arXiv:2011.10153v1 [math.DS] 19 Nov 2020

ENRICHED FUNCTIONAL LIMIT THEOREMS FOR DYNAMICAL SYSTEMS

ANA CRISTINA MOREIRA FREITAS, JORGE MILHAZES FREITAS, AND MIKE TODD

Abstract. We prove functional limit theorems for stochastic processes which have clusters of large values which, when summed and suitably normalised in time and space, get collapsed in a jump of the limiting process observed at the same time point. In order to keep track of the clustering information, which gets lost in the usual Skorohod topologies in the space of càdl'ag functions, we introduce a new space specially designed for that purpose which generalises the already more general spaces introduced by Whitt. Our main applications are to hyperbolic and non-uniformly expanding dynamical systems with heavy tailed observable functions maximised at dynamically linked maximal sets (such as periodic points). We also study limits of extremal processes and record times point processes for observables not necessarily heavy tailed. The applications studied include hyperbolic systems such as Anosov diffeomorphisms, but also non-uniformly expanding maps such as maps with critical points of Benedicks-Carleson type or indifferent fixed points such as Manneville-Pomeau maps. The main tool is a limit theorem for point processes with decorations derived from a bi-infinite sequence which we call the piling process, designed to keep the clustering information. This piling process is based on the tail process defined in the regularly varying context rather than the general stationary stochastic processes we consider here.

Contents

1. Introduction 2
2. Enriched functional limit theorems for non-uniformly hyperbolic dynamics 5
2.1. Dynamically defined stochastic processes 5
2.2. Applications to specific systems 6
2.3. The functional spaces 8
2.4. Rare events and point processes 12
2.5. Functional limit theorems for heavy tailed dynamical sums 13
2.6. Enriched extremal process dynamics in the presence of clustering 15
2.7. Record point processes 18

Date: November 23, 2020.
2010 *Mathematics Subject Classification.* 37A50, 37B20, 60G70, 60G55, 37A25.
Key words and phrases. Complete convergence, point processes, clustering, hitting times.
ACMF, JMF and MT were partially supported by FCT project PTDC/MAT-PUR/28177/2017, with national funds, and by CMUP, which is financed by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project with reference UIDB/00144/2020. We would like to thank Ian Melbourne for insightful discussions on Skorohod spaces and for various suggestions for improvement of this text.

3. General complete convergence of multi-dimensional cluster point processes 19
3.1. Identifying clusters 19
3.2. Dependence structure 20
3.3. Bookkeeping of clusters 22
3.4. Complete convergence of point processes 32
3.5. Application to jointly regularly varying sequences 34
4. Proofs of the dynamical enriched functional limit theorems 35
4.1. Proof of Theorem 2.5 37
4.2. Proof of Theorem 2.9 39
4.3. Proof of Theorem 2.11 42
Appendix A. Completeness and separability of the space F' 43
Appendix B. Weak and weak$^{\#}$ convergence 44
Appendix C. Convergence of point processes on non locally compact spaces 47
Appendix D. Asymptotic independence in disjoint time intervals 51
Appendix E. Remaining proofs for the dynamical examples 53
References 56

1. Introduction

The Donsker functional limit theorem gives an invariance principle for the sum of independent and identically distributed (i.i.d.) random variables with finite second moments. The functional limit is a Brownian motion which lives in the space C of continuous $\mathbb{R}^d$-valued functions defined on a subinterval I of the real line, usually taken as either $[0,1]$ or $[0,+\infty)$, equipped with the uniform norm. Invariance principles such as this have been proved for a large class of dynamical systems with good mixing properties ([HK82, DP84, MT02, FMT03, MN05, HM07, BM08, MN09, Gou10]). Note that weak convergence to a Brownian motion implies the existence of a Central Limit Theorem (CLT) for the ergodic sums of such systems.

For systems with weaker mixing properties such as those modelled by Young towers ([You98, You99]), with return time functions that are not square integrable, which is the case of certain Pomeau-Manneville ([PM80]) or Liverani-Saussol-Vaienti ([LSV99]) maps, where the existence of an indifferent fixed point is a source of non-uniform hyperbolicity, the CLT fails and the ergodic sums fall into the domain of attraction of α-stable laws ([Gou04]). In [Gou08], Gouëzel shows that the same regime applies to ergodic sums of unbounded heavy tailed observables for the doubling map. Later, in [TK10b], Tyran-Kamińska showed a functional limit theorem for heavy tailed ergodic sums of essentially uniformly expanding maps. The functional limit obtained was an α-stable Lévy process, which lives in the space D of càdlàg functions, *i.e.*, right continuous $\mathbb{R}$-valued functions with left limits defined on I, equipped with Skorohod's J_1 topology. In this scenario, the heavy tails allow for the occurrence of extremal observations

with a considerable impact in the Birkhoff sum, which are ultimately responsible for the appearance of the jumps in the limit Lévy process.

In [TK10b], Tyran-Kamińska also shows that in the case of certain LSV maps although the ergodic sums, when properly normalised, converge to an α-stable law (as proved in [Gou04]), there is no functional version of this result in the J_1 topology. The problem lies in the fact that there is a stacking or piling of observations due to the compression resulting from the time normalisation, which causes an accumulation of numerous small jumps that aggregate into a big jump in the limit and this is forbidden by the J_1 topology. This issue was solved in [MZ15] where through inducing arguments it was shown that a functional version encompassing convergence to an α-stable Lévy process does apply if one considers a weaker topology, namely, the Skorohod M_1 topology, which does not preclude the accumulation of many small jumps into a big jump. Then more recently in [MV20, JPZ20, JMP+19], the authors considered billiards with cusps. In [JPZ20], the authors prove convergence to a Lévy process in the M_1 topology for observables that do not change sign near the cusp. In order to allow more freedom in the piling of observations, namely, to allow sign changes of the observable near the cusp, in [MV20], an even weaker topology was considered, Skorohod's M_2 topology, so that convergence to a Lévy process would not be precluded. However, though the choice of weaker topologies does allow for functional limits, the information collapsed in a jump of the limiting Lévy motion is completely lost.

The main purpose of this paper is to develop tools in order to obtain enriched functional limit theorems that incorporate the collapsed information without any loss. Above there are two contexts where Lévy processes arise: 'good' dynamical systems and 'bad' (heavy) observables; and 'bad' dynamical systems (eg with cusps) and 'good' observables. The theory developed here principally applies to the former context (similar to [Gou08, TK10b]), though it does have some applications in the latter context, and introduces a framework for further advances.

We illustrate the power of these