{A}) = 1$. Further, for each $\epsilon > 0$, N' restricted to the set

$[0, \epsilon^{-1}] \times \{\boldsymbol{x} \in \tilde{l}_{0,0} : \|\boldsymbol{x}\|_\infty > \epsilon\}$, in distribution equals

$$\sum_{i=1}^{T} \delta_{(\epsilon^{-1}U_i, \epsilon V_i (Q_j^i)_j)} ,$$

where $(U_i)_i$ are i.i.d. uniform on $[0,1]$, $(V_i)_i$ are i.i.d. Pareto with index α, T is Poisson with intensity $\vartheta\epsilon^{-1-\alpha}$, and all of involved random elements are independent; see e.g. [LP17, Proposition 3.8] or [Res87, pp. 132–134]. In particular, together with the additional assumption on the Q_j's, this restricted N' satisfies all of the assumptions of Lemma 3.2.1 (simply use independence and the fact that U_i's and V_i's are continuous random variables). Letting $\epsilon \to 0$ shows that N' also satisfies all of these assumptions.

Hence, applying the continuous mapping theorem to (3.4) yields that

$$R_{N_n'} \xrightarrow{d} R_{N'} . \tag{3.5}$$

Step 2. Set $J_{n,i} = \{(i-1)r_n + 1, \dots, ir_n\}$ for $i \in \mathbb{N}$, so $\boldsymbol{X}_{n,i} = (X_j : j \in J_{n,i})$. Observe, for any nonnegative and measurable function f on $(0,\infty)$,

$$R_{N_n'}(f) = \sum_{i=1}^{\infty} f(i/k_n) \cdot R^{\boldsymbol{X}_{n,i}/a_n}(M^{N_n'}(\tfrac{i}{k_n}-)) ,$$

where for each $i \in \mathbb{N}$,

$$R^{\boldsymbol{X}_{n,i}/a_n}(M^{N_n'}(\tfrac{i}{k_n}-)) = \sum_{j \in J_{n,i}} \mathbb{1}_{\{X_j \text{ is a record}\}} .$$

Also,

$$R_{N_n}(f) = \sum_{j=1}^{\infty} f(j/n) \mathbb{1}_{\{X_j \text{ is a record}\}} = \sum_{j=1}^{\infty} \sum_{j \in J_{n,i}} f(j/n) \mathbb{1}_{\{X_j \text{ is a record}\}} .$$

Hence, each record time j/n of the process N_n appears at slightly altered time i/k_n in the process N_n' where $i = \lfloor j/r_n \rfloor + 1$ (i.e. such that $j \in J_{n,i}$). However, asymptotically the record times are very close since $r_n/n \to 0$.

Indeed, take $f \in LB_b^+(0,\infty)$ and let $0 < a < b$ be such that the support of f is contained in $(a, b]$. Without loss of generality assume that $f \le 1$ and that $|f(t) - f(s)| \le |t - s|$ for all $t, s \in (0,\infty)$. By the simple inequality $|e^{-x} - e^{-y}| \le |x - y| \wedge 1$ valid for all $x, y \ge 0$,

$$\left| \mathbb{E}[e^{-R_{N_n'}(f)}] - \mathbb{E}[e^{-R_{N_n}(f)}] \right| \le \mathbb{E}|R_{N_n'}(f) - R_{N_n}(f)| \wedge 1 . \tag{3.6}$$

Observe,

$$|R_{N_n'}(f) - R_{N_n}(f)| \le \sum_{i=1}^{\infty} \sum_{j \in J_{n,i}} |f(i/k_n) - f(j/n)| \mathbb{1}_{\{X_j \text{ is a record}\}} . \tag{3.7}$$

Recall that $k_n = \lfloor n/r_n \rfloor$, so for all $i \in \mathbb{N}, j \in J_{n,i}$, $j/n \le i/k_n$ and simple computation (cf. 3.35) yields that $|i/k_n - j/n| = (i/k_n - ir_n/n) + (ir_n/n - j/n) \le (i/k_n + 1)r_n/n$ which tends to zero as $n \to \infty$ uniformly over $i \le k_n T$ for every fixed $T > 0$ since $r_n/n \to 0$. In particular, for n large enough, $j/n \in (a, b]$ implies that the corresponding $i = \lfloor j/r_n \rfloor + 1$ satisfies $i/k_n \in (a, b+1]$ (1 is arbitrary here). Since $f(t) = 0$ for all $t \notin (a, b]$, one can therefore restrict the summation in (3.7) to i's such that $i/k_n \in (a, b+1]$ and by the Lipschitz property of f,

$$|R_{N_n'}(f) - R_{N_n}(f)| \le (b+2) \frac{r_n}{n} R_{N_n'}((a, b+1]) .$$

Since $R_{N_n'}((a, b+1]) \xrightarrow{d} R_{N'}((a, b+1])$ (assume w.l.o.g. that $\mathbb{P}(R_{N_n'}(\{a, b+1\}) = 0) = 1$), this shows that $|R_{N_n'}(f) - R_{N_n}(f)| \to 0$ in probability, and by (3.6) this further implies that

$$\left| \mathbb{E}[e^{-R_{N_n'}(f)}] - \mathbb{E}[e^{-R_{N_n}(f)}] \right| \to 0 .$$

Since $f \in LB_b^+(0, \infty)$ was arbitrary, together with (3.5) this yields the convergence statement of the theorem.

Step 3. To prove the representation of the limit, we rely on the theory of [Res87, Chapter 4]. Observe first that $N' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i \boldsymbol{Q}^i)}$ has records at exactly the same time instances as the process $M_0 = \sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$, since by the assumptions of the theorem and by definition of the sequences $\boldsymbol{Q}^i = (Q_j^i)_{j \in \mathbb{Z}}$, all of their components are in $[0, 1]$ with at least one of them being exactly equal to 1. Let $(Y(t) : t > 0)$ be a stochastic process defined by

$$Y(t) = \sup_{T_i \le t} P_i \, , \, t > 0 .$$

Because M_0 is a Poisson point process on $[0, \infty) \times (0, \infty)$ with intensity measure $Leb \times \mathrm{d}(-\vartheta y^{-\alpha})$, $(Y(t))_t$ is the so–called extremal process generated by the distribution function $F(y) = e^{-\vartheta y^{-\alpha}}$, $y \ge 0$. Observe, record times of M_0 correspond to jump times of the process $(Y(t))_t$; write them as a double sided sequence τ_n, $n \in \mathbb{Z}$, such that $\tau_n < \tau_{n+1}$ for each n. According to [Res87, Proposition 4.9], since F is continuous, $\sum_{n \in \mathbb{Z}} \delta_{\tau_n}$ is a Poisson point process with intensity $x^{-1} \mathrm{d}x$ on $(0, \infty)$.

Fix an arbitrary $s > 0$, and assume without loss of generality that τ_1 represents the first record time strictly greater than s, i.e. $\tau_1 = \inf\{\tau_i : \tau_i > s\}$. Denote the Markov chain of corresponding successive record values by U_n, $n \in \mathbb{Z}$ (i.e. $U_n = Y(\tau_n)$); they clearly satisfy $U_n < U_{n+1}$ and $U_0 = Y(s)$. It now follows from [Res87, Proposition 4.7 (iv)] that $\{U_n/U_{n-1}, n \geq 1\}$ is a sequence of i.i.d. random variables Pareto distributed with tail index α.

Because the record times τ_n and record values U_n for $n \geq 1$ of the point process M_0 match the records of the point process $N' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i \boldsymbol{Q}^i)}$ on the interval (s, ∞), we just need to count how many of them appear at any give time τ_n which are larger than the previous record U_{n-1}. If, say, $\tau_n = T_i$, that number corresponds to the number of Q_j^i's which after multiplication by the corresponding $U_n = P_i$ represent a record larger than U_{n-1}. Hence, that random number has the same distribution as

$$\kappa = R^{\boldsymbol{Q}}(U_0/U_1) .$$

Recall that $s > 0$ was arbitrary. Now since the point process $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ and therefore the sequence $\{U_n/U_{n-1}, n \geq 1\}$ is independent of the i.i.d. random elements $(\boldsymbol{Q}^i)$ and since U_1/U_0 has a Pareto distribution with tail index α, the claim follows.

□

3.3 Partial sums

3.3.1 Introduction

Donsker–type functional limit theorems represent one of the key developments in probability theory. They express invariance principles for rescaled random walks of the form

$$S_{\lfloor nt \rfloor} = X_1 + \cdots + X_{\lfloor nt \rfloor} , \quad t \in [0,1] . \tag{3.8}$$

Many extension of the original invariance principle exist, most notably allowing dependence between the steps X_i, or showing, like Skorohod did, that non–Gaussian limits are possible if the steps X_i have infinite variance. For a survey of invariance principles in the case of dependent variables in the domain of attraction of the Gaussian law, we refer to [MPU06], see also [Bra07] for a thorough survey of mixing conditions. In the case of a non–Gaussian limit, the limit of the processes $(S_{\lfloor nt \rfloor})_{t \in [0,1]}$ is not a continuous process in general. Hence, the limiting theorems of this type are placed in the space of càdlàg functions denoted by $D \equiv D([0,1], \mathbb{R})$ under one of the Skorohod topologies. The topology denoted by J_1 is the most widely used (often implicitly) and suitable for i.i.d. steps (see [Res07, Section 7.2]), but over the years many theorems involving dependent steps have been shown using other Skorohod topologies. Even in the case of a simple m–dependent linear process from a regularly varying distribution, it is known that the limiting theorem cannot be shown in the standard J_1 topology, see Avram and Taqqu [AT92, Theorem 1]. Moreover, there are examples of such processes for which none of the Skorohod topologies work, see [AT92,

p. 488].

However, as we found out, for all those processes and many other stochastic models relevant in applications, random walks do converge, but their limit exists in an entirely different space. To describe elements of such a space we use the concept of decorated càdlàg functions due to Whitt [Whi02], and denote the corresponding space by $E \equiv E([0,1],\mathbb{R})$. For the benefit of the reader, in Section 3.3.2, we briefly introduce this space closely following the exposition in Whitt [Whi02, Sections 15.4 and 15.5].

Our main result is a functional limit theorem in the space E for the partial sum process $S_{\lfloor nt \rfloor}$ of a regularly varying stationary time series with tail index $\alpha \in (0,2)$ (i.e. with infinite variance), see Section 3.3.3 and in particular Theorem 3.3.5. As a related goal we also study the running maximum of the random walk $S_{\lfloor nt \rfloor}$ for which, due to monotonicity, the limiting theorem can still be expressed in the familiar space D under the Skorohod's M_1 topology, see Section 3.3.4. In Section 3.3.5, as an another corollary we obtain sufficient conditions under which the partial sums process converges in the space D with respect to the M_2 topology, in particular recovering results of [BK14] for linear processes. Finally, proofs of certain technical auxiliary results are postponed to Section 2.5.

3.3.2 The space of decorated càdlàg functions - E

The elements of $E \equiv E([0,1],\mathbb{R})$ have the form

$$(x, J, \{I(t) : t \in J\})$$

where

- $x \in D([0,1],\mathbb{R})$;
- J is a countable subset of $[0,1]$ with $Disc(x) \subseteq J$, where $Disc(x)$ is the set of discontinuities of the càdlàg function x;
- for each $t \in J$, $I(t)$ is a closed bounded interval in $\mathbb{R}$ (called the *decoration*) such that $x(t), x(t-) \in I(t)$ for all $t \in J$.

Moreover, we assume that for each $\epsilon > 0$, there are at most finitely many times t for which the length of the interval $I(t)$ is greater than ϵ. This ensures that the graphs of elements in E, defined below, are compact subsets of $\mathbb{R}^2$ which allows one to impose a metric on E by using the Hausdorff metric on the space of graphs of elements in E.

Note that every triple $(x, J, \{I(t) : t \in J\})$ can be equivalently represented by a set-valued function

$$x'(t) := \begin{cases} I(t) & \text{if } t \in J\,, \\ \{x(t)\} & \text{if } t \notin J\,, \end{cases}$$

or by the graph of x' defined by

$$\Gamma_{x'} := \{(t,z) \in [0,1] \times \mathbb{R} : z \in x'(t)\}.$$

In the sequel, we will usually denote the elements of E by x'.

Let m denote the Hausdorff metric on the space of compact subsets of $\mathbb{R}^d$ (regardless of dimension) i.e. for compact subsets A, B,

$$m(A,B) = \sup_{x \in A} \|x - B\|_\infty \vee \sup_{y \in B} \|y - A\|_\infty \ ,$$

where $\|x - B\|_\infty = \inf_{y \in B} \|x - y\|_\infty$ with $\|\cdot\|_\infty$ being the sup–norm on $\mathbb{R}^d$. We then define a metric on E, denoted by m_E, by

$$m_E(x', y') = m(\Gamma_{x'}, \Gamma_{y'}) \ . \tag{3.9}$$

We call the topology induced by m_E on E the M_2 topology. This topology is separable, but the metric space (E, m_E) is not complete. Also, we define the uniform metric on E by

$$m^*(x', y') = \sup_{0 \le t \le 1} m(x'(t), y'(t)) \ , \tag{3.10}$$

Obviously, m^* is a stronger metric than m_E, i.e. for any $x', y' \in E$,

$$m_E(x', y') \le m^*(x', y'). \tag{3.11}$$

We will often use the following elementary fact: for $a \le b$ and $c \le d$ it holds that

$$m([a,b],[c,d]) \le |c - a| \vee |d - b|. \tag{3.12}$$

By a slight abuse of notation, we identify every $x \in D$ with an element in E represented by

$$(x, Disc(x), \{[x(t-), x(t)] : t \in Disc(x)\}) \ ,$$

where for any two real numbers a, b by $[a,b]$ we denote the closed interval $[\min\{a,b\}, \max\{a,b\}]$. Consequently, we identify the space D with the subset D' of E given by

$$D' = \{x' \in E : J = Disc(x) \text{ and for all } t \in J,\ I(t) = [x(t-), x(t)]\} \ . \tag{3.13}$$

For an element $x' \in D'$ we have

$$\Gamma_{x'} = \Gamma_x,$$

where Γ_x is the completed graph of x. Since the M_2 topology on D corresponds to the Hausdorff metric on the space of the completed graphs Γ_x, the map $x \to (x, Disc(x), \{[x(t-), x(t)] :$

$t \in Disc(x)\})$ is a homeomorphism from D endowed with the M_2 topology onto D' endowed with the M_2 topology. This yields the following lemma.

Lemma 3.3.1. *The space D endowed with the M_2 topology is homeomorphic to the subset D' in E with the M_2 topology.*

Remark 3.3.2. Because two elements in E can have intervals at the same time point, addition in E is in general not well behaved. However, problems disappear if one of the summands is a continuous function. In such a case, the sum is naturally defined as follows: consider an element $x' = (x, J, \{I(t) : t \in J\})$ in E and a continuous function b on $[0, 1]$, we define the element $x' + b$ in E by

$$x' + b = (x + b, J, \{I(t) + b(t) : t \in J\}) \, .$$

We now state a useful characterization of convergence in (E, m_E) in terms of the local-maximum function defined for any $x' \in E$ by

$$M_{t_1,t_2}(x') := \sup\{z : z \in x'(t), t_1 \le t \le t_2\}, \tag{3.14}$$

for $0 \le t_1 < t_2 \le 1$.

Theorem 3.3.3 (Theorem 15.5.1 Whitt [Whi02])**.** *For elements $x'_n, x' \in E$ the following are equivalent:*

(i) $x'_n \to x'$ in (E, m_E), i.e. $m_E(x'_n, x') \to 0$.

(ii) For all $t_1 < t_2$ in a countable dense subset of $[0, 1]$, including 0 and 1,

$$M_{t_1,t_2}(x'_n) \to M_{t_1,t_2}(x') \text{ in } \mathbb{R}$$

and

$$M_{t_1,t_2}(-x'_n) \to M_{t_1,t_2}(-x') \text{ in } \mathbb{R}.$$

3.3.3 Invariance principle in the space E

Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary $\mathbb{R}$–valued time series which is regularly varying with index $\alpha \in (0, 2)$ and let $(a_n)_n$ be a $\mathbb{R}$–valued sequence satisfying

$$n\mathbb{P}(|X_0| > a_n) \to 1 \, , \text{ as } n \to \infty \, . \tag{3.15}$$

As in the previous section, assume that for some $(r_n)_n$ satisfying $r_n \to \infty$ and $k_n = \lfloor n/r_n \rfloor \to \infty$, and blocks $\boldsymbol{X}_{n,i} = (X_{(i-1)r_n+1}, \ldots, X_{ir_n})$, $i = 1, \ldots, k_n$, it holds that

$$N'_n = \sum_{i=1}^{n} \delta_{(i/k_n, \boldsymbol{X}_{n,i}/a_n)} \xrightarrow{d} N' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i \boldsymbol{Q}^i)} \, , \tag{3.16}$$

in $\mathcal{M}_p([0,1] \times \tilde{l}_{0,0})$, where $\sum_{i=1}^{\infty} \delta_{(T_i,P_i)}$ is a Poisson process in $\mathcal{M}_p([0,1] \times (0,\infty))$ with intensity measure $\vartheta Leb \times d(-y^{-\alpha})$, and $\boldsymbol{Q}^i = (Q^i_j)_{j\in\mathbb{Z}}$, $i \geq 1$, i.i.d. elements of $\tilde{l}_0$, distributed as $\boldsymbol{Q} = (Q_j)_{j\in\mathbb{Z}}$ and independent of $\sum_{i=1}^{\infty} \delta_{(T_i,P_i)}$. The tail process of (X_i) is assumed to be in l_0 a.s., so in particular, the corresponding ϑ and $\boldsymbol{Q}$ from Section 2.3.3 are well defined.

For each $n \in \mathbb{N}$, consider now the partial sum process $(S_n(t))$ in $D([0,1])$ defined by

$$S_n(t) = \sum_{i=1}^{\lfloor nt \rfloor} \frac{X_i}{a_n} \;, \; t \in [0,1] \,, \tag{3.17}$$

and define also

$$V_n(t) = \begin{cases} S_n(t) & \text{if } 0 < \alpha < 1 \,, \\ S_n(t) - \lfloor nt \rfloor \mathbb{E}\left(\frac{X_1}{a_n} \mathbb{1}_{\{|X_1|/a_n \leq 1\}}\right) & \text{if } 1 \leq \alpha < 2 \,. \end{cases} \tag{3.18}$$

As usual, when $1 \leq \alpha < 2$, an additional condition is needed to deal with the small jumps.

Assumption 3.3.4. *For all $\delta > 0$,*

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \mathbb{P}\left(\max_{1\leq k\leq n} \left| \sum_{i=1}^{k} \{ X_i \mathbb{1}_{\{|X_i| \leq a_n\epsilon\}} - \mathbb{E}[X_i \mathbb{1}_{||X_i| \leq a_n\epsilon\}}]\} \right| > a_n\delta \right) = 0 \,. \tag{3.19}$$

Under above assumptions, it was essentially proved in [DH95] that the finite dimensional distributions of V_n converge to those of an α–stable Lévy process; a basic reference for Lévy processes is e.g. Sato [Sat99]. This result is strengthened in [BKS12] to convergence in the M_1 topology if $Q_j Q_{j'} \geq 0$ for all $j \neq j' \in \mathbb{Z}$, i.e. if all extremes within one cluster have the same sign. In the next theorem, we remove the latter restriction and establish convergence of the process V_n in the space E.

For that purpose, we will need to impose some moment conditions on the process $\boldsymbol{Q}$, cf. [DH95, Theorem 3.2]. As originally noted by [DH95, Theorem 2.6], convergence (3.16) and Fatou's lemma imply that $\vartheta \mathbb{E}[\sum_{j\in\mathbb{Z}} |Q_j|^\alpha] \leq 1$ (it was shown in [PS18, Equation (3.14)] that in fact equality holds as soon as ϑ and $\boldsymbol{Q}$ are well–defined). This in particular implies that for $\alpha \in (0,1]$,

$$\mathbb{E}\left[\left(\sum_{j\in\mathbb{Z}} |Q_j| \right)^\alpha \right] < \infty \,. \tag{3.20}$$

For $\alpha > 1$, this will have to be assumed. Furthermore, the case $\alpha = 1$, as usual, requires additional care. We will assume that

$$\mathbb{E}\left[\sum_{j\in\mathbb{Z}} |Q_j| \log \left(|Q_j|^{-1} \sum_{i\in\mathbb{Z}} |Q_i| \right) \right] < \infty \,, \tag{3.21}$$

where we use the convention $|Q_j| \log(|Q_j|^{-1} \sum_{i\in\mathbb{Z}} |Q_i|) = 0$ if $|Q_j| = 0$. Fortunately, it turns out that conditions (3.20) and (3.21) are not too restrictive, see Remark 3.3.7 below.

Finally, regular variation of X_0 and the choice of (a_n) imply that

$$n\mathbb{P}(X_0/a_n \in \cdot) \xrightarrow{v} \mu\,, \tag{3.22}$$

as $n \to \infty$ in $\mathcal{M}(\mathbb{R} \setminus \{0\})$ with measure μ on $\mathbb{R} \setminus \{0\}$ given by

$$\mu(dy) = p\alpha y^{-\alpha-1}\mathbb{1}_{(0,\infty)}(y)\mathrm{d}y + (1-p)\alpha(-y)^{-\alpha-1}\mathbb{1}_{(-\infty,0)}(y)\mathrm{d}y\,, \tag{3.23}$$

for $p = \mathbb{P}(\Theta_0 = 1)$, where $(\Theta_i)_{i\in\mathbb{Z}}$ is the spectral tail process of (X_i).

Theorem 3.3.5. *Let $(X_i : i \in \mathbb{Z})$ be a stationary $\mathbb{R}$-valued regularly varying time series with tail index $\alpha \in (0, 2)$ and assume that the convergence in (3.16) holds. If $\alpha \geq 1$ let Assumption 3.3.4 hold. For $\alpha > 1$, assume that (3.20) holds, and for $\alpha = 1$, assume that (3.21) holds. Then*

$$V_n \xrightarrow{d} V' = (V, \{T_i\}_{i\in\mathbb{N}}, \{I(T_i)\}_{i\in\mathbb{N}})\ ,$$

with respect to the M_2 topology on $E([0, 1], \mathbb{R})$, where

(i) V is an $\alpha-$stable Lévy process on $[0, 1]$ given by

$$V(\,\cdot\,) = \sum_{T_i\leq\cdot}\sum_{j\in\mathbb{Z}} P_iQ_j^i\ ,\ \ 0 < \alpha < 1\ , \tag{3.24a}$$

$$V(\,\cdot\,) = \lim_{\epsilon\to 0}\left(\sum_{T_i\leq\cdot}\sum_{j\in\mathbb{Z}} P_iQ_j^i\mathbb{1}_{\{|P_iQ_j^i|>\epsilon\}} - (\cdot)\int_{\epsilon<|x|\leq 1} x\mu(\mathrm{d}x)\right)\ ,\ \ 1 \leq \alpha < 2\ , \tag{3.24b}$$

where the series in (3.24a) is almost surely absolutely summable and, along some subsequence, (3.24b) holds almost surely uniformly on $[0, 1]$ with μ given in (3.23).

(ii) For all $i \in \mathbb{N}$,

$$I(T_i) = V(T_i-) + P_i\left[\inf_{k\in\mathbb{Z}}\sum_{j\leq k} Q_j^i\ , \sup_{k\in\mathbb{Z}}\sum_{j\leq k} Q_j^i\right].$$

Before proving the theorem, we make several remarks.

Remark 3.3.6. Note that for $\alpha < 1$, convergence of the point process is the only assumption of the theorem. Further, the previous result is applicable to any m–dependent time series which is regularly varying with index $\alpha \in (0, 2)$. Indeed, in this case $(Q_j)_j$ has at most finite number of nonzero terms so (3.20) and (3.21) are automatic. Moreover, the vanishing small values condition (3.19) holds in this case by [TK10a, Lemma 4.8].

Remark 3.3.7. By [PS18, Lemma 3.11], if $\alpha \in (1, 2)$, the condition (3.20) is equivalent to

$$\mathbb{E}\left[\left(\sum_{j=0}^{\infty}|\Theta_j|\right)^{\alpha-1}\right] < \infty \tag{3.25}$$

Furthermore, if $\alpha = 1$, [PS18, Lemma 3.14] shows that the condition (3.21) is then equivalent to

$$\mathbb{E}\left[\log\left(\sum_{j=0}^{\infty}|\Theta_j|\right)\right] < \infty\,. \tag{3.26}$$

These conditions are easier to check than conditions (3.20) and (3.21) since it is easier to determine the distribution of the spectral tail process than the distribution of $\boldsymbol{Q}$. In fact, it suffices to determine only the distribution of the forward spectral tail process $(\Theta_j)_{j\geq 0}$ which is often easier than determining the distribution of the whole spectral tail process. For example, it follows from the proof of [MW14, Theorem 3.2] that for functions of Markov chains satisfying a suitable drift condition (see [MW14] for details and examples), (3.25) and (3.26) hold.

Remark 3.3.8. The α–stable Lévy process V from Theorem 3.3.5 is the weak limit in the sense of finite–dimensional distributions of the partial sum process V_n, characterized by

$$\log\mathbb{E}[\mathrm{e}^{\mathrm{i}zV(1)}] = \begin{cases} \mathrm{i}az + \Gamma(1-\alpha)\cos(\pi\alpha/2)\sigma^\alpha|z|^\alpha\{1-\mathrm{i}\phi\,\mathrm{sgn}(z)\tan(\pi\alpha/2)\} & \alpha\neq 1\,,\\ \mathrm{i}az - \frac{\pi}{2}\sigma|z|\{1+\mathrm{i}\frac{2}{\pi}\phi\,\mathrm{sgn}(z)\log(|z|)\} & \alpha = 1\,, \end{cases} \tag{3.27}$$

with, denoting $x^{\langle\alpha\rangle} = x|x|^{\alpha-1} = x_+^\alpha - x_-^\alpha$,

$$\sigma^\alpha = \vartheta\mathbb{E}[|\sum_{j\in\mathbb{Z}}Q_j|^\alpha]\,,\quad \phi = \frac{\mathbb{E}[(\sum_{j\in\mathbb{Z}}Q_j)^{\langle\alpha\rangle}]}{\mathbb{E}[|\sum_{j\in\mathbb{Z}}Q_j|^\alpha]}$$

and

(i) $a = 0$ if $\alpha < 1$;

(ii) $a = (\alpha-1)^{-1}\alpha\vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}}Q_j^{\langle\alpha\rangle}\right]$ if $\alpha > 1$;

(iii) if $\alpha = 1$, then

$$a = \vartheta\left(c_0\mathbb{E}\left[\sum_{j\in\mathbb{Z}}Q_j\right] - \mathbb{E}\left[\sum_{j\in\mathbb{Z}}Q_j\log\left(\left|\sum_{j\in\mathbb{Z}}Q_j\right|\right)\right] - \mathbb{E}\left[\sum_{j\in\mathbb{Z}}Q_j\log\left(|Q_j|^{-1}\right)\right]\right)\,,$$

with $c_0 = \int_0^\infty(\sin y - y\mathbb{1}_{(0,1]}(y))y^{-2}\mathrm{d}y$.

These parameters were computed in [DH95, Remark 3.2, Theorem 3.2] but with a complicated expression for the location parameter a in the case $\alpha = 1$ (see [DH95, Remark 3.3]). The explicit expression given here, which holds under the assumption (3.21), is new; the proof is given in Lemma 3.3.21. As often done in the literature, if the sequence is assumed to be symmetric then assumption (3.21) is not needed and the location parameter is 0.

Moreover, [PS18, Corollary 3.12, Equation (3.21) and Lemma 3.14] imply that the scale, skewness and location parameters from Remark 3.3.8 can also be expressed in terms of the forward spectral tail process as follows:

$$\sigma^\alpha = \mathbb{E}\left[\left|\sum_{j=0}^{\infty}\Theta_j\right|^\alpha - \left|\sum_{j=1}^{\infty}\Theta_j\right|^\alpha\right],$$
$$\phi = \sigma^{-\alpha}\mathbb{E}\left[\left(\sum_{j=0}^{\infty}\Theta_j\right)^{\langle\alpha\rangle} - \left(\sum_{j=1}^{\infty}\Theta_j\right)^{\langle\alpha\rangle}\right],$$

$a = 0$ if $\alpha < 1$, $a = (\alpha-1)^{-1}\alpha\mathbb{E}[\Theta_0]$ if $\alpha > 1$ and

$$a = c_0\mathbb{E}[\Theta_0] - \mathbb{E}\left[\sum_{j=0}^{\infty}\Theta_j\log\left(\left|\sum_{j=0}^{\infty}\Theta_j\right|\right) - \sum_{j=1}^{\infty}\Theta_j\log\left(\left|\sum_{j=1}^{\infty}\Theta_j\right|\right)\right],$$

if $\alpha = 1$. It can be shown that these expressions coincide for $\alpha \neq 1$ with those in the literature, see e.g. [MW16, Theorem 4.3]. As already noted, the expression of the location parameter for $\alpha = 1$ under the assumption (3.21) (or (3.26)) is new.

Example 3.3.9. Consider again the infinite order moving average process (X_i) from Examples 2.3.7 and 2.4.5, in particular point process convergence (3.16) holds. Assuming that innovations are regularly varying with tail index $\alpha \in (0,2)$, [DR85] proved convergence of the finite–dimensional distributions of the partial sum process V_n. When $c_j \geq 0$ for all $j \in \mathbb{Z}$ (with an additional condition on c_j's when $\alpha \geq 1$ which is not needed, see [TK10b, Corollary 1]), [AT92] proved functional convergence of V_n in the M_1 topology (see [Whi02, Section 12.3] for details on this topology). Under less restrictions on the coefficients (see Example 3.3.17 below), [BK14] obtained convergence in the weaker M_2 topology. We also mention [BJL16] who, with no restrictions on the c_j's (and even weaker conditions than in (2.26) when $\alpha < 1$), proved functional convergence with respect to the S topology (which is weaker than the M_1 topology and incomparable with the M_2 topology, but makes the supremum functional not continuous).

Our Theorem 3.3.5 directly applies to the case of a finite order moving average process. To consider the case of an infinite order moving average process, assume that $\alpha < 1$; the case $\alpha \in [1,2)$ should be treated using m–dependent approximations as in Section 2.4.1. Applying Theorem 3.3.5, one obtains the convergence of the partial sum process $V_n \xrightarrow{d} V' = (V, \{T_i\}_{i\in\mathbb{N}}, \{I(T_i)\}_{i\in\mathbb{N}})$ in E where

$$V(\cdot) = \frac{\sum_{j\in\mathbb{Z}} c_j}{\max_{j\in\mathbb{Z}}|c_j|}\sum_{T_i\leq\cdot} P_iK_i\,,$$

and

$$I(T_i) = V(T_i-) + \frac{P_iK_i}{\max_{j\in\mathbb{Z}}|c_j|}\left[\inf_{k\in\mathbb{Z}}\sum_{j\leq k} c_j\,, \sup_{k\in\mathbb{Z}}\sum_{j\leq k} c_j\right].$$

For an illustration, assume that the innovations (ξ_i) are nonnegative, and consider the process

$$X_i = \xi_i + c\xi_{i-1}\,,\; i \in \mathbb{Z}\,.$$

In the case $c \geq 0$, the convergence of partial sum process in M_1 topology follows from [AT92]. On the other hand, for negative c's convergence fails in any of Skorohod's topology, but partial sums do have a limit in the sense described by our theorem as can be also guessed from Figure 3.2.

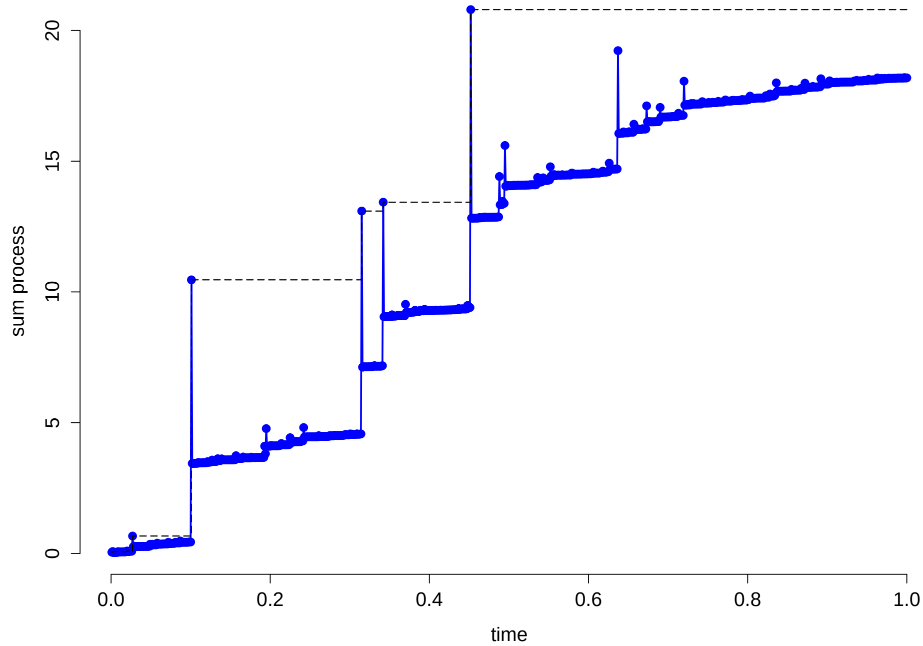

Figure 3.2: A simulated sample path of the process S_n in the case of linear sequence $X_i = \xi_i - 0.7\xi_{i-1}$ with index of regular variation $\alpha = 0.7$ in blue. Observe that due to downward "corrections" after each large jump, in the limit the paths of the process S_n cannot converge to a càdlàg function.

Remark 3.3.10. We do not exclude the case $\sum_{j\in\mathbb{Z}} Q_j = 0$ with probability one, as happens for instance in Example 3.3.9 with $c = -1$. In such a case, the càdlàg component V is simply the null process.

Example 3.3.11. Consider a stationary GARCH$(1,1)$ process

$$X_i = \sigma_i Z_i, \quad \sigma_i^2 = \alpha_0 + \alpha_1 X_{i-1}^2 + \beta_1 \sigma_{i-1}^2, \quad i \in \mathbb{Z}\,,$$

where $\alpha_0, \alpha_1, \beta_1 > 0$, and $(Z_i : i \in \mathbb{Z})$ is a sequence of i.i.d. random variables with mean zero and variance one. Under mild conditions, the process (X_i) is regularly varying and for some (r_n) satisfies all of the assumptions of Theorem 2.3.14, hence point process convergence in (3.16) holds. These hold for instance in the case of standard normal innovations Z_i and sufficiently small parameters α_1, β_1, see [BKS12, Example 4.4] and [MW14, Section 5.4]. Consider such a stationary GARCH$(1,1)$ process with tail index $\alpha \in (0,1)$. Since all the conditions of Theorem 3.3.5 are met, its partial sum process has a limit in the space E (cf. Figure 3.3).

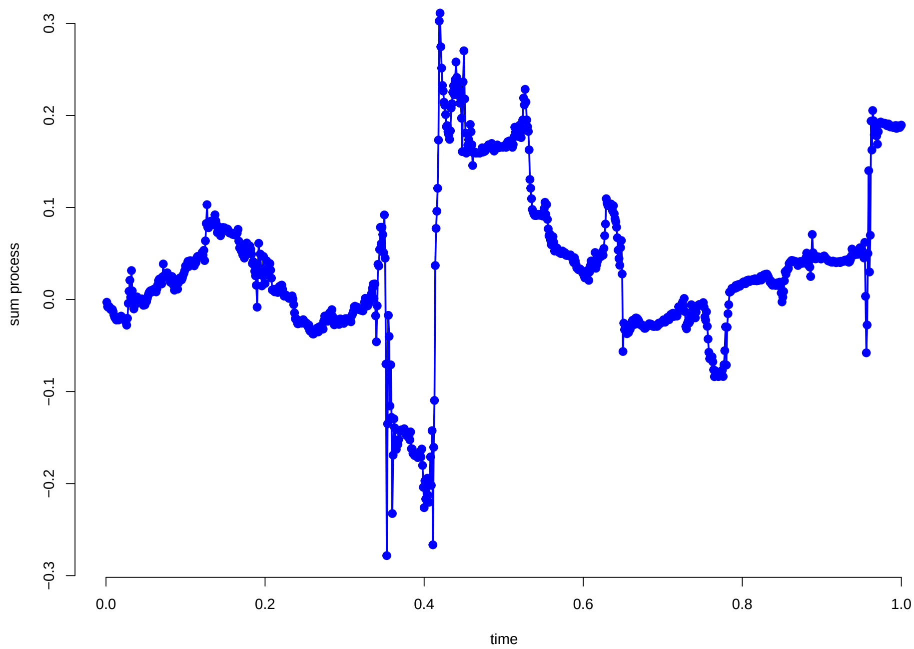

Figure 3.3: A simulated sample path of the process S_n in the case of GARCH$(1,1)$ process with parameters $\alpha_0 = 0.01$, $\alpha_1 = 1.45$ and $\beta_1 = 0.1$, and tail index α between 0.5 and 1.

Remark 3.3.12. If (3.20) holds, the sums $W_i = \sum_{j\in\mathbb{Z}} |Q_j^i|$ are almost surely finite and $(W_i)_{i\geq 1}$ is a sequence of i.i.d. random variables with $\mathbb{E}[W_1^\alpha] < \infty$. Furthermore, by independence of $\sum_{i=1}^\infty \delta_{P_i}$ and (W_i), it follows easily (use [Res87, Propositions 3.7 and 3.8]) that $\sum_{i=1}^\infty \delta_{P_i W_i}$ is also a Poisson process on $(0,\infty)$ with intensity measure $\vartheta \mathbb{E}[W_1^\alpha] \alpha y^{-\alpha-1} dy$. In particular, almost surely, for every $\delta > 0$ there exist at most finitely many indices i such that $P_i W_i > \delta$. By the dominated convergence theorem, this moreover implies that, almost surely,

$$\limsup_{\epsilon\to 0} \sup_{i\in\mathbb{N}} \sum_{j\in\mathbb{Z}} P_i |Q_j^i| \mathbb{1}_{\{P_i |Q_j^i| \leq \epsilon\}} \leq \delta\,,$$

for all $\delta > 0$, hence

$$\sup_{i\in\mathbb{N}} \sum_{j\in\mathbb{Z}} P_i |Q_j^i| \mathbb{1}_{\{P_i |Q_j^i| \leq \epsilon\}} \to 0 \tag{3.28}$$

as $\epsilon \to 0$. These facts will be used several times in the proof.

Proof of Theorem 3.3.5. The proof is split into the case $\alpha < 1$ which is simpler, and the case $\alpha \in [1, 2)$ where centering and truncation introduce additional technical difficulties.

(a) Assume first that $\alpha \in (0, 1)$. We divide the proof into several steps.

Step 1. For every $\epsilon > 0$, consider the functions s^ϵ, u^ϵ and v^ϵ defined on $\tilde{l}_0$ by

$$s^\epsilon(\boldsymbol{x}) = \sum_j x_j \mathbb{1}_{\{|x_j|>\epsilon\}} \ , \quad u^\epsilon(\boldsymbol{x}) = \inf_k \sum_{j \le k} x_j \mathbb{1}_{\{|x_j|>\epsilon\}} \ , \quad v^\epsilon(\boldsymbol{x}) = \sup_k \sum_{j \le k} x_j \mathbb{1}_{\{|x_j|>\epsilon\}} \ ,$$

where $\boldsymbol{x} = (x_j)_{j \in \mathbb{Z}} \in \tilde{l}_0$. Define the mapping $T^\epsilon : \mathcal{M}_p([0, 1] \times \tilde{l}_{0,0}) \to E$ by setting, for $\gamma = \sum_{i=1}^\infty \delta_{t_i, \boldsymbol{x}^i}$,

$$T^\epsilon \gamma = \left(\left(\sum_{t_i \le t} s^\epsilon(\boldsymbol{x}^i) \right)_{t \in [0,1]} , \{t_i : \|\boldsymbol{x}^i\|_\infty > \epsilon\}, \{I(t_i) : \|\boldsymbol{x}^i\|_\infty > \epsilon\} \right) ,$$

where

$$I(t_i) = \sum_{t_j < t_i} s^\epsilon(\boldsymbol{x}^j) + \left[\sum_{t_k = t_i} u^\epsilon(\boldsymbol{x}^k), \sum_{t_k = t_i} v^\epsilon(\boldsymbol{x}^k) \right] .$$

Since γ belongs $\mathcal{M}_p([0, 1] \times \tilde{l}_{0,0})$, there is only a finite number of points $(t_i, \boldsymbol{x}^i)$ such that $\|\boldsymbol{x}^i\|_\infty > \epsilon$ and furthermore, every $\boldsymbol{x}^i = (x^i_j)_{j \in \mathbb{Z}}$ has at most finitely many x^i_j's such that $|x^i_j| > \epsilon$. Therefore, the mapping T^ϵ is well-defined, that is, $T^\epsilon \gamma$ is a proper element of E.

Next, we define subsets of $\mathcal{M}_p([0, 1] \times \tilde{l}_{0,0})$

$$\Lambda_1 = \left\{ \sum_{i=1}^\infty \delta_{t_i, \boldsymbol{x}^i} : |x^i_j| \neq \epsilon, i \ge 1, j \in \mathbb{Z} \right\} ,$$
$$\Lambda_2 = \{ \sum_{i=1}^\infty \delta_{t_i, \boldsymbol{x}^i} : 0 < t_i < 1 \text{ and } t_i \neq t_j \text{ for every } i > j \ge 1 \}.$$

We claim that T^ϵ is continuous on the set $\Lambda_1 \cap \Lambda_2$. Assume that $\gamma_n \xrightarrow{v} \gamma = \sum_{i=1}^\infty \delta_{t_i, \boldsymbol{x}^i} \in \Lambda_1 \cap \Lambda_2$. By Proposition 1.2.8, this implies that finitely many points of γ_n in every bounded set B in $[0, 1] \times \tilde{l}_{0,0}$ such that $\gamma(\partial B) = 0$ converge pointwise to finitely many points of γ in B. In particular, this holds for $B = \{(t, \boldsymbol{x}) : \|\boldsymbol{x}\|_\infty > \epsilon\}$ and it follows that for all $t_1 < t_2$ in $[0, 1]$ such that $\gamma(\{t_1, t_2\} \times \tilde{l}_{0,0}) = 0$,

$$M_{t_1,t_2}(T^\epsilon(\gamma_n)) \to M_{t_1,t_2}(T^\epsilon(\gamma)) \text{ in } \mathbb{R}$$

and

$$M_{t_1,t_2}(-T^\epsilon(\gamma_n)) \to M_{t_1,t_2}(-T^\epsilon(\gamma)) \text{ in } \mathbb{R},$$

with the local-maximum function M_{t_1,t_2} defined as in (3.14). Since the set of all such times is dense in $[0,1]$ and includes 0 and 1, an application of Theorem 3.3.3 gives that

$$T^{\epsilon}(\gamma_n) \to T^{\epsilon}(\gamma)$$

in E endowed with the M_2 topology.

Recall the point process $N' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i \boldsymbol{Q}^i)}$ from (3.16). Since the mean measure of $\sum_{i=1}^{\infty} \delta_{(T_i,P_i)}$ does not have atoms, it is clear that $N' \in \Lambda_1 \cap \Lambda_2$ a.s. Continuous mapping theorem applied to convergence (3.16) now yields that

$$\tilde{S}'_{n,\epsilon} := T^{\epsilon}(N'_n) \xrightarrow{d} T^{\epsilon}(N') =: S'_{\epsilon} .$$

Step 2. Recall that $W_i = \sum_{j\in\mathbb{Z}} |Q_j^i|$ and $\sum_{i=1}^{\infty} \delta_{P_i W_i}$ is a Poisson point process on $(0,\infty)$ with intensity measure $\vartheta \mathbb{E}[W_i^{\alpha}] \alpha y^{-\alpha-1} dy$ (see Remark 3.3.12). Since $\alpha < 1$, by Campbell's theorem (see e.g. [Kin93, p. 28]) one can sum up the points $\{P_i W_i\}$, i.e.

$$\sum_{i=1}^{\infty} P_i W_i = \sum_{i=1}^{\infty} \sum_{j\in\mathbb{Z}} P_i |Q_j^i| < \infty \ \text{ a.s.} , \tag{3.29}$$

In particular, the process

$$V(t) = \sum_{T_i \le t} \sum_{j\in\mathbb{Z}} P_i Q_j^i \, , \, t \in [0,1] \, ,$$

is almost surely well–defined for each $t \in [0,1]$. Since $\sum_{i=1}^{\infty} \delta_{(T_i,P_i,\boldsymbol{Q}^i)}$ is also a Poisson process (on the space $[0,1] \times (0,\infty) \times \tilde{l}_{0,0}$) with intensity measure $Leb \times d(-\vartheta y^{-\alpha}) \times \mathbb{P}(\boldsymbol{Q} \in \cdot)$, it can be shown that V has stationary independent increments and is stochastically continuous (cf. the argument for the case $\alpha \ge 1$ given in the proof of 3.3.21 and also [Res07, pp. 151–153]). Moreover, by the dominated convergence theorem and (3.29), V is the almost sure uniform limit on $[0,1]$ of the piecewise constant process $t \mapsto \sum_{T_i \le t} \sum_{j\in\mathbb{Z}} P_i Q_j^i \mathbb{1}_{\{P_i > \epsilon\}}$ as $\epsilon \to 0$. Hence, the sample paths of V are càdlàg almost surely. Finally, it was shown in [DH95, Theorems 3.1 and 3.2, Remark 3.2] that $V(1)$ is α–stable with characteristic function as given in Remark 3.3.8. Hence, V is an α–stable Lévy process.

Define now the random element V' in $E([0,1],\mathbb{R})$ by

$$V' = \left(V, \{T_i\}_{i\in\mathbb{N}}, \{I(T_i)\}_{i\in\mathbb{N}} \right) \, , \tag{3.30}$$

where

$$\begin{aligned} I(T_i) &= V(T_i-) + \left[u(P_i \boldsymbol{Q}^i), v(P_i \boldsymbol{Q}^i) \right] \, , \\ u(\boldsymbol{x}) &= \inf_k \sum_{j\le k} x_j \, , \quad v(\boldsymbol{x}) = \sup_k \sum_{j \le k} x_j \, . \end{aligned}$$

Since for every $\delta > 0$ there are at most finitely many points P_iW_i such that $P_iW_i > \delta$ and $diam(I(T_i)) = v(P_i\boldsymbol{Q}^i) - u(P_i\boldsymbol{Q}^i) \leq P_iW_i$, V' is indeed a proper element of E a.s.

We now show that, as $\epsilon \to 0$, the limits S'_ϵ from the previous step converge to V' in (E, m) almost surely. Recall the uniform metric m^* on E defined in (3.10). By (3.29) and dominated convergence theorem

$$m^*(S'_\epsilon, V') = \sup_{0\leq t\leq 1} m(S'_\epsilon(t), V'(t)) \leq \sum_{i=1}^{\infty}\sum_{j\in\mathbb{Z}} P_i|Q_j^i|\mathbb{1}_{\{P_i|Q_j^i|\leq\epsilon\}} \to 0 \;, \tag{3.31}$$

almost surely as $\epsilon \to 0$. Indeed, let S_ϵ be the càdlàg part of S'_ϵ, i.e.

$$S_\epsilon(t) = \sum_{T_i\leq t} s^\epsilon(P_i\boldsymbol{Q}^i) = \sum_{T_i\leq t}\sum_{j\in\mathbb{Z}} P_iQ_j^i\mathbb{1}_{\{P_i|Q_j^i|>\epsilon\}} \;,\; t\in[0,1] \;.$$

If $t \notin \{T_i\}$ then

$$m(S'_\epsilon(t), V'(t)) = |S_\epsilon(t) - V(t)| \leq \sum_{T_i\leq t}\sum_{j\in\mathbb{Z}} P_i|Q_j^i|\mathbb{1}_{\{P_i|Q_j^i|\leq\epsilon\}}.$$

Further, when $t = T_k$ for some $k \in \mathbb{Z}$, by using (3.12) we obtain

$$\begin{aligned} m(S'_\epsilon(t), V'(t)) \leq\; & |(S_\epsilon(T_k-) + v^\epsilon(P_k\boldsymbol{Q}^k)) - (V(T_k-) + v(P_k\boldsymbol{Q}^k))| \\ & \vee |(S_\epsilon(T_k-) + u^\epsilon(P_k\boldsymbol{Q}^k)) - (V(T_k-) + u(P_k\boldsymbol{Q}^k))|. \end{aligned} \tag{3.32}$$

The first term on the right-hand side of the equation above is bounded by

$$\begin{aligned} & \left|\sum_{T_i<T_k}\sum_{j\in\mathbb{Z}} P_iQ_j^i\mathbb{1}_{\{P_i|Q_j^i|\leq\epsilon\}}\right| + \left|\sup_{l\in\mathbb{Z}}\sum_{j\leq l} P_kQ_j^k\mathbb{1}_{\{P_k|Q_j^k|>\epsilon\}} - \sup_{l\in\mathbb{Z}}\sum_{j\leq l} P_kQ_j^k\right| \\ & \leq \left|\sum_{T_i<T_k}\sum_{j\in\mathbb{Z}} P_iQ_j^i\mathbb{1}_{\{P_i|Q_j^i|\leq\epsilon\}}\right| + \sup_{l\in\mathbb{Z}}\left|\sum_{j\leq l} P_kQ_j^k\mathbb{1}_{\{P_k|Q_j^k|>\epsilon\}} - \sum_{j\leq l} P_kQ_j^k\right| \\ & \leq \sum_{T_i\leq T_k}\sum_{j\in\mathbb{Z}} P_i|Q_j^i|\mathbb{1}_{\{P_i|Q_j^i|\leq\epsilon\}}. \end{aligned}$$

Since, by similar arguments, one can obtain the same bound for the second term on the right-hand side of (3.32), (3.31) holds. It now follows from (3.11) that

$$S'_\epsilon \to V' \,, \text{ as } \epsilon \to 0 \,,$$

almost surely in (E, m).

Step 3. Recall the blocks for $\boldsymbol{X}_{n,i} = (X_{(i-1)r_n+1}, \ldots, X_{ir_n})$, $i = 1, \ldots, k_n$. Define a random element in E by

$$\tilde{S}'_n = \left(\left(\sum_{i/k_n \le t} s(\boldsymbol{X}_{n,i}/a_n) \right)_{t \in [0,1]}, \{i/k_n\}_{i=1}^{k_n}, \{I(i/k_n)\}_{i=1}^{k_n} \right),$$

where $s((x_j)_{j\in\mathbb{Z}}) = \sum_{j\in\mathbb{Z}} x_j$ and

$$I(i/k_n) = \sum_{j<i} s(\boldsymbol{X}_{n,j}/a_n) + [u(\boldsymbol{X}_{n,i}/a_n), v(\boldsymbol{X}_{n,i}/a_n)] \ .$$

By [Bil68, Theorem 4.2] and the previous two steps, to show that

$$\tilde{S}'_n \xrightarrow{d} V'$$

in (E, m_E), it suffices to prove that, for all $\delta > 0$,

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \mathbb{P}(m_E(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) > \delta) = 0 \ . \tag{3.33}$$

Note first that, by the same arguments as in the previous step, we have

$$m^*(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) \le \sum_{j=1}^{k_n r_n} \frac{|X_j|}{a_n} \mathbb{1}_{\{|X_j| \le a_n \epsilon\}} \ .$$

Hence, by (3.11) and Markov's inequality,

$$\begin{aligned} \mathbb{P}(m_E(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) > \delta) &\le \frac{k_n r_n}{\delta a_n} \mathbb{E}\left[|X_1| \mathbb{1}_{\{|X_1| \le a_n \epsilon\}} \right] \\ &= \frac{k_n r_n}{n\delta} \cdot \frac{\mathbb{E}\left[|X_1| \mathbb{1}_{\{|X_1| \le a_n \epsilon\}} \right]}{a_n \epsilon \mathbb{P}(|X_1| > a_n \epsilon)} \cdot \epsilon n \mathbb{P}(|X_1| > a_n \epsilon) \ . \end{aligned}$$

Since $|X_1|$ is regularly varying with index $\alpha < 1$, Karamata's theorem (see [BDM16, Appendix B.4]) and the choice of (a_n) now imply that

$$\limsup_{n\to\infty} \mathbb{P}(m_E(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) > \delta) \le \frac{\alpha}{(1-\alpha)\delta} \epsilon^{1-\alpha} \to 0 \, ,$$

as $\epsilon \to 0$ since $1 - \alpha > 0$, i.e. (3.33) holds. Hence,

$$\tilde{S}'_n \xrightarrow{d} V'$$

in (E, m_E).

Step 4. Finally, by [Bil68, Theorem 4.1], to show that the original partial sum process S_n from (3.17), and therefore V_n since $\alpha \in (0,1)$, also converges in distribution to V' in

(E, m_E), it suffices to prove that

$$m_E(S_n, \tilde{S}'_n) \xrightarrow{\mathbb{P}} 0 \, . \tag{3.34}$$

Recall that $k_n = \lfloor n/r_n \rfloor$ so $\frac{ir_n}{n} \le \frac{i}{k_n}$ for all $i = 0, 1, \dots, k_n$ and moreover

$$\frac{i}{k_n} - \frac{ir_n}{n} = \frac{i}{k_n}\left(1 - \frac{k_n r_n}{n}\right) \le 1 - \frac{\lfloor n/r_n \rfloor}{n/r_n} = 1 - \left(1 - \frac{\{n/r_n\}}{n/r_n}\right) \le \frac{r_n}{n} \, . \tag{3.35}$$

Let $d_{n,i}$ for $i = 0, \dots, k_n - 1$ be the Hausdorff distance between restrictions of graphs Γ_{S_n} and $\Gamma_{\tilde{S}'_n}$ on time intervals $(\frac{ir_n}{n}, \frac{(i+1)r_n}{n}]$ and $(\frac{i}{k_n}, \frac{i+1}{k_n}]$, respectively (see Figure 3.4).

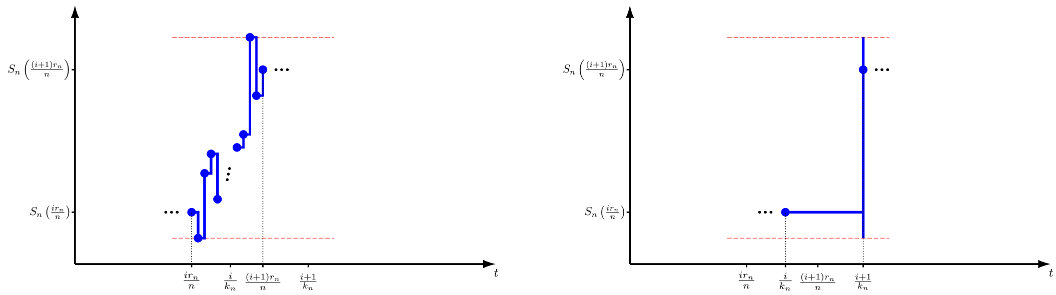

Figure 3.4: Restrictions of graphs Γ_{S_n} and $\Gamma_{\tilde{S}'_n}$ on time intervals $(\frac{ir_n}{n}, \frac{(i+1)r_n}{n}]$ and $(\frac{i}{k_n}, \frac{i+1}{k_n}]$, respectively.

First note that, by (3.35), the time distance between any two points on these graphs is at most $2r_n/n$. Further, by construction, S_n and $\tilde{S}'_n$ have the same range of values on these time intervals. More precisely,

$$\bigcup_{t \in (\frac{ir_n}{n}, \frac{(i+1)r_n}{n}]} \{z \in \mathbb{R} \, : \, (t, z) \in \Gamma_{S_n}\} = \bigcup_{t \in (\frac{i}{k_n}, \frac{i+1}{k_n}]} \{z \in \mathbb{R} \, : \, (t, z) \in \Gamma_{\tilde{S}'_n}\} = \tilde{S}'_n((i+1)/k_n) \, .$$

Therefore, the distance between the graphs comes only from the time component, i.e.

$$d_{n,i} \le \frac{2r_n}{n} \to 0 \, , \quad \text{as } n \to \infty \, ,$$

for all $i = 0, 1, \dots, k_n - 1$.

Moreover, if we let d_{n,k_n} be the Hausdorff distance between the restriction of the graph Γ_{S_n} on $(\frac{k_n r_n}{n}, 1]$ and the interval $(1, \tilde{S}'_n(1))$, it holds that

$$d_{n,k_n} \le \frac{r_n}{n} \vee \sum_{j=k_n r_n + 1}^{n} \frac{|X_j|}{a_n} \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$. Hence, (3.34) holds since

$$m_E(S_n, \tilde{S}'_n) \le \bigvee_{i=0}^{k_n} d_{n,i},$$

and this finishes the proof in the case $\alpha \in (0,1)$.

(b) Assume now that $\alpha \in [1,2)$. As shown in Step 1. in the proof of (a), for every $\epsilon > 0$ it holds that

$$\tilde{S}'_{n,\epsilon} \xrightarrow{d} S'_\epsilon \tag{3.36}$$

in E, where $\tilde{S}'_{n,\epsilon} = T^\epsilon(N'_n)$ and $S'_\epsilon = T^\epsilon(N')$. For every $\epsilon > 0$ define a càdlàg process $S_{n,\epsilon}$ by

$$S_{n,\epsilon}(t) = \sum_{i=1}^{\lfloor nt \rfloor} \frac{X_i}{a_n} \mathbb{1}_{\{|X_i|/a_n > \epsilon\}}\,,\ t \in [0,1]\,.$$

Using the same arguments as in Step 4. in the proof of (a), it holds that, as $n \to \infty$,

$$m_E(S_{n,\epsilon}, \tilde{S}'_{n,\epsilon}) \xrightarrow{\mathbb{P}} 0\,. \tag{3.37}$$

By [Bil68, Theorem 4.1], (3.36) and (3.37) imply that for every $\epsilon > 0$,

$$S_{n,\epsilon} \xrightarrow{d} S'_\epsilon \tag{3.38}$$

in (E, m_E).

Since $\alpha \in [1,2)$ we need to introduce centering. For $\epsilon > 0$, define càdlàg process by setting, for $t \in [0,1]$,

$$V_{n,\epsilon}(t) = S_{n,\epsilon}(t) - \lfloor nt \rfloor \mathbb{E}\left(\frac{X_1}{a_n} \mathbb{1}_{\{\epsilon < |X_1|/a_n \le 1\}}\right)\,.$$

From (3.22) we have, for any $t \in [0,1]$, as $n \to \infty$,

$$\lfloor nt \rfloor \mathbb{E}\left(\frac{X_1}{a_n} \mathbb{1}_{\{\epsilon < |X_1|/a_n \le 1\}}\right) \to t \int_{\{x\,:\,\epsilon < |x| \le 1\}} x\mu(dx). \tag{3.39}$$

Since the limit function above is continuous and the convergence is uniform on $[0,1]$, (3.38) and Lemma 3.3.18 yield that for every $\epsilon > 0$,

$$V_{n,\epsilon} \xrightarrow{d} V'_\epsilon \tag{3.40}$$

in E, where V'_ϵ is given by (see Remark 3.3.2)

$$V'_\epsilon(t) = S'_\epsilon(t) - t \int_{\{x\,:\,\epsilon < |x| \le 1\}} x\mu(dx)\,.$$

Let V_ϵ be the càdlàg part of V'_ϵ, i.e.,

$$V_\epsilon(t) = \sum_{T_i \le t} s^\epsilon(P_i \boldsymbol{Q}^i) - t \int_{\{x\,:\,\epsilon<|x|\le 1\}} x\mu(dx)\ . \tag{3.41}$$

By Lemma 3.3.19, there exist an α-stable Lévy process V and a sequence $\epsilon_k \to 0$, as $k \to \infty$, such that

$$\|V_{\epsilon_k} - V\|_\infty := \sup_{t\in[0,1]} |V_{\epsilon_k}(t) - V(t)| \to 0\,, \tag{3.42}$$

as $k \to \infty$ almost surely. Next, as argued in Step 2. in the proof of (a),

$$V' = \left(V, \{T_i\}_{i\in\mathbb{N}}, \{I(T_i)\}_{i\in\mathbb{N}}\right)\ , \tag{3.43}$$

where

$$I(T_i) = V(T_i-) + \left[u(P_i\boldsymbol{Q}^i), v(P_i\boldsymbol{Q}^i)\right]\ , \tag{3.44}$$

is a proper element of E. Also, one can argue similarly to the proof of (3.31) to conclude that

$$m^*(V', V'_{\epsilon_k}) \le \|V - V_{\epsilon_k}\|_\infty + \sup_{i\in\mathbb{N}} \sum_{j\in\mathbb{Z}} P_i|Q^i_j| \mathbb{1}_{\{P_i|Q^i_j|\le\epsilon\}}\ .$$

Now it follows from (3.11), (3.28) and (3.42) that

$$V'_{\epsilon_k} \to V' \tag{3.45}$$

in (E, m_E) as $k \to \infty$ almost surely.

Finally, by (3.40), (3.45) and [Bil68, Theorem 4.2], to show that the original (centered) partial sum process V_n from (3.18) satisfies

$$V_n \xrightarrow{d} V'\,, \tag{3.46}$$

in (E, m_E), it suffices to prove that, for all $\delta > 0$,

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \mathbb{P}(m_E(V_{n,\epsilon}, V_n) > \delta) = 0\ . \tag{3.47}$$

But this follows from Assumption 3.3.4 and (3.11) since

$$m^*(V_{n,\epsilon}, V_n) \le \max_{1\le k\le n} \left| \sum_{i=1}^k \left(\frac{X_i}{a_n} \mathbb{1}_{\{|X_i|/a_n\le\epsilon\}} - \mathbb{E}\left[\frac{X_i}{a_n}\mathbb{1}_{\{|X_i|/a_n\le\epsilon\}}\right]\right)\right|.$$

Hence (3.46) holds and this finishes the proof.

□

3.3.4 Supremum of the partial sum process

We next show that the supremum of the partial sum process converges in distribution in D endowed with the M_1 topology, where the limit is the "running supremum" of the limit process V' from Theorem 3.3.5. See [Whi02, Section 12.3] for details on the M_1 topology.

Let V be the α–stable Lévy process defined in (3.24) and define the process V^+ on $[0,1]$ by

$$V^+(t) = \begin{cases} V(t)\,, & t \notin \{T_j\}_{j\in\mathbb{N}} \\ V(t-) + \sup_{k\in\mathbb{Z}} \sum_{j\le k} P_i Q_j^i\,, & t = T_i \text{ for some } i \in \mathbb{N}. \end{cases}$$

Define V^- analogously using infimum instead of supremum. Note that V^+ and V^- need not be right-continuous at the jump times T_j. However, their partial supremum or infimum are càdlàg functions.

Theorem 3.3.13. *Under the same conditions as in Theorem 3.3.5, it holds that*

$$\left(\sup_{s\le t} V_n(s)\right)_{t\in[0,1]} \xrightarrow{d} \left(\sup_{s\le t} V^+(s)\right)_{t\in[0,1]}\,,$$

and

$$\left(\inf_{s\le t} V_n(s)\right)_{t\in[0,1]} \xrightarrow{d} \left(\inf_{s\le t} V^-(s)\right)_{t\in[0,1]}\,,$$

jointly in $D([0,1],\mathbb{R})$ endowed with the M_1 topology.

Proof. We prove the result only for the supremum of the partial sum process since the infimum case is completely analogous and joint convergence holds since we are applying the continuous mapping argument to the same process.

Define the mapping $\sup : E([0,1],\mathbb{R}) \to D([0,1],\mathbb{R})$ by

$$\sup(x')(t) = \sup\{z : z \in x'(s), 0 \le s \le t\}\,.$$

Note that $\sup(x')$ is non-decreasing and since for every $\delta > 0$ there are at most finitely many times t for which the $diam(x'(t))$ is greater than δ, by [Whi02, Theorem 15.4.1.] it follows easily that this mapping is well-defined, i.e. that $\sup(x')$ is indeed an element in D. Also, by construction,

$$\sup(V') = \left(\sup_{s\le t} V^+(s)\right)_{t\in[0,1]}$$

and

$$\sup(V_n) = \left(\sup_{s\le t} V_n(s)\right)_{t\in[0,1]}\,.$$

Define the subset of E by

$$\Lambda = \{x' \in E \;:\; x'(0) = \{0\}\}$$

and assume that $x'_n \to x'$ in (E, m_E), where $x'_n, x' \in \Lambda$. By Theorem 3.3.3 it follows that

$$\sup(x'_n)(t) = M_{0,t}(x'_n) \to M_{0,t}(x') = \sup(x')(t)$$

for all t in a dense subset of (0,1], including 1. Also, the convergence trivially holds for $t = 0$ since $\sup(x'_n)(0) = \sup(x')(0) = 0$ for all $n \in \mathbb{N}$. Since $\sup(x')$ is non-decreasing for all $x' \in E$, we can apply [Whi02, Corollary 12.5.1] and conclude that

$$\sup(x'_n) \to \sup(x')$$

in D endowed with M_1 topology. Since $V_n, V' \in \Lambda$ almost surely, by Theorem 3.3.5 and continuous mapping argument it follows that

$$\left(\sup_{s\le t} V_n(s)\right)_{t\in[0,1]} \xrightarrow{d} \left(\sup_{s\le t} V^+(s)\right)_{t\in[0,1]}$$

in D endowed with M_1 topology. □

Remark 3.3.14. Note that when $\sum_{j\in\mathbb{Z}} Q_j = 0$ a.s., the limit for the supremum of the partial sum process in Theorem 3.3.13 is simply a so called Fréchet extremal process. For an illustration of the general limiting behavior of running maxima in the case of a linear processes, consider again the moving average of order 1 from Example 3.3.9. Figure 3.2 shows a path (dashed line) of the running maxima of the MA(1) process $X_t = \xi_t - 0.7\xi_{t-1}$.

3.3.5 M_2 convergence of the partial sum process

We can now characterize the convergence of the partial sum process in the M_2 topology in $D([0,1])$ by an appropriate condition on the tail process of the sequence $(X_i : i \in \mathbb{Z})$.

Assumption 3.3.15. *The sequence $(Q_j)_{j\in\mathbb{Z}}$ almost surely satisfies*

$$-\left(\sum_{j\in\mathbb{Z}} Q_j\right)_- = \inf_{k\in\mathbb{Z}} \sum_{j\le k} Q_j \le \sup_{k\in\mathbb{Z}} \sum_{j\le k} Q_j = \left(\sum_{j\in\mathbb{Z}} Q_j\right)_+ . \tag{3.48}$$

Note that this assumption ensures that $\sum_{j\in\mathbb{Z}} Q_j \neq 0$ and that the limit process V' from Theorem 3.3.5 has sample paths in the subset D' of E which was defined in (3.13). By Lemma 3.3.1, Theorem 3.3.5 and the continuous mapping theorem, the next result follows immediately.

Theorem 3.3.16. *If, in addition to conditions in Theorem 3.3.5, Assumption 3.3.15 holds, then*

$$V_n \xrightarrow{d} V$$

in $D([0,1],\mathbb{R})$ endowed with the M_2 topology.

Since the supremum functional is continuous with respect to the M_2 topology, this result implies that the limit of the running supremum of the partial sum process is the running supremum of the limiting α-stable Lévy process as in the case of i.i.d. random variables.

Example 3.3.17. For the linear process $X_t = \sum_{j\in\mathbb{Z}} c_j \xi_{t-j}$ from Examples 2.3.7 and 2.4.5, the condition (3.48) can be expressed as

$$-\left(\sum_{j\in\mathbb{Z}} c_j\right)_- = \inf_{k\in\mathbb{Z}} \sum_{j\leq k} c_j \leq \sup_{k\in\mathbb{Z}} \sum_{j\leq k} c_j = \left(\sum_{j\in\mathbb{Z}} c_j\right)_+ . \tag{3.49}$$

This is exactly [BK14, Condition 3.2]. Note that (3.49) implies that

$$\left|\sum_{j\in\mathbb{Z}} c_j\right| > 0 .$$

3.3.6 Postponed proofs

On continuity of addition in E

Lemma 3.3.18. *Suppose that $(x_n : n \in \mathbb{N})$ is a sequence in $D([0,1],\mathbb{R})$ and $x' = (x, S, \{I(t) : t \in S\})$ an element in E such that $x_n \to x'$ in E. Suppose also that $(b_n : n \in \mathbb{N})$ is a sequence in $D([0,1],\mathbb{R})$ which converges uniformly to a continuous function b on $[0,1]$. Then $x_n - b_n \to x' - b$ in (E, m_E) where*

$$x' - b := (x - b, S, \{I(t) - b(t) : t \in S\}) .$$

Proof. Recall the definition of m_E given in (3.9). By Whitt [Whi02, Theorem 15.5.1.] to show that $x_n - b_n \to x' - b$ in E, it suffices to prove that

$$\sup_{(t,z)\in\Gamma_{x_n-b_n}} \|(t,z) - \Gamma_{x'-b}\|_\infty \to 0. \tag{3.50}$$

Take an arbitrary $\epsilon > 0$. Note that b is uniformly continuous so by the conditions of the lemma there exists $0 < \delta \leq \epsilon$ and $n_0 \in \mathbb{N}$ such that

(i) $|t - s| < \delta \Rightarrow |b(t) - b(s)| < \epsilon$,

(ii) $m_E(x_n, x') < \delta$, for all $n \geq n_0$ and

(iii) $|b_n(t) - b(t)| < \epsilon$, for all $t \in [0,1]$.

Also, since b is continuous, it easily follows that $|b_n(t) - b_n(t-)| \le 2\epsilon$ for all $n \ge n_0$ and $t \in [0,1]$.

Take $n \ge n_0$ and a point $(t,z) \in \Gamma_{x_n - b_n}$, i.e.

$$z \in [(x_n(t-) - b_n(t-)) \wedge (x_n(t) - b_n(t)), (x_n(t-) - b_n(t-)) \vee (x_n(t) - b_n(t))].$$

Since $|b_n(t) - b_n(t-)| \le 2\epsilon$ there exists $z' \in [x_n(t-) \wedge x_n(t), x_n(t-) \vee x_n(t)]$ (i.e. $(t,z') \in \Gamma_{x_n}$), such that

$$|(z' - b_n(t)) - z| \le 2\epsilon.$$

Next, since $m_E(x_n, x') < \delta$, there exists a point $(s,y) \in \Gamma_{x'}$ such that

$$|s-t| \vee |y - z'| < \delta.$$

Note that $(s, y - b(s)) \in \Gamma_{x'-b}$ and by previous arguments

$$\begin{aligned} |(y - b(s)) - z| &= |(y - b(s)) - z + (z' - b_n(t)) - (z' - b_n(t)) + b(t) - b(t)| \\ &\le |y - z'| + |b(t) - b(s)| + |b_n(t) - b(t)| + |(z' - b_n(t)) - z| \\ &\le \delta + \epsilon + \epsilon + 2\epsilon \\ &\le 5\epsilon. \end{aligned}$$

Also, $|s - t| < \delta \le \epsilon$. Hence, for all $n \ge n_0$,

$$\sup_{(t,z)\in\Gamma_{x_n - b_n}} \|(t,z) - \Gamma_{x'-b}\|_\infty \le 5\epsilon.$$

and since ϵ was arbitrary, (3.50) holds. $\square$

A lemma for partial sum convergence in E

Lemma 3.3.19. *Let $\alpha \in [1,2)$ and let the assumptions of Theorem 3.3.5 hold. Then there exists an α-stable Lévy process V on $[0,1]$ and a sequence $\epsilon_k \nearrow 0$ such that, as $\epsilon \to 0$, the process V_{ϵ_k} from (3.41) converges uniformly a.s. to V, i.e. $\|V_{\epsilon_k} - V\| := \sup_{t\in[0,1]} |V_{\epsilon_k}(t) - V(t)| \to 0$. The characteristic function of $V(1)$ equals the one given in Remark 3.3.8.*

Remark 3.3.20. The proof below does not use the vanishing small values condition (3.19).

Proof. Recall that for all $\epsilon > 0$ and $t \in [0,1]$,

$$\begin{aligned} V_\epsilon(t) &= \sum_{T_i \le t} s^\epsilon(P_i \boldsymbol{Q}^i) - t \int_{\{x\,:\,\epsilon<|x|\le 1\}} x\mu(dx) \\ &= \sum_{T_i \le t} \sum_{j\in\mathbb{Z}} P_i Q_j^i \mathbb{1}_{\{P_i|Q_j^i|>\epsilon\}} - t \int_{\{x\,:\,\epsilon<|x|\le 1\}} x\mu(dx)\,, \end{aligned}$$

where

$$\mu(dx) = p\alpha x^{-\alpha-1}\mathbb{1}_{(0,\infty)}(x)\mathrm{d}x + (1-p)\alpha(-x)^{-\alpha-1}\mathbb{1}_{(-\infty,0)}(x)\mathrm{d}x$$

for $p = \mathbb{P}(\Theta_0 = 1)$. We first show that the centering term can be expressed as an expectation of a functional of the limiting point process N'. More precisely, we show that for all $\epsilon > 0$,

$$\int_{\{x\,:\,\epsilon<|x|\le 1\}} x\mu(dx) = \vartheta \int_0^\infty \mathbb{E}\left[y \sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon<y|Q_j|\le 1\}} \right] \alpha y^{-\alpha-1}\mathrm{d}y\,. \tag{3.51}$$

First, by [PS18, Equation (3.21)],

$$\vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j|Q_j|^{\alpha-1}\right] = 2p-1\,. \tag{3.52}$$

so by Fubini's theorem, if $\alpha > 1$

$$\begin{aligned} \vartheta \int_0^\infty \mathbb{E}\left[y \sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon<y|Q_j|\le 1\}} \right] \alpha y^{-\alpha-1}\mathrm{d}y &= \alpha\vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j \int_{\epsilon|Q_j|^{-1}}^{|Q_j|^{-1}} y^{-\alpha}\mathrm{d}y\right] \\ &= \frac{\alpha}{\alpha-1}(\epsilon^{-\alpha+1}-1)\vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j|Q_j|^{\alpha-1}\right] \\ &= \frac{\alpha}{\alpha-1}(\epsilon^{-\alpha+1}-1)(2p-1)\,, \end{aligned}$$

and if $\alpha = 1$ the same term equals $\log(\epsilon^{-1})(2p-1)$. Note that the use of Fubini's theorem is justified since the same calculation as above shows that the above integral converges absolutely since $\mathbb{E}[\sum_{j\in\mathbb{Z}}|Q_j|^\alpha] < \infty$. The equality in (3.51) now follows by the definition of the measure μ. Hence, for all $t \in [0,1]$,

$$V_\epsilon(t) = \sum_{T_i \le t} s^\epsilon(P_i \boldsymbol{Q}^i) - t\vartheta \int_0^\infty \mathbb{E}\left[y \sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon<y|Q_j|\le 1\}} \right] \alpha y^{-\alpha-1}\mathrm{d}y\,.$$

Set $W = \sum_{j\in\mathbb{Z}} Q_j$ and for $\delta > 0$ and $\epsilon \ge 0$ define

$$m_{\epsilon,\delta} = \vartheta \int_0^\infty \mathbb{E}\left[y \sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon<y|Q_j|\le 1,\,\delta<yW\}} \right] \alpha y^{-\alpha-1}dy\,.$$

By assumptions (3.20) and (3.21) $m_{0,\delta}$ is well defined for all $\delta > 0$, and moreover dominated convergence theorem yields that $\lim_{\epsilon\to 0} m_{\epsilon,\delta} = m_{0,\delta}$ for all $\delta > 0$. Indeed, if $\alpha = 1$, for all $\delta > 0$

$$\vartheta \int_0^\infty \mathbb{E}\left[y \sum_{j\in\mathbb{Z}} |Q_j| \mathbb{1}_{\{y|Q_j|\le 1,\, \delta < yW\}} \right] \alpha y^{-\alpha-1} \mathrm{d}y \le \vartheta \mathbb{E}\left[\sum_{j\in\mathbb{Z}} |Q_j| \int_{\frac{\delta\wedge 1}{W}}^{\frac{1}{|Q_j|}} y^{-1}\mathrm{d}y \right]$$
$$= \vartheta \mathbb{E}\left[\sum_{j\in\mathbb{Z}} |Q_j| \log(|Q_j|^{-1}) + W\log W + \log((\delta\wedge 1)^{-1})W \right] ,$$

which is finite by assumption (3.21), and if $\alpha > 1$ using a similar calculation together with the assumption $\mathbb{E}[W^\alpha] < \infty$ one obtains the same conclusion.

Recall from Remark 3.3.12 that $W_i = \sum_{j\in\mathbb{Z}} |Q^i_j|$, $i \in \mathbb{N}$, is a sequence of i.i.d. random variables so $\sum_{i=1}^\infty \delta_{P_i W_i}$ is a Poisson point process on $(0,\infty)$ with intensity measure $\vartheta\mathbb{E}[W_1^\alpha]\alpha y^{-\alpha-1}dy$. Since $\mathbb{E}[W_1^\alpha] < \infty$, for every $\delta > 0$ there a.s. exists at most finitely many points P_iW_i such that $P_iW_i > \delta$. Hence, for every $\delta > 0$ and $\epsilon \ge 0$ the process

$$V_{\epsilon,\delta}(t) = \sum_{T_i\le t} s^\epsilon(P_i\boldsymbol{Q}^i)\mathbb{1}_{\{\delta<P_iW_i\}} - tm_{\epsilon,\delta} = \sum_{T_i\le t}\sum_{j\in\mathbb{Z}} P_iQ^i_j \mathbb{1}_{\{\epsilon<P_i|Q^i_j|,\, \delta<P_iW_i\}} - tm_{\epsilon,\delta} ,$$

for $t \in [0,1]$, is a well–defined random element of $D[0,1]$, and moreover, as $\epsilon \to 0$, $V_{\epsilon,\delta}$ converges uniformly almost surely to

$$V_{0,\delta}(\cdot) = \sum_{T_i\le \cdot}\sum_{j\in\mathbb{Z}} P_iQ^i_j\mathbb{1}_{\{P_iW_i>\delta\}} - (\cdot)\vartheta\int_0^\infty \mathbb{E}\left[y\sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\delta<yW\}}\right]\alpha y^{-\alpha-1}dy .$$

Next, fix a positive sequence $\{\delta_k\}$ with $\delta_k \searrow 0$ as $k\to\infty$. We prove that V_{0,δ_k} converges uniformly almost surely in $D([0,1])$.

Note, $\sum_{i\ge 1}\delta_{T_i,P_i,\boldsymbol{Q}^i}$ is a Poisson point process on $[0,1]\times(0,\infty)\times\tilde{l}_{0,0}$ with intensity measure $\vartheta Leb \times d(-y^{-\alpha}) \times \mathbb{P}(\boldsymbol{Q} \in \cdot)$. In particular, the process $V_{0,\delta}$ has independent increments with respect to δ, that is for every $\delta < \delta'$, $V_{0,\delta} - V_{0,\delta'}$ is independent of $V_{0,\delta'}$. Moreover, by Campbell's theorem (see e.g. [Kin93, p. 28]),

$$\mathrm{Var}(V_{0,\delta}(1) - V_{0,\delta'}(1)) = \vartheta\int_0^\infty y^2\mathbb{E}\left[\left(\sum_{j\in\mathbb{Z}} Q_j\right)^2 \mathbb{1}_{\{\delta<yW\le\delta'\}}\right]\alpha y^{-\alpha-1}\mathrm{d}y$$
$$\le \vartheta\mathbb{E}\left[W^2\int_0^{\delta'/W}\alpha y^{-\alpha+1}\mathrm{d}y\right] = \frac{\vartheta\alpha(\delta')^{2-\alpha}}{(2-\alpha)}\mathbb{E}[W^\alpha] .$$

Therefore, $\lim_{\delta'\to 0}\sup_{\delta<\delta'}\mathrm{Var}(V_{0,\delta}(1) - V_{0,\delta'}(1)) = 0$ and now arguing exactly as in the proof of [Res07, Proposition 5.7, Property 2] shows that $(V_{0,\delta_k})_k$ is almost surely a Cauchy sequence in $D([0,1])$ with respect to the supremum metric $\|\cdot\|_\infty$. Since the space $D([0,1])$ is complete under this metric, we obtain the existence of a process $V = (V(t) : t\in[0,1])$

with paths in $D([0,1])$ almost surely and such that $\lim_{k\to\infty} \|V_{0,\delta_k} - V\|_\infty = 0$ almost surely.

Lemma 3.3.21. *The process V satisfies $V(0) = 0$ a.s. and has independent stationary increments and is stochastically continuous, hence V is a Lévy process. Moreover, V is α–stable with characteristic function of $V(1)$ given in Remark 3.3.8.*

The proof of the previous result is given right after the end of the present proof. There only remains to prove that for all $u > 0$,

$$\lim_{\delta\to 0} \limsup_{\epsilon\to 0} \mathbb{P}(\|V_\epsilon - V_{\epsilon,\delta}\|_\infty > u) = 0 \,. \tag{3.53}$$

Indeed, this would imply that $\|V_\epsilon - V\|_\infty \to 0$ in probability and hence that, along some subsequence, V_ϵ converges to V uniformly almost surely. Since for $\delta \le 1$, $yW = \sum_{j\in\mathbb{Z}} y|Q_j| \le \delta$ implies that $y|Q_j| \le \delta \le 1$ for all $j \in \mathbb{Z}$, we have that

$$\begin{aligned} V_\epsilon(t) - V_{\epsilon,\delta}(t) = \sum_{T_i\le t} \sum_{j\in\mathbb{Z}} P_i Q_j^i \mathbb{1}_{\{\epsilon < P_i|Q_j^i|,\, P_iW_i\le\delta\}} \\ - t\vartheta \int_0^\infty \mathbb{E}\left[y\sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon<y|Q_j|,\, yW\le\delta\}} \right] \alpha y^{-\alpha-1}\mathrm{d}y \,. \end{aligned}$$

The process $V_\epsilon - V_{\epsilon,\delta}$ is a càdlàg martingale, thus applying Doob–Meyer's inequality yields

$$\begin{aligned} \mathbb{P}\left(\|V_\epsilon - V_{\epsilon,\delta}\|_\infty > u\right) &\le u^{-2}\,\mathrm{Var}(V_\epsilon(1) - V_{\epsilon,\delta}(1)) \\ &= u^{-2}\vartheta \int_0^\infty y^2 \mathbb{E}\left[\left(\sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon<y|Q_j|,\, yW\le\delta\}} \right)^2 \right] \alpha y^{-\alpha-1}\mathrm{d}y \\ &\le u^{-2}\vartheta \mathbb{E}\left[W^2 \int_0^{\delta/W} \alpha y^{-\alpha+1}\mathrm{d}y \right] = \frac{\vartheta\alpha\delta^{2-\alpha}}{u^2(2-\alpha)} \mathbb{E}[W^\alpha] \end{aligned}$$

and hence (3.53) holds since $\alpha < 2$. □

Proof of Lemma 3.3.21. Recall, $V(t)$ is the almost sure limit of $V_{0,\delta_k}(t)$ as $k \to \infty$ for all $t \in [0,1]$. Observe further that for all $0 \le s < t \le 1$,

$$V_{0,\delta_k}(t) - V_{0,\delta_k}(s) = \sum_{s<T_i\le t} \sum_{j\in\mathbb{Z}} P_i Q_j^i \mathbb{1}_{\{\delta_k < P_iW_i\}} - (t-s)m_{0,\delta} \,.$$

Note also that $V_{0,\delta_k}(0) = 0$ almost surely for all $k \in \mathbb{N}$ and hence $V(0) = 0$. Further, recall that $\sum_{i\ge 1} \delta_{T_i,P_i,\boldsymbol{Q}^i}$ is a Poisson point process on $[0,1] \times (0,\infty) \times \tilde{l}_{0,0}$ with intensity measure $\vartheta Leb \times d(-y^{-\alpha}) \times \mathbb{P}(\boldsymbol{Q} \in \cdot)$. In particular, V has independent increments (see [Res07, pp. 151]), and by [Kin93, p. 28, Equation (3.17)] for all $k \in \mathbb{N}, 0 \le s < t \le 1, z \in \mathbb{R}$,

$$\begin{aligned}\log \mathbb{E}[e^{\mathrm{i}z(V_{0,\delta_k}(t)-V_{0,\delta_k}(s))}] &= (t-s)\vartheta \int_0^\infty \mathbb{E}\left[\left\{e^{\mathrm{i}zyS}-1\right\}\mathbb{1}\{\delta < yW\}\right]\alpha y^{-\alpha-1}dy - \mathrm{i}z(t-s)m_{0,\delta}\\ &= (t-s)\log \mathbb{E}[e^{\mathrm{i}zV_{0,\delta_k}(1)}] = \log \mathbb{E}[e^{\mathrm{i}zV_{0,\delta_k}(t-s)}]\,,\end{aligned}$$

where $W = \sum_{j\in\mathbb{Z}} |Q_j|$ and $S = \sum_{j\in\mathbb{Z}} Q_j$. Letting $k \to \infty$ yields that for all $z \in \mathbb{R}$, $\mathbb{E}[e^{\mathrm{i}z(V_0(t)-V_0(s))}] = \mathbb{E}[e^{\mathrm{i}zV_0(1)}]^{t-s} = \mathbb{E}[e^{\mathrm{i}zV_0(t-s)}]$ for all such s,t. This implies that V has stationary increments and that it is stochastically continuous (see [Res07, pp. 153]).

It remains to show that $V(1)$ has an α-stable distribution with the stated parameters. We show this only for the case $\alpha = 1$ since the case $\alpha > 1$ is similar.

Recall, $m_{0,\delta} = \vartheta\int_0^\infty \mathbb{E}\left[y\sum_{j\in\mathbb{Z}} Q_j\mathbb{1}_{\{\delta<yW\}}\right]y^{-2}dy$. Since $yW \le 1$ implies that $y|Q_j| \le 1$ for all $j \in \mathbb{Z}$, for all $\delta < 1$ we have that

$$\begin{aligned}V_{0,\delta}(1) = \Bigg(\sum_{i,j} P_i Q_j^i \mathbb{1}\{\delta < P_i W_i\} - \vartheta\int_0^\infty \mathbb{E}\left[y\sum_{j\in\mathbb{Z}} Q_j \mathbb{1}_{\{\delta<yW\le 1\}}\right]y^{-2}dy\Bigg)\\ - \vartheta\int_0^\infty \mathbb{E}\left[y\sum_{j\in\mathbb{Z}} Q_j\mathbb{1}\{y|Q_j|\le 1,\ 1<yW\}\right]y^{-2}dy\,.\end{aligned}$$

By Fubini's theorem, the last term above is equal to

$$\vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j\log(W)\right] + \vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j \log(|Q_j|^{-1})\right]\,,$$

(with the usual convention $0\log 0 = 0$). Therefore, for all $z \in \mathbb{R}$ and $\delta < 1$

$$\begin{aligned}\log\mathbb{E}\left[e^{\mathrm{i}zV_{0,\delta}(1)}\right] = \vartheta\int_0^\infty \mathbb{E}\left[\left\{e^{\mathrm{i}zyS}-1-\mathrm{i}zyS\mathbb{1}\{yW\le 1\}\right\}\mathbb{1}\{\delta<yW\}\right]y^{-2}dy\\ - \mathrm{i}z\vartheta\mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j\log(|Q_j|^{-1}W)\right]\,.\end{aligned} \tag{3.54}$$

Since for all $\delta < 1$, using the fact that $|e^{\mathrm{i}z} - 1 - \mathrm{i}z| \le |z|^2/2 \le |z|^2$ for all $z \in \mathbb{R}$ (see for example [Sat99, Lemma 8.6]) and $\mathbb{E}[W] < \infty$,

$$\begin{aligned}&\mathbb{E}\left[\int_0^\infty \left|e^{\mathrm{i}zyS}-1-\mathrm{i}zyS\mathbb{1}\{yW\le 1\}\right|\mathbb{1}\{\delta<yW\}y^{-2}dy\right]\\ &\le \mathbb{E}\left[\int_{1/W}^\infty \left|e^{\mathrm{i}zyS}-1\right|y^{-2}dy\right] + \mathbb{E}\left[\int_{\delta/W}^{1/W}\left|e^{\mathrm{i}zyS}-1-\mathrm{i}zyS\right|y^{-2}dy\right]\\ &\le 2\mathbb{E}[W] + |z|^2(1-\delta)\mathbb{E}[W] \le (2+|z|^2)\mathbb{E}[W] < \infty\,,\end{aligned}$$

by the dominated convergence theorem, as $\delta \to 0$ the first term on the right side of (3.54) tends to

$$
\begin{aligned}
\vartheta \int_0^\infty & \mathbb{E}\left[\left\{e^{\mathrm{i}zyS} - 1 - \mathrm{i}zyS\mathbb{1}\{yW \le 1\}\right\}\right] y^{-2}dy \\
&= \vartheta\mathbb{E}\left[\int_0^\infty \left\{e^{-\mathrm{i}zyS} - 1 - \mathrm{i}zyS\mathbb{1}_{(0,1]}(y)\right\} y^{-2}dy\right] + \mathrm{i}z\vartheta\mathbb{E}\left[S\int_{1/W}^1 y^{-1}dy\right] .
\end{aligned}
$$

Altogether, using the integral from [Sat99, Page 85] we get that

$$
\begin{aligned}
\lim_{\delta\to 0}\log\mathbb{E}\left[e^{\mathrm{i}zV_{0,\delta}(1)}\right] &= -\vartheta\frac{\pi}{2}|z|\mathbb{E}[|S|] - i\vartheta z\log|z|\mathbb{E}[S] - \mathrm{i}z\vartheta\mathbb{E}[S\log|S|] + \mathrm{i}c_0\vartheta z\mathbb{E}[S] \\
&\quad + \mathrm{i}z\vartheta\mathbb{E}[S\log W] - \mathrm{i}z\vartheta\mathbb{E}[S\log W] - \vartheta\mathbb{E}[\sum_{j\in\mathbb{Z}} Q_j\log(|Q_j|^{-1})] ,
\end{aligned}
$$

where

$$
c_0 = \int_0^\infty \frac{\sin y - y\mathbb{1}_{(0,1]}(y)}{y^2}dy .
$$

Setting $\sigma = \vartheta\mathbb{E}[|S|]$, $\phi = \frac{\mathbb{E}[S]}{\mathbb{E}[|S|]}$ and

$$
a = \vartheta\left(c_0\mathbb{E}[S] - \mathbb{E}[S\log|S|] - \mathbb{E}\left[\sum_{j\in\mathbb{Z}} Q_j\log(|Q_j|^{-1})\right]\right) ,
$$

since the term $\mathrm{i}z\vartheta\mathbb{E}[S\log W]$ cancels out, we obtain that

$$
\begin{aligned}
\log\mathbb{E}\left[e^{\mathrm{i}zV(1)}\right] &= \lim_{k\to\infty}\log\mathbb{E}\left[e^{\mathrm{i}zV_{0,\delta_k}(1)}\right] \\
&= \lim_{\delta\to 0}\log\mathbb{E}\left[e^{\mathrm{i}zV_{0,\delta}(1)}\right] \\
&= -\frac{\pi}{2}\sigma|z|\left(1 + \mathrm{i}\frac{2}{\pi}\phi\,\mathrm{sgn}(z)\log|z|\right) + \mathrm{i}az .
\end{aligned}
$$

$\square$

Chapter 4

Sequence alignment problem

The local alignment problem was studied extensively in the probabilistic setting, see for instance [AGG89, DKZ94b, Han06] and references therein. We first explain its key ingredients and also state our main result in that context.

4.1 Introduction

Let $(A_i)_{i\in\mathbb{N}}$ and $(B_i)_{i\in\mathbb{N}}$ be two independent i.i.d. sequences taking values in a finite alphabet E. Also, let A and B be independent random variables distributed as A_1 and B_1, respectively. For a fixed score function $s : E \times E \to \mathbb{R}$ and for all $i, j \in \mathbb{N}$ and $m = 0, 1, \ldots, i \wedge j$ (where $i \wedge j := \min\{i, j\}$), let

$$S^m_{i,j} = \sum_{k=0}^{m-1} s(A_{i-k}, B_{j-k})$$

be the score of aligning segments $A_{i-m+1}, \ldots, A_i$ and $B_{j-m+1}, \ldots, B_j$. Further, for all $i, j \in \mathbb{N}$ define

$$S_{i,j} = \max\{S^m_{i,j} : 0 \leq m \leq i \wedge j\}\,. \tag{4.1}$$

From biological perspective it is essential to understand the extremal distributional properties of the random matrix $(S_{i,j} : 1 \leq i, j \leq n)$ as $n \to \infty$. The following simple assumption is standard in this context, cf. Dembo et al. [DKZ94b].

Assumption 4.1.1. *The distribution of $s(A, B)$ is nonlattice, i.e. $\mathbb{P}(s(A, B) \in \delta\mathbb{Z}) < 1$ for all $\delta > 0$, and satisfying*

$$\mathbb{E}[s(A, B)] < 0 \quad \textit{and} \quad \mathbb{P}(s(A, B) > 0) > 0\,. \tag{4.2}$$

The lattice case is excluded for simplicity in the sequel. It is known to be conceptually similar, although technicaly more involved. Note further that, like [DKZ94b] and [Han06], we consider only gapless local alignments.

Denote by μ_A and μ_B the distributions of A and B, respectively, and assume for simplicity that $\mu_A(e)$, $\mu_B(e) > 0$ for each letter e in the alphabet E. By Assumption 4.1.1 there exists the unique strictly positive solution α^* of the Lundberg equation

$$m(\alpha^*) := \mathbb{E}[e^{\alpha^* s(A,B)}] = 1\,. \tag{4.3}$$

Let μ^* be the (exponentially tilted) probability measure on $E \times E$ given by

$$\mu^*(a,b) = e^{\alpha^* s(a,b)} \mu_A(a) \mu_B(b)\;,\; a, b \in E\,. \tag{4.4}$$

For two probability measures μ and ν on a finite set F, denote by $H(\nu|\mu)$ the relative entropy of ν with respect to μ, i.e.

$$H(\nu|\mu) = \sum_{x \in F} \nu(x) \log \frac{\nu(x)}{\mu(x)}\,.$$

Dembo et al. [DKZ94b] introduce one final condition on the tilted probability measure μ^*. It essentially restricts extremal dependence within the field $(S_{i,j})$ in a way which seems biologically meaningful and exactly suits their, as well as our, asymptotic analysis of the field.

Assumption 4.1.2 (Condition (E') in [DKZ94b])**.** *It holds that*

$$H(\mu^*|\mu_A \times \mu_B) > 2\,\{H(\mu_A^*|\mu_A) \vee H(\mu_B^*|\mu_B)\}\,, \tag{4.5}$$

where μ_A^* and μ_B^* denote the marginals of μ^*.

Note that (4.5) holds automatically if $\mu_A = \mu_B$ and if the score function s is symmetric (i.e. $s(a,b) = s(b,a)$) and not of the form $s(a,b) = s(a) + s(b)$, see [DKZ94a, Section 3].

Under Assumptions 4.1.1 and 4.1.2, Dembo et al. [DKZ94b] (see also Hansen [Han06]) showed that the distribution of the maximal local alignment score $M_n = \max_{1 \le i,j \le n} S_{i,j}$, asymptotically follows a Gumbel distribution. More precisely, as $n \to \infty$, for a certain constant $K^* > 0$,

$$\mathbb{P}\left(M_n - \frac{2\log(n)}{\alpha^*} \le x\right) \to e^{-K^* e^{-\alpha^* x}}\,,\; x \in \mathbb{R}\,. \tag{4.6}$$

Observe that the field $(S_{i,j})$ consists of dependent random variables. For instance, simple arguments can be given (cf. (4.9) below) showing that any extreme score, i.e. score exceeding a given large threshold, will be followed by a run of extreme scores along the diagonal. This phenomena is illustrated in Figure 4.1 for both real life and simulated sequences. The approach of [DKZ94b] is based on showing that the number of such extreme clusters, as both sample size and the threshold tend to infinity, becomes asymptotically Poisson distributed.

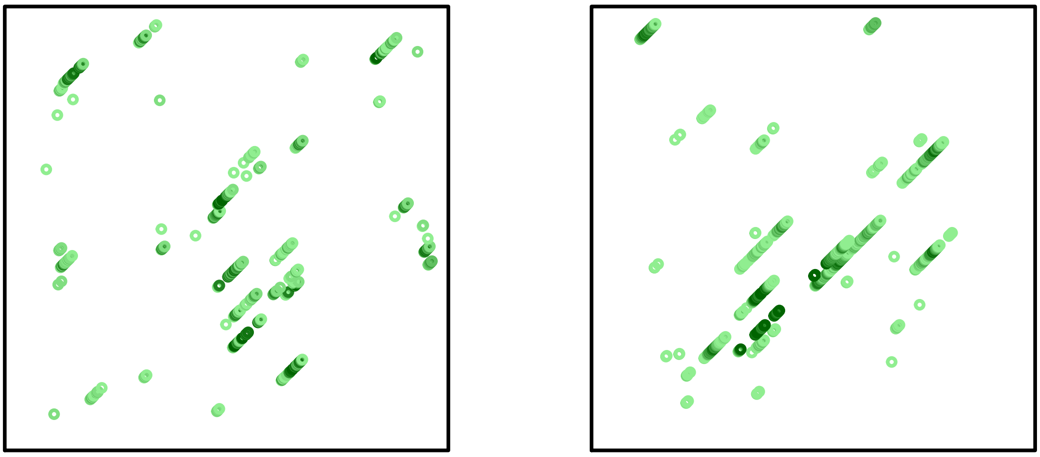

Figure 4.1: Heatmap of large local scores for alignments of two simulated sequences (left) and two regions of human and fruit-fly genome (right). Each sequence is 1000 nucleotides long.

In the sequel, we show that one can give much more detailed information about the structure within extreme clusters. In particular, following the method below one can deduce the asymptotic distribution of arbitrary functionals of upper order statistics of the field $(S_{i,j})$.

Observe that for each $i, j \in \mathbb{N}$, $S_{i,j}$ can be seen as the maximum of a truncated random walk $(S^m_{i,j})_{m=0,\dots,i\wedge j}$ which by (4.2) has negative drift. It can be rigorously shown, see Remark 4.3.3, that in all our asymptotic considerations this truncation and related edge effects can be ignored. Therefore we assume throughout that sequences (A_i) and (B_i) extend over all integers $i \in \mathbb{Z}$. This makes scores $S^m_{i,j}$ well defined for all $i, j \in \mathbb{Z}$ and $m \geq 0$, and consequently we update the original field of scores $(S_{i,j})$ as follows

$$S_{i,j} = \sup\{S^m_{i,j} : m \geq 0\}\,,\ i, j \in \mathbb{Z}\,. \tag{4.7}$$

By construction, the field $(S_{i,j})$ is stationary. Moreover, by the classical Cramér-Lundberg theory (see e.g. [Asm03, Part C, XIII.5]), Assumption 4.1.1 implies that the tail of $S_{i,j}$ is asymptotically exponential, or more precisely

$$\mathbb{P}(S_{i,j} > u) \sim Ce^{-\alpha^* u}\,,\ \text{as}\ \ u \to \infty\,, \tag{4.8}$$

for some $C > 0$. Note that, in the language of extreme value theory, marginal distribution of the field $(S_{i,j})$, belongs to the maximum domain of attraction of the Gumbel distribution. In this light, the limiting result (4.6) may not be very surprising, but its proof remains quite involved due to the clustering of extremal scores of the field $(S_{i,j})$. Observe, the field

$(S_{i,j})$ satisfies the following simple (Lindley) recursion along any diagonal, namely

$$S_{i,j} = (S_{i-1,j-1} + \varepsilon_{i,j})_+ \,, \tag{4.9}$$

where random variables $\varepsilon_{i,j} = s(A_i, B_j)$ have negative mean.

Our main result in this context, strengthens (4.6) to convergence in distribution of point processes based on the $S_{i,j}$'s. The key observation is that, under Assumptions 4.1.1 and 4.1.2, the transformed field $(e^{S_{i,j}})$ is jointly regularly varying with index α^*, see Proposition 4.2.1 below. Distribution of its spectral tail field $(\Theta_{i,j})_{i,j\in\mathbb{Z}}$ can be described in detail using two auxiliary independent i.i.d. sequences $(\varepsilon_i)_{i\geq 1}$ and $(\varepsilon_i^*)_{i\geq 1}$ whose distributions correspond to the distributions of $s(A,B)$ under the product measure $\mu_A \times \mu_B$ and under the tilted measure μ^* from (4.4), respectively. Set $S_0^\varepsilon = 0$ and

$$S_m^\varepsilon = \sum_{i=1}^m \varepsilon_i \,, \text{ for } m \geq 1 \quad \text{ and } S_m^\varepsilon = -\sum_{i=1}^{-m} \varepsilon_i^* \,, \text{ for } m \leq -1 \,. \tag{4.10}$$

Then $\Theta_{i,j} = 0$ for all $i \neq j$ and $\Theta_{m,m} = e^{S_m^\varepsilon}$, for $m \in \mathbb{Z}$. Observe, (S_m^ε) is an asymmetric double sided random walk.

Theorem 4.1.3. *Under Assumptions 4.1.1 and 4.1.2, as $n \to \infty$,*

$$\sum_{i,j=1}^n \delta_{\left(\frac{(i,j)}{n},\, S_{i,j} - \frac{2\log(n)}{\alpha^*}\right)} \xrightarrow{d} \sum_{k\in\mathbb{N}} \sum_{m\in\mathbb{Z}} \delta_{(\boldsymbol{T}_k, \tilde{P}_k + \tilde{Q}_m^k)}$$

in the $\mathcal{M}_p([0,1]^2 \times \mathbb{R})$, where the Poisson cluster process in the limit has the following components

(i) $\sum_{k\in\mathbb{N}} \delta_{(\boldsymbol{T}_k, \tilde{P}_k)}$ *is a Poisson point process on $[0,1]^2 \times \mathbb{R}$ with intensity measure Leb $\times$ $d(-\vartheta C e^{-\alpha^* u})$ where, for an exponential random variable Γ with parameter α^* independent of (S_m^ε),*

$$\vartheta = \mathbb{P}(\sup_{m\geq 1} S_m^\varepsilon + \Gamma \leq 0) \,;$$

(ii) $(\tilde{Q}_m^k)_{m\in\mathbb{Z}}$, $k \in \mathbb{N}$ *are i.i.d. two-sided $\mathbb{R}$-valued sequences, independent of* $\sum_{k\in\mathbb{N}} \delta_{(\boldsymbol{T}_k, \tilde{P}_k)}$ *and with common distribution equal to the distribution of the random walk $(S_m^\varepsilon)_m$ conditioned on staying negative for $m < 0$ and nonpositive for $m > 0$.*

For the convergence in distribution above, bounded sets in $[0,1]^2 \times \mathbb{R}$ are those which are contained in sets $[0,1]^2 \times (x, \infty)$, $x \in \mathbb{R}$. Observe that application of the theorem yields (4.6) at once with the following new expression for the key constant therein

$$K^* = \vartheta C \,.$$

The expression for the extremal index ϑ appears frequently in applications, see e.g. de Haan et al. [dHRRdV89] where an algorithm for its numerical computation is suggested. The constant C arising from (4.8) is also frequently encountered in the literature, for various expressions of C we refer to [Asm03, Part C, XIII.5].

The distribution of random walks conditioned to stay negative (or positive) is discussed in detail by Tanaka [Tan89] and Biggins [Big03]. Consequently, one can apply those results to simulate and precisely describe the distribution of $\tilde{Q}_m^k$'s. Putting all these ingredients together, one can use Theorem 4.1.3 to give a probabilistic and geometric intepretation of the plots in Figure 4.1. Note that this type of limiting behaviour was conjectured already by Metzler et al. [MGW02] who suggested a marked Poissonian model of local alignments with essentially same features.

The rest of the chapter is devoted to the proof of Theorem 4.1.3 using the theory of regularly varying random fields described Chapter 2. For some of the key technical results in our analysis we are indepted to Hansen [Han06] who even allows sequences (A_i) and (B_i) to be Markov chains. In the i.i.d. setting, however, the corresponding proofs, which rely on change of measure arguments, are much less involved. For an alternative approach based on combinatorial arguments see Dembo et al. [DKZ94b].

This chapter is based on the paper [BP18b].

4.2 The tail field

Recall, $(A_i)_{i\in\mathbb{Z}}$ and $(B_i)_{i\in\mathbb{Z}}$ are independent i.i.d. sequences, $S_{i,j}^m = \sum_{k=0}^{m-1} s(A_{i-k}, B_{j-k})$ for $i,j \in \mathbb{Z}$ and $m \geq 0$, and $S_{i,j} = \sup\{S_{i,j}^m : m \geq 0\}$ for $i,j \in \mathbb{Z}$. We assume throughout that Assumptions 4.1.1 and 4.1.2 hold.

Consider a positive stationary field $\boldsymbol{X} = (X_{i,j} : i,j \in \mathbb{Z})$ defined by

$$X_{i,j} = e^{S_{i,j}},\ i,j \in \mathbb{Z}\,.$$

Observe that by (4.8), for $\alpha^* > 0$ satisfying $\mathbb{E}[e^{\alpha^* s(A,B)}] = 1$,

$$\mathbb{P}(X_{i,j} > u) \sim C u^{-\alpha^*},\quad \text{as } u \to \infty\,, \tag{4.11}$$

i.e. marginal distributions of $\boldsymbol{X}$ are regularly varying. More importantly, the transformed field $\boldsymbol{X}$ has a tail field, i.e. it is jointly regularly varying. Observe, (4.9) implies that $\boldsymbol{X}$ satisfies

$$X_{m,m} = \max\{X_{m-1,m-1} e^{s(A_m,B_m)}, 1\}\,,$$

for each $m \in \mathbb{Z}$, where $X_{m-1,m-1}$ is independent of the i.i.d. sequence $(e^{s(A_k,B_k)} : k \geq m)$, cf. Example 2.3.8.

Proposition 4.2.1. *The field $\boldsymbol{X}$ is regularly varying with tail index α^* and with the spectral tail field $\boldsymbol{\Theta} = (\Theta_{i,j} : i,j \in \mathbb{Z})$ satisfying*

(i) $\Theta_{i,j} = 0$ *for* $i,j \in \mathbb{Z}$, $i \neq j$.

(ii) $\Theta_{m,m} = e^{S_m^\varepsilon}$ *for* $m \in \mathbb{Z}$, *where* $S_0^\varepsilon = 0$ *and*

$$S_m^\varepsilon = \sum_{i=1}^{m} \varepsilon_i\,, \text{ for } m \geq 1 \quad \text{ and } S_m^\varepsilon = -\sum_{i=1}^{-m} \varepsilon_i^*\,, \text{ for } m \leq -1\,, \tag{4.12}$$

for independent i.i.d. sequences $(\varepsilon_i)_{i\geq 1}$ *and* $(\varepsilon_i^*)_{i\geq 1}$ *whose distributions correspond to the distributions of* $s(A,B)$ *under the product measure* $\mu_A \times \mu_B$ *and under the tilted measure* μ^* *from (4.4), respectively.*

The tail field $\boldsymbol{Y} = (Y_{i,j})_{i,j\in\mathbb{Z}}$ of $\boldsymbol{X}$ is therefore given by $Y_{i,j} = Y\Theta_{i,j}$ where Y satisfies $\mathbb{P}(Y \geq y) = y^{-\alpha^*}$ for $y \geq 1$ and is independent from $\boldsymbol{\Theta}$. Observe, $\mathbb{E}[\varepsilon_1] = \mathbb{E}[s(A,B)] < 0$ and since the moment generating function $m(\alpha) = \mathbb{E}[e^{\alpha s(A,B)}]$ is strictly convex and $m(0) = m(\alpha^*) = 1$,

$$\mathbb{E}[\varepsilon_1^*] = \mathbb{E}[s(A,B)e^{\alpha^* s(A,B)}] = \frac{\mathrm{d}m}{\mathrm{d}\alpha}(\alpha^*) > 0\,.$$

This implies that $\mathbb{P}(\lim_{|m|\to\infty} S_m^\varepsilon = -\infty) = 1$ so $\boldsymbol{\Theta}$ and $\boldsymbol{Y}$ are elements of l_0 almost surely. In particular, by (2.21),

$$0 < \vartheta = \mathbb{P}(\sup_{(i,j)\succ_l (0,0)} Y_{i,j} \leq 1) = \mathbb{P}(Y \max_{m\geq 1} \Theta_{m,m} \leq 1) = \mathbb{P}(\log Y + \max_{m\geq 1} S_m^\varepsilon \leq 0)\,, \tag{4.13}$$

where, note, $\log Y$ is a standard exponential random variable with index α^*. Also, by (2.24), the distribution of the random element $\boldsymbol{Q} = (Q_{i,j})_{i,j\in\mathbb{Z}}$ in $\tilde{l}_0$ is determined by

$$Q_{i,j} = 0 \text{ for } i \neq j\,,\ (Q_{m,m})_{m\in\mathbb{Z}} \overset{d}{=} \Big(e^{S_m^\varepsilon}, m \in \mathbb{Z} \mid \sup_{m\leq -1} S_m^\varepsilon < 0, \sup_{m\geq 1} S_m^\varepsilon \leq 0\Big)\,. \tag{4.14}$$

To prove Proposition 4.2.1 we need two auxiliary lemmas. The first one is a rough estimate using Markov inequality, see Section 4.4 for the proof.

Lemma 4.2.2. *There exist a constant* $c_0 > 0$ *such that*

$$\lim_{u\to\infty} e^{2\alpha^* u}\mathbb{P}\Big(\max_{m > c_0 u} S_{0,0}^m \geq 0\Big) = 0\,.$$

Before we state the second lemma, observe first that, using $\mathbb{E}[e^{\alpha^* s(A,B)}] = 1$, for all $u \geq 0$ and any integer $m \geq 0$,

$$\mathbb{P}(S_{0,0}^m \geq u) = \mathbb{E}[e^{-\alpha^* S_{0,0}^m} e^{\alpha^* S_{0,0}^m} \mathbb{1}\Big\{S_{0,0}^m \geq u\Big\}] \leq e^{-\alpha^* u}\mathbb{P}^*(S_{0,0}^m \geq u) \leq e^{-\alpha^* u}\,,$$

where the tilted measure $\mathbb{P}^*$ makes pairs (A_k, B_k) for $k = -m+1, \ldots, 0$, independent and distributed according to the measure μ^*. The following result is a special case of [Han06, Lemma 5.11]. The proof relies on change of measure arguments and the Azuma–Hoeffding inequality for martingales. The key fact is that, whenever $\mu^* \neq \mu_A^* \times \mu_B^*$ (which holds under (4.5)),

$$\mathbb{E}_{\nu_A \times \nu_B}[s(A,B)] < \mathbb{E}_{\mu^*}[s(A,B)]$$

for all $\nu_A \in \{\mu_A, \mu_A^*\}$ and $\nu_B \in \{\mu_B, \mu_B^*\}$, where $\mathbb{E}_\mu$ denotes expectation assuming (A,B) is distributed according to μ, see [DKZ94a, beginning of Section 3].

Lemma 4.2.3 ([Han06, Lemma 5.11])**.** *There exists an $0 < \epsilon_0 < 1$ such that for all $u > 0$,*

$$\sup_{\substack{i,j\in\mathbb{Z},\, i\neq j\\ m,l\geq 0}} \mathbb{P}(S_{0,0}^m > u, S_{i,j}^l > u) \leq 2e^{-(1+\epsilon_0)\alpha^* u}\,.$$

Since the proof given in [Han06, Lemma 5.11] is simpler in the i.i.d. setting, for convenience, we present it in Section 4.4. We are now in position to prove Proposition 4.2.1.

Proof of Proposition 4.2.1. Let $\boldsymbol{\Theta}$ be from the statement of the proposition. We first show that, as $u \to \infty$,

$$X_{0,0}^{-1}\boldsymbol{X}_I \;\Big|\; X_{0,0} > u \xrightarrow{d} \boldsymbol{\Theta}_I\,, \tag{4.15}$$

for all $I \subseteq \mathbb{Z}^2 \setminus \{(m,m) : m \leq -1\}$. Since $X_{0,0}$ is regularly varying with index α^*, this will prove the regular variation property of $\boldsymbol{X}$ and show that the spectral tail field $\boldsymbol{\Theta}' = (\Theta'_{i,j})_{i,j\in\mathbb{Z}}$ of $\boldsymbol{X}$ satisfies

$$(\Theta'_{i,j} : (i,j) \in \mathbb{Z}^2 \setminus \{(m,m) : m \leq -1\}) \stackrel{d}{=} (\Theta_{i,j} : (i,j) \in \mathbb{Z}^2 \setminus \{(m,m) : m \leq -1\})\,, \tag{4.16}$$

see Remark 2.2.4.

Observe, by (4.9), for each $m \geq 1$,

$$X_{m,m} = \max\{X_{m-1,m-1}e^{s(A_m,B_m)}, 1\}\,.$$

Now since $X_{0,0}$ is regularly varying and independent of the i.i.d. sequence $(e^{s(A_k,B_k)})_{k\geq 1}$, [Seg07, Theorem 2.3] implies that for all $m \geq 0$, as $u \to \infty$,

$$X_{0,0}^{-1}(X_{0,0}, X_{1,1}, \ldots, X_{m,m}) \;\Big|\; X_{0,0} > u \xrightarrow{d} \Big(1, e^{s(A_1,B_1)}, \ldots, \prod_{k=1}^m e^{s(A_k,B_k)}\Big) \stackrel{d}{=} (\Theta_{0,0}, \Theta_{1,1}, \ldots, \Theta_{m,m})\,.$$

Since $\Theta_{i,j} = 0$ for all $i, j \in \mathbb{Z}$, $i \neq j$, (4.15) will follow if we show that for all such i, j,

$$\begin{aligned}\mathbb{P}(X_{i,j} > X_{0,0}\eta \mid X_{0,0} > u) &\le \mathbb{P}(X_{i,j} > u\eta \mid X_{0,0} > u) \\ &= \mathbb{P}(S_{i,j} > \log u + \log \eta \mid S_{0,0} > \log u) \to 0\,, \text{ as } u \to \infty\,,\end{aligned} \tag{4.17}$$

for all $\eta \in (0, 1)$.

Fix now $i, j \in \mathbb{Z}$ such that $i \neq j$. Using (4.8) and Lemmas 4.2.2 and 4.2.3, for every $M \geq 0$,

$$\begin{aligned}\limsup_{u\to\infty} \mathbb{P}(S_{i,j} > u - M \mid S_{0,0} > u) &= \limsup_{u\to\infty} C^{-1} e^{\alpha^* u} \mathbb{P}(S_{0,0} > u, S_{i,j} > u - M) \\ &\le \limsup_{u\to\infty} C^{-1} e^{\alpha^* u} \mathbb{P}\left(\max_{1\le m\le c_0 u} S_{0,0}^m > u - M,\ \max_{1\le l\le c_0 u} S_{i,j}^l > u - M \right) \\ &\le \limsup_{u\to\infty} 2C^{-1} e^{(1+\epsilon)\alpha^* M} e^{-\epsilon\alpha^* u} = 0\,,\end{aligned}$$

hence (4.17) holds.

Finally, we extend (4.16) to equality in distribution on whole $\mathbb{R}^{\mathbb{Z}^2}$. First, fix $m \geq 1$ and note that by (2.10) and $\mathbb{E}[e^{\alpha^* s(A,B)}] = 1$,

$$\mathbb{P}(\Theta'_{-m,-m} > 0) = \mathbb{E}[\Theta_{m,m}^{\alpha^*}] = 1\,.$$

Further, for arbitrary bounded measurable function $h : \mathbb{R}^{2m+1} \to \mathbb{R}$, using (2.10) and (4.16),

$$\begin{aligned}\mathbb{E}[h(\Theta'_{-m,-m}, \ldots, \Theta'_{m,m})] &= \mathbb{E}[h(\Theta_{m,m}^{-1}(\Theta_{0,0}, \ldots, \Theta_{2m,2m}))\Theta_{m,m}^{\alpha^*}] \\ &= \mathbb{E}\Big[h(e^{-\sum_{k=1}^m \varepsilon_k}, e^{-\sum_{k=2}^m \varepsilon_k}, \ldots, e^{\varepsilon_m}, 1, e^{\varepsilon_{m+1}}, \ldots, e^{\sum_{k=m+1}^{2m} \varepsilon_k}) \prod_{k=1}^m e^{\alpha^* \varepsilon_k}\Big]\,.\end{aligned}$$

By definition of $(\Theta_{k,k})_{k\in\mathbb{Z}}$, this implies that

$$\mathbb{E}[h(\Theta'_{-m,-m}, \ldots, \Theta'_{m,m})] = \mathbb{E}[h(\Theta_{-m,-m}, \ldots, \Theta_{m,m})]\,.$$

□

4.3 Checking assumptions of Theorem 2.3.14

In view of (4.11), define

$$a_n = (Cn^2)^{1/\alpha^*},\ n \in \mathbb{N}\,,$$

so that $\lim_{n\to\infty} n^2 \mathbb{P}(X_{0,0} > a_n) = 1$. The proof of the following result is postponed to Section 4.4.

Proposition 4.3.1. *The random field $\boldsymbol{X}$ satisfies Assumption 2.3.2 for every sequence of positive integers (r_n) such that $\lim_{n\to\infty} r_n = \infty$ and $\lim_{n\to\infty} r_n/n^\epsilon = 0$ for all $\epsilon > 0$.*

Take now sequences of positive integers (l_n) and (r_n) such that

$$\lim_{n\to\infty} \log n/l_n = \lim_{n\to\infty} l_n/r_n = \lim_{n\to\infty} r_n/n^\epsilon = 0$$

for all $\epsilon > 0$ and set $k_n = \lfloor n/r_n \rfloor$. Recall blocks of indices $J_{n,\boldsymbol{i}} \subseteq \{1,\dots,k_n r_n\}^2$ of size r_n^2 from (2.11) and blocks $\boldsymbol{X}_{n,\boldsymbol{i}} := \boldsymbol{X}_{J_{n,\boldsymbol{i}}}$ for $\boldsymbol{i} \in I_n := \{1,\dots,k_n\}^2$. To show that $\boldsymbol{X}_{n,\boldsymbol{i}}$'s satisfy the asymptotic independence condition of Theorem 2.3.14, we will apply Proposition 1.4.5. However, to use it we first need to alter the original blocks.

First, cut off edges of $J_{n,\boldsymbol{i}}$'s by l_n, more precisely, define

$$\tilde{J}_{n,\boldsymbol{i}} := \{(i,j) : (\boldsymbol{i}-\boldsymbol{1})\cdot r_n + \boldsymbol{1} \le (i,j) \le \boldsymbol{i}\cdot r_n - l_n\cdot \boldsymbol{1}\},\ \boldsymbol{i} \in I_n\,.$$

Further, for all $i,j \in \mathbb{Z}$ and $m \in \mathbb{N}$ let $\varepsilon_{i,j}^m$ be the empirical measure on E^2 of the sequence (A_{i-k}, B_{j-k}), $k = 0,\dots,m-1$, i.e.

$$\varepsilon_{i,j}^m = \sum_{k=0}^{m-1} \delta_{(A_{i-k},B_{j-k})}\,.$$

Notice, the score $S_{i,j}^m = \sum_{k=0}^{m-1} s(A_{i-k}, B_{j-k})$ is then equal to the integral of the score function s w.r.t. $\varepsilon_{i,j}^m$. Further, for $\eta > 0$ denote by B_η the set of all probability measures ν on E^2 such that $||\nu - \mu^*|| := \sum_{a,b\in E} |\nu(a,b) - \mu^*(a,b)| < \eta$.

Let $b_n = \log a_n$, then for all $\eta > 0$ and $i,j \in \mathbb{Z}$ define the random variable $\tilde{S}_{i,j} = \tilde{S}_{i,j}(n,\eta)$ by

$$\tilde{S}_{i,j} = \max\{S_{i,j}^m : 1 \le m \le c_0 b_n,\ \varepsilon_{i,j}^m \in B_\eta\} \tag{4.18}$$

with $c_0 > 0$ from Lemma 4.2.2 and $\max\emptyset := 0$, and define modified blocks $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}} = \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}(\eta)$ in $\tilde{l}_0$ by

$$\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}} = (e^{\tilde{S}_{i,j}} : (i,j) \in \tilde{J}_{n,\boldsymbol{i}})\,. \tag{4.19}$$

It turns out that by restricting to the $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$'s one does not lose any relevant information. To understand the role of $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$'s, observe that for any nonnegative and measurable function f on $[0,1]^2 \times \tilde{l}_{0,0}$,

$$\begin{aligned}
&\left|\mathbb{E}\left[e^{-\sum_{\boldsymbol{i}\in I_n} f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\right] - \prod_{\boldsymbol{i}\in I_n} \mathbb{E}\left[e^{-f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\right]\right| \\
&\quad\le \left|\mathbb{E}\left[e^{-\sum_{\boldsymbol{i}\in I_n} f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\right] - \mathbb{E}\left[e^{-\sum_{\boldsymbol{i}\in I_n} f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}\right]\right| \\
&\quad+ \left|\prod_{\boldsymbol{i}\in I_n} \mathbb{E}\left[e^{-f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\right] - \prod_{\boldsymbol{i}\in I_n} \mathbb{E}\left[e^{-f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}\right]\right| \\
&\quad+ \left|\mathbb{E}[e^{-\sum_{\boldsymbol{i}\in I_n} f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}] - \prod_{\boldsymbol{i}\in I_n} \mathbb{E}[e^{-f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}]\right| =: I_1 + I_2 + I_3\,.
\end{aligned} \tag{4.20}$$

Recall now the convergence determining family $\mathcal{F}_0'$ from Section 2.3.6. The proof of the following result is in Section 4.4.

Lemma 4.3.2. *For every $\eta > 0$ and every $f \in \mathcal{F}_0'$, $I_1 + I_2 \to 0$ as $n \to \infty$.*

Remark 4.3.3. In particular, since $I_1 \to 0$ for all $f \in \mathcal{F}_0'$, point processes $\sum_{\boldsymbol{i}\in I_n} \delta_{(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}})}$, which are based on $\tilde{S}_{i,j}$'s, converge in distribution if and only if point processes $\sum_{\boldsymbol{i}\in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}})}$, which are based on $S_{i,j}$'s from (4.7), do, and in that case their limits coincide. Similarly, one can show that the former (and therefore the latter) convergence is equivalent to convergence of point processes of blocks based on nonstationary scores from (4.1). In particular, point process convergence results given below hold even with $S_{i,j}$'s from (4.7) replaced with the ones from (4.1).

By (4.20) and Lemma 4.3.2, to show that $(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)$'s are $AI(\mathcal{F}_0')$, it is sufficient to find at least one $\eta > 0$ such that $I_3 \to 0$ for all $f \in \mathcal{F}_0'$, i.e. that $(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)$'s are $AI(\mathcal{F}_0')$. For that purpose, we apply the following variant of Proposition 1.4.5. Recall, for $\boldsymbol{x} = (x_{\boldsymbol{j}})_{\boldsymbol{j}\in\mathbb{Z}^2} \in \tilde{l}_0$, $\|\boldsymbol{x}\|_\infty = \max_{\boldsymbol{j}\in\mathbb{Z}^2} |x_{\boldsymbol{j}}|$, in particular,

$$\|\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}\|_\infty = \max_{\boldsymbol{j}\in\tilde{J}_{n,\boldsymbol{i}}} |X_{\boldsymbol{j}}| \, , \; \boldsymbol{i} \in I_n \, .$$

Corollary 4.3.4. *Let for each $n \in \mathbb{N}$, $(\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}} : \boldsymbol{i} \in I_n)$ be identically distributed random elements in $\tilde{l}_0$ and (a_n) a sequence such that for all $\epsilon > 0$,*

$$\limsup_{n\to\infty} k_n^2 \mathbb{P}(\|\tilde{\boldsymbol{X}}_{n,\boldsymbol{1}}\|_\infty > a_n\epsilon) < \infty \, . \tag{4.21}$$

If there exists a neighborhood structure $B_n(\boldsymbol{i}) \subseteq I_n$, $n \in \mathbb{N}$, $\boldsymbol{i} \in I_n$, such that, denoting $\|B_n\| = \max_{\boldsymbol{i}\in I_n} |B_n(\boldsymbol{i})|$,

(i) As $n \to \infty$, $\|B_n\|/k_n^2 \to 0$ and for all $\epsilon > 0$,

$$k_n^2 \|B_n\| \max_{\substack{\boldsymbol{i}\in I_n \\ \boldsymbol{i}\neq\boldsymbol{j}\in B_n(\boldsymbol{i})}} \mathbb{P}(\|\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}\|_\infty > a_n\epsilon, \, \|\tilde{\boldsymbol{X}}_{n,\boldsymbol{j}}\|_\infty > a_n\epsilon) \to 0 \, ; \tag{4.22}$$

(ii) For n big enough, $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$ is independent of $\sigma(\tilde{\boldsymbol{X}}_{n,\boldsymbol{j}} : \boldsymbol{j} \notin B_n(\boldsymbol{i}))$ for each $\boldsymbol{i} \in I_n$.

Then the family $((\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n) : n \in \mathbb{N}, \, \boldsymbol{i} \in I_n)$ is $AI(\mathcal{F}_0')$.

Proof. First, observe that for any sequence $\epsilon_m \searrow 0$ sets $K_m' = [0,1]^2 \times \{\boldsymbol{x} \in \tilde{l}_{0,0} : \|\boldsymbol{x}\|_\infty > \epsilon_m\}$, $m \in \mathbb{N}$, form a base for the family of bounded sets of $[0,1]^2 \times \tilde{l}_{0,0}$. Next, regardless of ordering of $I_n = \{1, \dots, k_n\}^2$, $|\tilde{B}_n(\boldsymbol{i})| \le |B_n(\boldsymbol{i})|$ for all $\boldsymbol{i} \in I_n$. Since $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$'s are identically distributed,

$$\begin{aligned} b_{n,1}^m &= \sum_{\boldsymbol{i}\in I_n} \sum_{\boldsymbol{j}\in\tilde{B}_n(\boldsymbol{i})} \mathbb{P}((\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n) \in K_m') \cdot \mathbb{P}((\boldsymbol{j}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{j}}/a_n) \in K_m') \\ &\le k_n^2 \|B_n\| \mathbb{P}(\|\tilde{\boldsymbol{X}}_{n,\boldsymbol{1}}\|_\infty > a_n\epsilon_m)^2 \, . \end{aligned}$$

In view of (4.21), $\limsup_{n\to\infty} b^m_{n,1} \le (const.) \limsup_{n\to\infty} \|B_n\|/k_n^2 = 0$ for all $m \in \mathbb{N}$. Similarly, (4.22) implies that $\lim_{n\to\infty} b^m_{n,2} = 0$ for all $m \in \mathbb{N}$, and by (ii), $b_{n,3}(f) = 0$ for every measurable function $f \ge 0$ on $[0,1]^d \times \tilde{l}_{0,0}$ and n big enough. Applying Proposition 1.4.5 finishes the proof. $\square$

For every $\boldsymbol{i} = (i_1, i_2) \in I_n$ define its neighborhood $B_n(\boldsymbol{i})$ by

$$B_n(\boldsymbol{i}) = \{\boldsymbol{j} = (j_1, j_2) \in I_n : i_1 = j_1 \text{ or } i_2 = j_2\}.$$

Observe, $|B_n(\boldsymbol{i})| = 2k_n - 1$ for all $\boldsymbol{i} \in I_n$ and hence $\lim_{n\to\infty} \|B_n\|/k_n^2 = 0$. Further, by (4.11) for all $\epsilon > 0$,

$$\begin{aligned}\limsup_{n\to\infty} k_n^2 \mathbb{P}(\|\tilde{\boldsymbol{X}}_{n,\boldsymbol{1}}\|_\infty > a_n\epsilon) &\le \limsup_{n\to\infty} k_n^2 r_n^2 \mathbb{P}(e^{\tilde{S}_{0,0}} > a_n\epsilon) \\ &\le \limsup_{n\to\infty} k_n^2 r_n^2 \mathbb{P}(X_{0,0} > a_n\epsilon) = \epsilon^{-\alpha^*} < \infty .\end{aligned}$$

Next, recall that $S^m_{i,j} = \sum_{k=0}^{m-1} s(A_{i-k}, B_{j-k})$ so by (4.18), for every $n \in \mathbb{N}$,

$$\tilde{S}_{i,j} \in \sigma(A_{i-\lfloor c_0 b_n\rfloor+1}, \ldots, A_i, B_{j-\lfloor c_0 b_n\rfloor+1}, \ldots, B_j).$$

By construction of $\tilde{J}_{n,\boldsymbol{i}}$'s and choice of (l_n) such that, in particular, $\lim_{n\to\infty} c_0 b_n/l_n = \lim_{n\to\infty} l_n/r_n = 0$, this implies that, for n large enough, $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$ and blocks $(\tilde{\boldsymbol{X}}_{n,\boldsymbol{j}} : \boldsymbol{j} \notin B_n(\boldsymbol{i}))$ are constructed from completely different sets of A_k's and B_k's, and therefore independent.

Further, when $\boldsymbol{j} \in B_n(\boldsymbol{i})$, $\boldsymbol{j} \neq \boldsymbol{i}$, arbitrary scores $S^m_{i,j}$ and $S^l_{i',j'}$ which build blocks $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$ and $\tilde{\boldsymbol{X}}_{n,\boldsymbol{j}}$, respectively (i.e. $(i,j) \in \tilde{J}_{n,\boldsymbol{i}}, (i',j') \in \tilde{J}_{n,\boldsymbol{j}}$ and $1 \le m, l \le c_0 b_n$), for n large enough, depend on completely different sets of variables from at least one of the sequences (A_k) or (B_k). Thus, the following result, which is a special case of [Han06, Corollary 5.4], applies. It follows from [Han06, Lemma 5.3] under condition (12) in [Han06], which, when (A_i) and (B_i) are i.i.d. sequences, is equivalent to Assumption 4.1.2, see [Han06, Remark 3.8].

Lemma 4.3.5 ([Han06, Corollary 5.4]). *There exist constants $\epsilon_2, \eta > 0$ such that for all $u > 0$*

$$\mathbb{P}(S^m_{0,0} > u, S^l_{i,j} > u,\ \varepsilon^m_{0,0}, \varepsilon^l_{i,j} \in B_\eta) \le e^{-(3/2+\epsilon_2)\alpha^* u}$$

uniformly over all $i, j \in \mathbb{Z}$ and $m, l \in \mathbb{N}$ such that $\min\{i,j\} < -m+1$ or $\max\{i-l, j-l\} > 0$.

For completeness, in Section 4.4 we present a (lengthy) proof of the previous result by following the lines of [Han06, Lemma 5.3, Remark 3.8]. For a different (and, in this i.i.d. setting, probably better) argument, cf. Dembo et al. [DKZ94b, pp. 2032–2033]. Note, the fact that E is finite is here exploited.

Take now $\eta > 0$ from the previous result and recall the corresponding $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$'s. For n big enough and every $\epsilon > 0$ we get that

$$k_n^2 \|B_n\| \max_{\substack{\boldsymbol{i} \in I_n \\ \boldsymbol{i} \neq \boldsymbol{j} \in B_n(\boldsymbol{i})}} \mathbb{P}(\|\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}\|_\infty > a_n\epsilon,\ \|\tilde{\boldsymbol{X}}_{n,\boldsymbol{j}}\|_\infty > a_n\epsilon)$$
$$\leq k_n^2 2k_n r_n^4 (c_0 b_n)^2 e^{-(3/2+\epsilon_2)\alpha^*(b_n + \log\epsilon)} \sim (const.)\, n^3 r_n b_n^2 n^{-3-2\epsilon_2} \to 0\,,$$

as $n \to \infty$, by the choice of (r_n) and since $b_n \sim 2\log n/\alpha^*$.

Hence by Corollary 4.3.4, for this η, blocks $\tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}$, and therefore the original blocks $\boldsymbol{X}_{n,\boldsymbol{i}}$, satisfy the asymptotic independence condition. We can now apply Theorem 2.3.14.

Theorem 4.3.6. *Under Assumptions 4.1.1 and 4.1.2, for any sequence of positive integers (r_n) such that $\lim_{n\to\infty} r_n = \infty$ and $\lim_{n\to\infty} r_n/n^\epsilon \to 0$ for all $\epsilon > 0$,*

$$\sum_{\boldsymbol{i} \in I_n} \delta_{(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/(Cn^2)^{1/\alpha^*})} \xrightarrow{d} \sum_{k \in \mathbb{N}} \delta_{(\boldsymbol{T}_k, P_k(Q_{i,j}^k)_{i,j\in\mathbb{Z}})}$$

in $\mathcal{M}_p([0,1]^2 \times \tilde{l}_{0,0})$ where the limit is described in Theorem 2.3.14, with $\vartheta > 0$ given by (4.13) and distribution of $\boldsymbol{Q}$ in $\tilde{l}_0$ determined by (4.14).

In particular, application of Corollary 2.3.15 yields that

$$\sum_{i,j=1}^{n} \delta_{((i,j)/n, X_{i,j}/(Cn^2)^{1/\alpha^*})} \xrightarrow{d} \sum_{k\in\mathbb{N}} \sum_{i,j\in\mathbb{Z}} \delta_{(\boldsymbol{T}_k, P_k Q_{i,j}^k)} \tag{4.23}$$

in $\mathcal{M}_p([0,1]^2 \times (0,\infty))$. Consider now space $\mathcal{M}_p([0,1]^2 \times \mathbb{R})$ with a set $B \subseteq [0,1]^2 \times \mathbb{R}$ being bounded if $B \subseteq [0,1]^2 \times (x,\infty)$ for some $x \in \mathbb{R}$. It is easy to see that

$$\sum_{k\in\mathbb{N}} \delta_{(\boldsymbol{t}_k, x_k)} \mapsto \sum_{k\in\mathbb{N}} \delta_{(\boldsymbol{t}_k, \log(x_k C^{1/\alpha^*}))}$$

is a well–defined mapping from $\mathcal{M}_p([0,1]^2 \times (0,\infty))$ to $\mathcal{M}_p([0,1]^2 \times \mathbb{R})$ which is also continuous with respect to vague topologies on these spaces (see e.g. [Res87, Proposition 3.18]). Theorem 4.1.3 now follows easily from (4.23) via continuous mapping theorem and standard Poisson process transformation (see e.g. [Res87, Proposition 3.7]).

4.4 Postponed proofs

Proof of Lemma 4.2.2

By Markov inequality, for any $\lambda \geq 0$ and all $u > 0$

$$\mathbb{P}\left(\max_{m > c_0 u} S_{0,0}^m \geq 0\right) \leq \sum_{l=0}^{\infty} \mathbb{P}\left(S_{0,0}^{\lceil c_0 u\rceil + l} \geq 0\right) \leq \sum_{l=0}^{\infty} \mathbb{E}\left[e^{\lambda S_{0,0}^{\lceil c_0 u\rceil + l}}\right] = \sum_{l=0}^{\infty} m(\lambda)^{\lceil c_0 u\rceil + l}\,,$$

where $m(\lambda) = \mathbb{E}[e^{\lambda s(A,B)}]$ is the moment generating function of $s(A,B)$. Fix any $0 < \lambda_0 < \alpha^*$. By strict convexity of m and $m(\alpha^*) = 1$, $0 < m(\lambda_0) < 1$ and in particular

$$\mathbb{P}\left(\max_{m \geq c_0 u} S_{0,0}^m \geq 0\right) \leq e^{c_0 u \log m(\lambda_0)} \sum_{l=0}^{\infty} m(\lambda_0)^l.$$

Since the series above is summable, taking c_0 strictly larger than $-2\alpha^*/\log m(\lambda_0)$ finishes the proof.

Proof of Lemma 4.2.3

Fix an $u > 0$ and take $m, l \geq 0$ arbitrary. Note, since E is finite, $\|s\| := \sup_{(a,b)\in E^2} |s(a,b)| < \infty$, and in particular, there is nothing to prove if m or l is smaller than $u/\|s\|$. Assume therefore that $m \wedge l \geq u/\|s\|$.

Denote by δ_A and δ_B the number of variables of the sequence (A_k) and (B_k), respectively, which appear both in $S_{0,0}^m$ and $S_{i,j}^l$. Also, fix any $0 < \gamma < 1/2$.

Case 1. Assume that $\delta_A \vee \delta_B \leq \gamma u/\|s\|$. In this case $S_{0,0}^m$ consist of at least $(1-2\gamma)u/\|s\|$ summands $s(A_k, B_k)$ which are independent of $S_{i,j}^l$. Let $S_{0,0}^*$ be the sum of those $s(A_k, B_k)$'s. Therefore,

$$\begin{aligned} \mathbb{P}(S_{0,0}^m > u, S_{i,j}^l > u) &\leq \mathbb{P}(S_{0,0}^* \geq (1-2\gamma)u, S_{i,j}^l > u) \\ &= \mathbb{P}(S_{0,0}^* \geq (1-2\gamma)u)\mathbb{P}(S_{i,j}^l > u) \leq e^{-(2-2\gamma)\alpha^* u}\,. \end{aligned} \tag{4.24}$$

Case 2. Assume w.l.o.g. that $\delta_A > \gamma u/\|s\|$. Decompose $S_{0,0}^m = S_1 + S_2$ and $S_{i,j}^l = \tilde{S}_2 + S_3$ where S_2 and $\tilde{S}_2$, respectively, is the sum of scores $s(A_k, B_k)$ and $s(A_k, B_{k+(j-i)})$, respectively, for δ_A number of k's. In particular,

$$\mathbb{P}(S_{0,0}^m > u, S_{i,j}^l > u) \leq \mathbb{P}(S_{0,0}^m > u, \tilde{S}_2 \geq S_2) + \mathbb{P}(S_{i,j}^l > u, S_2 \geq \tilde{S}_2)\,. \tag{4.25}$$

A change of measure yields

$$\mathbb{P}(S_{0,0}^m > u, \tilde{S}_2 \geq S_2) \leq e^{-\alpha^* u}\mathbb{P}^*(\tilde{S}_2 \geq S_2),$$

where, under $\mathbb{P}^*$, pairs (A_k, B_k) from S_2 are i.i.d. and distributed according to μ^*, and pairs $(A_k, B_{k+(j-i)})$ from $\tilde{S}_2$ are independent and distributed according to $\mu_A^* \times \mu_B^*$ or $\mu_A^* \times \mu_B$. The key thing now is that, whenever $\mu^* \neq \mu_A^* \times \mu_B^*$ (which holds under (4.5)),

$$\mathbb{E}_{\nu_A \times \nu_B}[s(A,B)] - \mathbb{E}_{\mu^*}[s(A,B)] < 0 \tag{4.26}$$

for all $\nu_A \in \{\mu_A, \mu_A^*\}$ and $\nu_B \in \{\mu_B, \mu_B^*\}$, where $\mathbb{E}_\mu$ denotes expectation assuming (A,B) is distributed according to μ, see [DKZ94a, beginning of Section 3]. An application of [Han06, Theorem 5.7] (with $\eta = 6\|s\|$) which is a consequence of the Azuma-Hoeffding inequality, together with $\delta_A > \gamma u/\|s\|$, yields that $\mathbb{P}^*(\tilde{S}_2 \geq S_2) \leq e^{-\epsilon\alpha^* u}$ with $\epsilon = \epsilon(\gamma) = \zeta^2\gamma/(\alpha^* 72\|s\|^3)$ where $\zeta < 0$ is the maximum in (4.26). By analogous arguments, one obtains the same bound for the second term in (4.25). Together with (4.24) and optimizing over γ, this yields the result with $\epsilon_0 = \max_{0<\gamma<1/2} \min\{1-2\gamma, \epsilon(\gamma)\}$.

Proof of Proposition 4.3.1

Let (r_n) be an arbitrary sequence of positive integers satisfying $r_n \to \infty$ and $r_n/n^\epsilon \to 0$ for all $\epsilon > 0$. We have to show that for an arbitrary $u > 0$

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{m<|(i,j)|\leq r_n} X_{i,j} > a_n u \,\middle|\, X_{0,0} > a_n u \right) = 0 \,. \tag{4.27}$$

We deal with the diagonal elements using arguments from [Bas00, Lemma 4.1.4]. First, notice that by (4.9), for each $k \geq 1$ we can decompose

$$X_{k,k} = \max\{e^{\max_{0\leq l\leq k} S^l_{k,k}}, X_{0,0} e^{S^k_{k,k}}\} \,,$$

with $(S^l_{k,k})_{0\leq l\leq k}$ being independent of $X_{0,0}$. Hence, using stationarity,

$$\mathbb{P}\left(\max_{m<|k|\leq r_n} X_{k,k} > a_n u \,\middle|\, X_{0,0} > a_n u \right) \leq 2 \sum_{k=m+1}^{r_n} \mathbb{P}(X_{k,k} > a_n u \mid X_{0,0} > a_n u)$$
$$\leq 2r_n \mathbb{P}(e^{\max_{0\leq l\leq r_n} S^l_{0,0}} > a_n u) + 2 \sum_{k=m+1}^{r_n} \mathbb{P}(X_{0,0} e^{S^k_{k,k}} > a_n u \mid X_{0,0} > a_n u).$$

Since $r_n/n^2 \to 0$, the choice of (a_n) and (4.11) imply that

$$2r_n \mathbb{P}(e^{\max_{0\leq l\leq r_n} S^l_{0,0}} > a_n u) \leq 2r_n \mathbb{P}(X_{0,0} > a_n u) \to 0\,, \text{ as } n \to \infty\,.$$

For the second term, take an arbitrary $0 < \lambda_0 < \alpha^*$ so in particular $0 < m(\lambda_0) = \mathbb{E}[e^{\lambda_0 s(A,B)}] < 1$ by strict convexity of m. Apply Markov's inequality and use independence between $X_{0,0}$ and $S^k_{k,k}$ to obtain

$$\sum_{k=m+1}^{r_n} \mathbb{P}(X_{0,0} e^{S^k_{k,k}} > a_n u \mid X_{0,0} > a_n u) \leq \frac{\mathbb{E}[X_{0,0}^{\lambda_0} \mathbb{1}\{X_{0,0} > a_n u\}]}{(a_n u)^{\lambda_0} \mathbb{P}(X_{0,0} > a_n u)} \sum_{k=m+1}^{r_n} m(\lambda_0)^k \,.$$

Variant of Karamata's theorem (see [BDM16, Appendix B.4], also [BGT87, pp. 26–28]) now implies that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{m<|k|\leq r_n} X_{k,k} > a_n u \,\middle|\, X_{0,0} > a_n u \right) \leq \frac{\alpha^*}{\alpha^* - \lambda_0} \lim_{m\to\infty} \sum_{k=m+1}^{\infty} m(\lambda_0)^k = 0\,.$$

It remains to deal with the non diagonal terms. More precisely, in order to obtain

(4.27), we will show that, denoting $b_n = \log a_n$ and $M = \log y$,

$$\begin{aligned}&\limsup_{n\to\infty} \mathbb{P}\left(\max_{|(i,j)|\le r_n,\, i\neq j} S_{i,j} > b_n + M \,\middle|\, S_{0,0} > b_n + M \right) \\ &\qquad = C^{-1} e^{\alpha^* M} \limsup_{n\to\infty} e^{\alpha^* b_n} \mathbb{P}\left(\max_{|(i,j)|\le r_n,\, i\neq j} S_{i,j} > b_n + M,\, S_{0,0} > b_n + M \right) = 0\,.\end{aligned}$$

Notice that $e^{\alpha^* b_n} = Cn^2$. First, since $r_n/n \to 0$, stationarity and Lemma 4.2.2 give

$$\begin{aligned}\limsup_{n\to\infty} e^{\alpha^* b_n} \mathbb{P}\left(\max_{\substack{|(i,j)|\le r_n \\ k > c_0 b_n}} S^k_{i,j} \ge 0 \right) &\le \limsup_{n\to\infty} e^{\alpha^* b_n} (2r_n+1)^2 \mathbb{P}\left(\max_{k > c_0 b_n} S^k_{0,0} \ge 0 \right) \\ &\le \limsup_{n\to\infty} \frac{(2r_n+1)^2}{Cn^2} = 0\,.\end{aligned}$$

Now by Lemma 4.2.3 there exist an $\epsilon_0 > 0$ such that

$$\begin{aligned}&\limsup_{n\to\infty} e^{\alpha^* b_n} \mathbb{P}\left(\max_{|(i,j)|\le r_n,\, i\neq j} S_{i,j} > b_n + M,\, S_{0,0} > b_n + M \right) \\ &= \limsup_{n\to\infty} e^{\alpha^* b_n} \mathbb{P}\left(\max_{\substack{|(i,j)|\le r_n,\, i\neq j \\ 1\le l \le c_0 b_n}} S^l_{i,j} > b_n + M,\, \max_{1\le k\le c_0 b_n} S^k_{0,0} > b_n + M \right) \\ &\le \limsup_{n\to\infty} e^{\alpha^* b_n} (2r_n+1)^2 (c_0 b_n)^2 2 e^{-(1+\epsilon')\alpha^* b_n} \\ &= 2c_0^2 C^{-(1+\epsilon)} \limsup_{n\to\infty} \left(\frac{2r_n+1}{n^{\epsilon_0/2}} \right)^2 \left(\frac{b_n}{n^{\epsilon_0/2}} \right)^2 = 0\,,\end{aligned}$$

where the last equality follows by the choice of (r_n) and since $b_n \sim \frac{2}{\alpha^*}\log n$.

Proof of Lemma 4.3.2

First, we need the following simple result proved by a change of measure argument and a large deviation bound for empirical measures, cf. the proof of [Han06, Lemma 5.14, Equation (54)].

Lemma 4.4.1. *For all $\eta > 0$ there exists an $\epsilon_1 > 0$ such that*

$$\lim_{u\to\infty} e^{(1+\epsilon_1)\alpha^* u} \sup_{m\ge 1} \mathbb{P}(S^m_{0,0} > u, \varepsilon^m_{0,0} \notin B_\eta) = 0\,.$$

Proof. Fix $\eta > 0$ and denote $A_m(u) = \{S^m_{0,0} > u, \varepsilon^m_{\mathbf{0}} \notin B_\eta\}$ for $m \ge 1$ and $u > 0$. Note that, since $S^m_{\mathbf{0}} = \sum_{k=0}^{m-1} s(A_{-k}, B_{-k})$, $\mathbb{P}(A_m(u)) = 0$ whenever $m \le u/\|s\|$, so for fixed $u > 0$ we only need to deal with $\mathbb{P}(A_m(u))$ for $m > u/\|s\|$.

First, a change of measure yields

$$\mathbb{P}(A_m(u)) = \mathbb{E}\left[\frac{\exp(\alpha^* S_{0,0}^m)}{\exp(\alpha^* S_{0,0}^m)}\mathbb{1}_{A_m(u)}\right] \le e^{-\alpha^* u}\mathbb{P}^*(\varepsilon_{0,0}^m \notin B_\eta)\,,$$

where $\mathbb{P}^*$ makes $(A_{-k}, B_{-k}), k = 0, \dots, m-1$, i.i.d. elements of E^2 with common distribution μ^*. By Sanov's theorem (see [DZ10, Theorem 2.1.10])

$$\limsup_{m\to\infty} \frac{1}{m} \log \mathbb{P}^*(\varepsilon_{0,0}^m \notin B_\eta) \le -\inf_{\pi\notin B_\eta} H(\pi \mid \mu^*)\,.$$

Since, for a sequence of probability measures (π_n) on E^2, $H(\pi_n \mid \mu^*) \to 0$ implies that $\|\pi_n - \mu^*\| \to 0$, for we can find a constant $c = c(\eta) > 0$ such that $\inf_{\pi\notin B_\eta} H(\pi \mid \mu^*) > c$. Hence, for all $m > u/\|s\|$ with u large enough

$$\mathbb{P}^*(\varepsilon_{0,0}^m \notin B_\eta) \le e^{-mc} \le e^{-uc/\|s\|}\,.$$

To finish the proof, it suffices to take $\epsilon_1 := \frac{c}{\|s\|\alpha^*} > 0$. □

Proof of Lemma 4.3.2. Take an arbitrary $f \in \mathcal{F}_0' \subseteq CB_b^+([0,1]^2 \times \tilde{l}_{0,0})$ and let $\epsilon > 0$ be such that $f(\boldsymbol{t}, (x_{i,j})_{i,j}) = f(\boldsymbol{t}, (x_{i,j}\mathbb{1}_{\{|x_{i,j}|>\epsilon\}})_{i,j})$ for all $\boldsymbol{t} \in [0,1]^2$ and $(x_{i,j})_{i,j\in\mathbb{Z}} \in \tilde{l}_{0,0}$ with $f(\boldsymbol{t}, \boldsymbol{0}) = 0$.

By the elementary inequality $|\prod_{i=1}^k a_i - \prod_{i=1}^k b_i| \le \sum_{i=1}^k |a_i - b_i|$ valid for all $k \ge 1$ and $a_i, b_i \in [0,1]$ (see e.g. [Dur10, Lemma 3.4.3]),

$$\begin{aligned}\Big|\mathbb{E}\left[e^{-\sum_{\boldsymbol{i}\in I_n} f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\right] - \mathbb{E}\Big[e^{-\sum_{\boldsymbol{i}\in I_n} f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}\Big]\Big| \\ + \Big|\prod_{\boldsymbol{i}\in I_n} \mathbb{E}\left[e^{-f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)}\right] - \prod_{\boldsymbol{i}\in I_n} \mathbb{E}\Big[e^{-f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}\Big]\Big| \\ \le 2\sum_{\boldsymbol{i}\in I_n} \mathbb{E}\Big|e^{-f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)} - e^{-f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}\Big|\,. \end{aligned} \tag{4.28}$$

Further, denote by $J_{r_n} := \{1, \dots, r_n\}^2 = J_{n,\boldsymbol{1}}$ and $\tilde{J}_{r_n} := \{1, \dots, r_n - l_n\}^2 = \tilde{J}_{n,\boldsymbol{1}}$. Using stationarity we get that

$$\sum_{\boldsymbol{i}\in I_n} \mathbb{E}\Big|e^{-f(\boldsymbol{i}/k_n, \boldsymbol{X}_{n,\boldsymbol{i}}/a_n)} - e^{-f(\boldsymbol{i}/k_n, \tilde{\boldsymbol{X}}_{n,\boldsymbol{i}}/a_n)}\Big| \le k_n^2(A_1 + A_2 + A_3)\,, \tag{4.29}$$

where

$$\begin{aligned} A_1 &= \mathbb{P}(X_{i,j} > a_n\epsilon \;\text{ for some } (i,j) \in J_{r_n} \setminus \tilde{J}_{r_n})\,, \\ A_2 &= \mathbb{P}(\max_{m>c_0 b_n} e^{S_{i,j}^m} > a_n\epsilon \;\text{ for some } (i,j) \in \tilde{J}_{r_n})\,, \\ A_3 &= \mathbb{P}(e^{S_{i,j}^m} > a_n\epsilon \text{ and } \varepsilon_{i,j}^m \notin B_\eta \;\text{ for some } (i,j) \in \tilde{J}_{r_n},\, 1 \le m \le c_0 b_n)\,. \end{aligned}$$

Observe, $|J_{r_n} \setminus \tilde{J}_{r_n}| \le 2r_n l_n$ and $|\tilde{J}_{r_n}| \le r_n^2$, and recall that $k_n r_n \sim n$ as $n \to \infty$, so using stationarity and then (4.11), Lemma 4.2.2 and Lemma 4.4.1, respectively,

$$\begin{aligned}
\limsup_{n\to\infty} k_n^2 A_1 &\le \limsup_{n\to\infty} 2k_n^2 r_n l_n \mathbb{P}(X_{0,0} > a_n \epsilon) = (const.) \limsup_{n\to\infty} l_n / r_n = 0\,, \\
\limsup_{n\to\infty} k_n^2 A_2 &\le \limsup_{n\to\infty} k_n^2 r_n^2 \mathbb{P}(\max_{m > c_0 b_n} S_{0,0}^m \ge 0) \le \limsup_{n\to\infty} n^{-2} = 0\,, \\
\limsup_{n\to\infty} k_n^2 A_3 &\le \limsup_{n\to\infty} k_n^2 r_n^2 c_0 b_n \sup_{m\ge 1} \mathbb{P}(S_{0,0}^m > b_n + \log\epsilon,\, \varepsilon_{i,j}^m \notin B_\eta) \\
&\le (const.) \limsup_{n\to\infty} b_n / n^{2\epsilon_1} = 0\,.
\end{aligned}$$

Therefore, the right hand side, and then also the left hand side, of (4.29) tends to 0 as $n \to \infty$, and by (4.28) this proves the lemma. $\square$

Proof of Lemma 4.3.5

Recall, $S_{0,0}^m = \sum_{k=0}^{m-1} s(A_{-k}, B_{-k})$ and $S_{i,j}^l = \sum_{k=0}^{l-1} s(A_{i-k}, B_{j-k})$. Assume that $i, j \in \mathbb{Z}$ and $l, m \in \mathbb{N}$ are such that $j < -m+1$ or $j-l > 0$ and set $m' = |[-m+1, 0] \cap [i-l+1, i]| \ge 0$.

In words, scores $S_{0,0}^m$ and $S_{i,j}^l$ depend on completely different sets of B–variables and exactly m' same A–variables are used in both sums. The case when only B–variables are shared is completely analogous and therefore omitted.

For all $\eta, u > 0$ denote by $A(u, \eta)$ the event $\{S_{0,0}^m > u, S_{i,j}^l > u,\ \varepsilon_{0,0}^m, \varepsilon_{i,j}^l \in B_\eta\}$ and for a function $g : E^2 \to \mathbb{R}$, by a slight abuse of notation, denote

$$m(g) = \mathbb{E}[e^{g(A_0, B_0)}],\ m_1(g) = \mathbb{E}[e^{g(A_0,B_0) + g(A_0,B_1)}]\,.$$

Also, for any measure μ on E^2 and function $g : E^2 \to \mathbb{R}$, denote $\mu(g) = \sum_{a,b\in E} g(a,b)\mu(a,b)$.

Step 1. ([Han06, Lemma 5.3]) Take arbitrary $g : E^2 \to \mathbb{R}$ and $\epsilon > 0$. Using a change of measure argument, we show that there exist a constant $\eta = \eta(g, \epsilon) > 0$ such that for all $u > 0$

$$\mathbb{P}(A(u,\eta)) \le \exp\left(-u\left(\frac{2\mu^*(g) - \log m_1(g)}{\mu^*(s)} - \epsilon\right)\right)\,. \tag{4.30}$$

First, observe that $l\varepsilon_{i,j}^l(g) = \sum_{k=1}^l g(A_{i-k}, B_{j-k})$, in particular $l\varepsilon_{i,j}^l(s) = S_{i,j}^l$. Further, define a random variable L_g by

$$\begin{aligned}
L_g &= \frac{\exp(m\varepsilon_{0,0}^m(g) + l\varepsilon_{i,j}^l(g))}{\mathbb{E}\left[\exp(m\varepsilon_{0,0}^m(g) + l\varepsilon_{i,j}^m(g))\right]} \\
&= \exp\left(m\varepsilon_{0,0}^m(g) + l\varepsilon_{i,j}^l(g) - (m+l-2m')\log m(g) - m'\log m_1(g)\right)\,.
\end{aligned}$$

Independence of A_0, B_0 and B_1, Fubini's theorem and Jensen's inequality, imply that

$m_1(g) \geq m(g)^2$ and this further implies that

$$L_g \geq \exp\left(m\varepsilon_{0,0}^m(g) + l\varepsilon_{i,j}^l(g) - (m+l)\log m_1(g)/2\right) . \tag{4.31}$$

For every $\epsilon' > 0$ take an $\eta' > 0$ such that $\pi \in B_{\eta'}$ implies that

$$|\pi(g) - \mu^*(g)| \vee |\pi(s) - \mu^*(s)| \leq \epsilon' .$$

Then, on $A(u, \eta')$

$$m\varepsilon_{0,0}^m(g) + l\varepsilon_{i,j}^l(g) \geq (m+l) \cdot (\mu^*(g) - \epsilon')$$

and further, $m + l > 2u/(\mu^*(s) + \epsilon')$ (recall that $\mu^*(s) > 0$ under (4.5)). Hence, by (4.31), on $A(u, \eta')$

$$L_g \geq \exp\left(u\frac{2(\alpha^*(g) - \epsilon') - \log m_1(g)}{\mu^*(s) + \epsilon'}\right) .$$

Since $\mathbb{E}[L_g] = 1$ this gives the bound

$$\mathbb{P}(A(u, \eta')) = \mathbb{E}\left[\frac{L_g}{L_g} \cdot \mathbb{1}_{A(u,\eta')}\right] \leq \exp\left(-u\frac{2(\mu^*(g) - \epsilon') - \log m_1(g)}{\mu^*(s) + \epsilon'}\right) . \tag{4.32}$$

Note that if g and ϵ are such that $2\mu^*(g) - \log m_1(g) \leq \epsilon\mu^*(s)$, (4.30) is trivial. Otherwise, (4.30) follows from (4.32) by taking $\epsilon' > 0$ such that

$$\frac{2\mu^*(g) - \log m_1(g)}{\mu^*(s)} - \epsilon = \frac{2(\mu^*(g) - \epsilon') - \log m_1(g)}{\mu^*(s) + \epsilon'}$$

and setting $\eta = \eta'$.

Step 2. ([Han06, Remark 3.8]) To finish the proof, it suffices to find a function g and $\epsilon > 0$ for which

$$\frac{\mu^*(g) - \log m_1(g)}{\mu^*(s)} - \epsilon \geq \left(\frac{3}{2} + \epsilon\right)\alpha^* \tag{4.33}$$

We show that this is always possible under Assumption (4.1.2).

First, since finite measures on E^2 can be considered as elements of the finite–dimensional space $\mathbb{R}^{E^2}$, by Cramér's theorem (see e.g. [DZ10, Theorem 2.2.30]), the sequence

$$\frac{1}{n}\sum_{i=0}^{n-1}\left[\delta_{(A_i,B_i)} + \delta_{(A_i,B_{i+n})}\right] , \; n \in \mathbb{N}$$

of random elements in $\mathbb{R}^{E^2}$ satisfies the large deviations principle with the rate function I_1 satisfying

$$I_1(2\mu^*) = \sup_{g\in\mathbb{R}^{E^2}} (2\mu^*(g) - \log\mathbb{E}[e^{g(A_0,B_0)+g(A_0,B_n)}]) = \sup_{g\in\mathbb{R}^{E^2}} (2\mu^*(g) - \log m_1(g)) .$$

On the other hand, by Sanov's theorem (see [DZ10, Theorem 2.1.10]), the sequence

$$\frac{1}{n}\sum_{i=0}^{n-1}\delta_{(A_i,B_i,B_{i+n})}\;,\; n\in\mathbb{N}$$

of empirical measures on E^3 satisfies the large deviations principle with the rate function $H(\,\cdot\mid\mu_A\times\mu_B^2)$.

The contraction principle (see [DZ10, Theorem 4.2.1]) used with the continuous function $h(\pi)(a,b)=\sum_c[\pi(a,b,c)+\pi(a,c,b)]$ for $a,b\in E^2$ and probability measures π on E^3, and uniqueness of the rate function (see [DZ10, Lemma 4.1.4]) imply that

$$I_1(2\mu^*)=\inf\{H(\pi\mid\mu_A\times\mu_B^2):\pi_{1,2}+\pi_{1,3}=2\mu^*\}\,,$$

where the infimum is taken over all probability measures π on E^3 with $\pi_{1,2}$ and $\pi_{1,3}$ being measures on E^2 given by $\pi_{1,2}(a,b):=\sum_{c\in E}\pi(a,b,c)$ and $\pi_{1,3}(a,b):=\sum_{c\in E}\pi(a,c,b)$ for $a,b\in E$. We now show that the infimum above is obtained for the measure π^* defined by

$$\pi^*(a,b,c)=\frac{\mu^*(a,b)\cdot\mu^*(a,c)}{\mu_A^*(a)}>0\,,\;\; a,b,c\in E\,.$$

First, observe that $\pi^*_{1,2}=\pi^*_{1,3}=\mu^*$. Next,

$$\begin{aligned}H(\pi^*\mid\mu_A\times\mu_B^2)&=\sum_{a,b,c\in E}\frac{\mu^*(a,b)\cdot\mu^*(a,c)}{\mu_A^*(a)}\log\left(\frac{\mu_A(a)e^{\alpha^*(s(a,b)+s(a,c))}}{\mu_A^*(a)}\right)\\&=2\alpha^*\mu^*(s)-H(\mu_A^*\mid\mu_A)\,.\end{aligned}$$

On the other hand, any probability measure π on E^3 satisfying $\pi_{1,2}+\pi_{1,3}=2\mu^*$ necessarily satisfies

$$\pi_1(a):=\sum_{b,c\in E}\pi(a,b,c)=\frac{1}{2}\sum_{b\in E}\pi_{1,2}(a,b)+\frac{1}{2}\sum_{c\in E}\pi_{1,3}(a,c)=\sum_{b\in E}\mu^*(a,b)=\mu_A^*(a)\,,$$

for all $a\in E$. Using this we get that for all such π,

$$H(\pi\mid\mu_A\times\mu_B^2)=H(\pi\mid\pi^*)+2\alpha^*\mu^*(s)-H(\mu_A^*\mid\mu_A)\geq H(\pi^*\mid\mu_A\times\mu_B^2)\,,$$

and hence

$$\sup_{g\in\mathbb{R}^{E^2}}(2\mu^*(g)-\log m_1(g))=I_1(2\mu^*)=2\alpha^*\mu^*(s)-H(\mu_A^*\mid\mu_A)\,.$$

By definition of μ^* one can easily show that $\alpha^*\mu^*(s)=H(\mu^*\mid\mu_A\times\mu_B)$, so we conclude that under (4.1.2),

$$\sup_{g \in \mathbb{R}^{E^2}} (2\mu^*(g) - \log m_1(g)) = 2H(\mu^* \mid \mu_A \times \mu_B) - H(\mu^*_A \mid \mu_A)$$
$$> \frac{3}{2} H(\mu^* \mid \mu_A \times \mu_B) = \frac{3}{2}\alpha^* \mu^*(s)\,.$$

Hence, for $\epsilon > 0$ small enough there exists a function g on E^2 such that

$$\frac{2\mu^*(g) - \log m_1(g)}{\mu^*(s)} - \epsilon \geq \left(\frac{3}{2} + \epsilon\right)\alpha^*\,,$$

i.e. (4.33) holds.

Bibliography

[AGG89] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[Arr98] R. Arratia. On the central role of scale invariant Poisson processes on $(0, \infty)$. In *Microsurveys in discrete probability (Princeton, NJ, 1997)*, volume 41 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 21–41. Amer. Math. Soc., Providence, RI, 1998.

[Asm03] S. Asmussen. *Applied probability and queues*, volume 51 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 2003. Stochastic Modelling and Applied Probability.

[AT92] F. Avram and M. S. Taqqu. Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.*, 20(1):483–503, 1992.

[Ban80] R. Banys. On superpositions of random measures and point processes. In *Mathematical statistics and probability theory (Proc. Sixth Internat. Conf., Wisła, 1978)*, volume 2 of *Lecture Notes in Statist.*, pages 26–37. Springer, New York-Berlin, 1980.

[Bas00] B. Basrak. *The sample autocorrelation function of non-linear time series*. PhD thesis, Rijksuniversiteit Groningen Groningen, Netherlands, 2000.

[BDM16] D. Buraczewski, E. Damek, and T. Mikosch. *Stochastic models with power-law tails*. Springer Series in Operations Research and Financial Engineering. Springer, 2016. The equation $X = AX + B$.

[Ber88] S. K. Berberian. Borel spaces, 1988. URL: `https://web.ma.utexas.edu/mp_arc/c/02/02-156.pdf`.

[BGT87] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1987.

[Big03] J. D. Biggins. Random walk conditioned to stay positive. *J. London Math. Soc. (2)*, 67(1):259–272, 2003.

[Bil68] P. Billingsley. *Convergence of probability measures.* New York, Wiley, 1968.

[BJL16] R. Balan, A. Jakubowski, and S. Louhichi. Functional convergence of linear processes with heavy-tailed innovations. *J. Theoret. Probab.*, 29(2):491–526, 2016.

[BK14] B. Basrak and D. Krizmanić. A limit theorem for moving averages in the α-stable domain of attraction. *Stochastic Processes and their Applications*, 124(2):1070–1083, 2014.

[BKS12] B. Basrak, D. Krizmanić, and J. Segers. A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.*, 40(5):2008–2033, 2012.

[BP18a] B. Basrak and H. Planinić. A note on vague convergence of measures. *arXiv e-prints*, page arXiv:1803.07024, March 2018.

[BP18b] B. Basrak and H. Planinić. Compound Poisson approximation for random fields with application to sequence alignment. *ArXiv e-prints*, September 2018.

[BPS18] B. Basrak, H. Planinić, and P. Soulier. An invariance principle for sums and record times of regularly varying stationary sequences. *Probab. Theory Related Fields*, 172(3-4):869–914, 2018.

[Bra93] R. C. Bradley. Equivalent mixing conditions for random fields. *Ann. Probab.*, 21(4):1921–1926, 1993.

[Bra07] R. C. Bradley. *Introduction to strong mixing conditions. Vol. 1.* Kendrick Press, Heber City, UT, 2007.

[BS09] B. Basrak and J. Segers. Regularly varying multivariate time series. *Stochastic Process. Appl.*, 119(4):1055–1080, 2009.

[BT16] B. Basrak and A. Tafro. A complete convergence theorem for stationary regularly varying multivariate time series. *Extremes*, 19(3):549–560, 2016.

[DH95] R. A. Davis and T. Hsing. Point process and partial sum convergence for weakly dependent random variables with infinite variance. *The Annals of Probability*, 23(2):879–917, 1995.

[dHRRdV89] L. de Haan, S. I. Resnick, H. Rootzén, and C. G. de Vries. Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.*, 32(2):213–224, 1989.

[DHS17] C. Dombry, E. Hashorva, and P. Soulier. Tail measure and tail spectral process of regularly varying time series. *arXiv preprint arXiv:1710.08358*, 2017.

[DKZ94a] A. Dembo, S. Karlin, and O. Zeitouni. Critical phenomena for sequence matching with scoring. *The Annals of Probability*, 22(4):1993–2021, 1994.

[DKZ94b] A. Dembo, S. Karlin, and O. Zeitouni. Limit distribution of maximal non-aligned two-sequence segmental score. *Ann. Probab.*, 22(4):2022–2039, 1994.

[DR85] R. Davis and S. Resnick. Limit theory for moving averages of random variables with regularly varying tail probabilities. *The Annals of Probability*, 13(1):179–195, 1985.

[Dud64] R. M. Dudley. On sequential convergence. *Trans. Amer. Math. Soc.*, 112:483–507, 1964.

[Dur10] R. Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010.

[DVJ03] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.

[DZ10] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[EMD16] F. Eyi-Minko and C. Dombry. Extremes of independent stochastic processes: a point process approach. *Extremes*, 19(2):197–218, 2016.

[Fra65] S. P. Franklin. Spaces in which sequences suffice. *Fund. Math.*, 57:107–115, 1965.

[Gri63] B. Grigelionis. The convergence of stepwise random processes to a Poisson process. *Teor. Verojatnost. i Primenen.*, 8:189–194, 1963.

[Han06] N. R. Hansen. Local alignment of Markov chains. *Ann. Appl. Probab.*, 16(3):1262–1296, 2006.

[HL06] H. Hult and F. Lindskog. Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)*, 80(94):121–140, 2006.

[Hu66] S.-T. Hu. *Introduction to general topology.* Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, 1966.

[Jan18] A. Janßen. Spectral tail processes and max-stable approximations of multivariate regularly varying time series. *Stochastic Processes and their Applications*, 2018.

[Kal83] O. Kallenberg. *Random measures.* Akademie-Verlag, Berlin, third edition, 1983.

[Kal17] O. Kallenberg. *Random measures, theory and applications*, volume 77 of *Probability Theory and Stochastic Modelling.* Springer, Cham, 2017.

[Kar91] A. Karr. *Point Processes and Their Statistical Inference*, volume 7. CRC Press, 1991.

[Kin93] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability.* The Clarendon Press, Oxford University Press, New York, 1993. Oxford Science Publications.

[Kri10] D. Krizmanić. *Functional limit theorems for weakly dependent regularly varying time series.* PhD thesis, 2010.

[LP17] G. Last and M. Penrose. *Lectures on the Poisson Process.* Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.

[LRR14] F. Lindskog, S. I. Resnick, and J. Roy. Regularly varying measures on metric spaces: hidden regular variation and hidden jumps. *Probab. Surv.*, 11:270–314, 2014.

[MGW02] D. Metzler, S. Grossmann, and A. Wakolbinger. A Poisson model for gapped local alignments. *Statist. Probab. Lett.*, 60(1):91–100, 2002.

[Mor77] T. Mori. Limit distributions of two-dimensional point processes generated by strong-mixing sequences. *Yokohama Math. J.*, 25(2):155–168, 1977.

[MPU06] F. Merlevède, M. Peligrad, and S. Utev. Recent advances in invariance principles for stationary sequences. *Probab. Surv.*, 3:1–36, 2006.

[MS10] T. Meinguet and J. Segers. Regularly varying time series in Banach spaces. arXiv:1001.3262, 2010.

[MW14] T. Mikosch and O. Wintenberger. The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probability Theory and Related Fields*, 159(1-2):157–196, 2014.

[MW16] T. Mikosch and O. Wintenberger. A large deviations approach to limit theory for heavy-tailed time series. *Probab. Theory Related Fields*, 166(1-2):233–269, 2016.

[Nak88] B. S. Nakhapetyan. An approach to proving limit theorems for dependent random variables. *Theory of Probability & Its Applications*, 32(3):535–539, 1988.

[PS18] H. Planinić and P. Soulier. The tail process revisited. *Extremes*, 21(4):551–579, 2018.

[Res87] S. I. Resnick. *Extreme values, regular variation and point processes.* Applied Probability, Vol. 4,. New York, Springer-Verlag, 1987.

[Res07] S. I. Resnick. *Heavy-Tail Phenomena.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

[Ros85] M. Rosenblatt. *Stationary sequences and random fields.* Birkhäuser Boston, Inc., Boston, MA, 1985.

[RS98] S. Resnick and C. Stărică. Tail index estimation for dependent data. *Ann. Appl. Probab.*, 8(4):1156–1183, 1998.

[Sat99] K.-i. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 1999. Translated from the 1990 Japanese original, Revised by the author.

[Sch05] D. Schuhmacher. Distance estimates for dependent superpositions of point processes. *Stochastic Process. Appl.*, 115(11):1819–1837, 2005.

[Seg07] J. Segers. Multivariate regular variation of heavy-tailed markov chains. *arXiv preprint math/0701411*, 2007.

[SO12] G. Samorodnitsky and T. Owada. Tail measures of stochastic processes or random fields with regularly varying tails. preprint, 2012.

[SZM17] J. Segers, Y. Zhao, and T. Meinguet. Polar decomposition of regularly varying time series in star-shaped metric spaces. *Extremes*, 20(3):539–566, 2017.

[Tan89] H. Tanaka. Time reversal of random walks in one-dimension. *Tokyo J. Math.*, 12(1):159–174, 1989.

[TK10a] M. Tyran-Kamińska. Convergence to Lévy stable processes under some weak dependence conditions. *Stochastic Process. Appl.*, 120(9):1629–1650, 2010.

[TK10b] M. Tyran-Kamińska. Functional limit theorems for linear processes in the domain of attraction of stable laws. *Statist. Probab. Lett.*, 80(11-12):975–981, 2010.

[Whi02] W. Whitt. *Stochastic-process limits*. Springer-Verlag, New York, 2002.

[Wil04] S. Willard. *General topology*. Dover Publications, Inc., Mineola, NY, 2004. Reprint of the 1970 original [Addison-Wesley, Reading, MA; MR0264581].

[WS18] L. Wu and G. Samorodnitsky. Regularly varying random fields. *arXiv preprint arXiv:1809.04477*, 2018.

Summary

For a stationary, regularly varying and weakly dependent $\mathbb{R}$–valued time series $(X_i)_{i\in\mathbb{Z}}$, the language of point processes offers a nice probabilistic way to deduce the distribution of various functionals of the sample $X_1, \ldots, X_n$ as $n \to \infty$. This includes functionals whose behavior for large n essentially depends only on extreme X_i's. The main idea is to establish a so–called complete convergence result, that is convergence of point processes $\sum_{i=1}^{n} \delta_{(i/n, X_i/a_n)}$ for a suitable sequence (a_n) and then apply continuous mapping arguments. In this context, the limiting point process is a Poisson cluster process which can be fully described by the so–called tail process of (X_i).

However, in this type of point process convergence the information on the temporal order of extreme X_i's belonging to the same cluster is lost in the limit due to scaling of time. As one of the main contributions of the thesis, we present a new type of complete convergence result which preserves this kind of information. It applies to stationary regularly varying time series satisfying standard extremal dependence assumptions. Our approach is based on dividing the sample $X_1, \ldots, X_n$ into smaller blocks and then considering these blocks, instead of only individual X_i's, as points of a point process on a certain infinite–dimensional Polish space. Along the way, we revisit the notion of vague convergence of measures relying on an abstract theory of bounded sets and discuss general Poisson approximation theory for point processes on Polish spaces.

The order preserving convergence allows us to prove new limit results for record times and partial sums of the underlying time series. In particular, we show that rescaled record times, under an additional assumption, converge in distribution to a certain scale invariant compound Poisson process. This extends the well known result in the i.i.d. case. Furthermore, when the index of regular variation is in the interval $(0, 2)$, we obtain a new functional limit theorem for partial sums which applies to a variety of time series for which standard convergence in the space of càdlàg functions cannot hold. The main novelty is that the convergence is placed in the larger space of so–called decorated càdlàg functions equipped with an extension of Skorohod's M_2 topology. Corollaries of this result are discussed.

Finally, we use the language of stationary regularly varying random fields to revisit the well known problem of local alignment of sequences. For that purpose, we extend the notion of the tail process and the corresponding point process convergence theory to

$\mathbb{R}$–valued random fields indexed over d–dimensional integer lattice. In the course of this extension we introduce the concept of anchoring which further clarifies the link between the tail process and the asymptotic distribution of a cluster of extremes.

Keywords: Vague convergence; Point process; Poisson approximation; Regular variation; Time series; Stationarity; Tail process; Complete convergence result; Functional limit theorem; Record times; Random fields; Local sequence alignment.

Sažetak

Za stacionaran vremenski niz $(X_i)_{i\in\mathbb{Z}}$ s vrijednostima u $\mathbb{R}$ kaže se da je regularno varirajući ako su mu sve konačno–dimenzionalne distribucije višedimenzionalno regularno varirajuće. To svojstvo ekvivalentno je postojanju takozvanog repnog procesa pomoću kojeg se na intuitivan način mogu opisati ekstremalna svojstva tog vremenskog niza. Za takav niz, uz pretpostavku slabe zavisnosti, teorija točkovnih procesa nudi lijep vjerojatnosni pristup za određivanje distribucije raznih funkcionala uzorka $X_1, \ldots, X_n$ kada $n \to \infty$. Tu spadaju funkcionali čije ponašanje za velike n u principu ovisi samo o ekstremnim vrijednostima niza (X_i). Glavna ideja je prvo pokazati konvergenciju točkovnih procesa $\sum_{i=1}^{n} \delta_{(i/n, X_i/a_n)}$ za prikladan niz (a_n) te na nju primijeniti takozvane tehnike neprekidnih preslikavanja. Budući da iz te iste točkovne konvergencije možemo odrediti ponašanje puno različitih funkcionala, takvu konvergenciju često zovemo potpuna točkovna konvergencija. Nadalje, u ovom kontekstu granični točkovni proces je Poissonov proces s klasterima koji je potpuno određen repnim procesom niza (X_i).

Ipak, u ovom obliku potpune točkovne konvergencije, zbog skaliranja vremena, u limesu se gubi informacija o vremenskom poretku ekstremnih observacija koje su dio istog klastera. Jedan od glavnih doprinosa ove teze je predstavljanje novog oblika potpune točkovne konvergencije koji čuva ovaj poredak. On vrijedi za stacionarne regularno varirajuće vremenske nizove uz standardne pretpostavke o ekstremalnoj zavisnosti. Naš pristup baziran je na dijeljenju uzorka $X_1, \ldots, X_n$ na manje blokove i tretiranjem tih blokova, umjesto pojedinačnih observacija, kao točaka točkovnog procesa na određenom beskonačno–dimenzionalnom poljskom prostoru. Usput, dajemo alternativan pristup takozvanoj *vague* konvergenciji mjera koristeći apstraktnu teoriju ograničenih skupova i diskutiramo generalne uvjete pod kojima određeni točkovni procesi na poljskim prostorima konvergiraju po distribuciji prema Poissonovom procesu.

Potpuna točkovna konvergencija koja čuva poredak omogućava nam da pokažemo nove granične rezultate za vremena rekorda i parcijalne sume vremenskog niza u pozadini. Točnije, pokazujemo da reskalirana vremena rekorda, uz dodatnu pretpostavku na repni proces niza, konvergiraju po distribuciji prema određenom složenom Poissonovom procesu koji je invarijantan na skaliranje. Ovo poopćuje poznati rezultat u slučaju niza nezavisnih i jednako distribuiranih slučajnih varijabli. Nadalje, kada je indeks regularne varijacije u intervalu $(0, 2)$, pokazujemo novi funkcionalni granični teorem za parcijalne sume

koji je primjenjiv na velik broj vremenskih nizova za koje standardna konvergencija u prostoru càdlàg funkcija ne može vrijediti. Glavna novina je da se konvergencija gleda u većem prostoru takozvanih dekoriranih càdlàg funkcija uz topologiju koja je ekstenzija Skorohodove M_2 topologije. Također, diskutiramo i korolare ovog rezultata.

Na kraju, koristimo jezik stacionarnih regularno varirajućih slučajnih polja kako bi dali novi uvid u klasični problem lokalnog poravnanja dvaju nizova znakova. U tu svrhu, proširujemo pojam repnog procesa i teoriju potpune točkovne konvergencije na slučajna polja sa skupom indeksa $\mathbb{Z}^d$. U sklopu te ekstenzije predstavljamo ideju usidrenja koja dodatno razjašnjuje vezu između repnog procesa i granične distribucije klastera ekstremnih observacija.

Ključne riječi: Vague konvergencija; Točkovni proces; Poissonova aproksimacija; Regularna varijacija; Vremenski niz; Stacionarnost; Repni proces; Potpuna točkovna konvergencija; Funkcionalni granični teorem; Vremena rekorda; Slučajna polje; Lokalno poravnanje nizova.

Curriculum vitae

Hrvoje Planinić was born on 12th February 1990 in Split, Croatia. He received his primary and secondary education in Zagreb. In 2014, he graduated at the Department of Mathematics, Faculty of Science, University of Zagreb, with master thesis *Martingale techniques and mixing times of Markov chains* under supervision of prof. Zoran Vondraček. Right after, he enrolled in the doctoral program in mathematics at the same department under supervision of prof. Bojan Basrak and became a member of the *Probability seminar*.

In the beginning of 2015, he was employed as a research and teaching assistant at the Department of Mathematics, Faculty of Science, University of Zagreb, within the project *Stochastic methods in analytical and applied problems* (PI: Zoran Vondraček) financed by the Croatian Science Foundation (3526, 2014-2018). He currently participates in the work of bilateral Hungarian–Croatian project *Statistical inference for branching processes with immigration* (Croatian PI: Bojan Basrak, Hungarian PI: Gyula Pap).

He was a teaching assistant for several courses including *Markov chains* and *Probability and statistics*.

He attended a number of conferences and workshops at which he gave 3 talks and 1 poster presentation. He had 5 short research visits during which he gave 1 seminar talk. He co–authored 4 papers, 2 of which are published:

1. B. Basrak, H. Planinić, and P. Soulier. An invariance principle for sums and record times of regularly varying stationary sequences. *Probab. Theory Related Fields*, 172(3-4):869–914, 2018.

2. H. Planinić and P. Soulier. The tail process revisited. *Extremes*, 21(4):551–579, 2018.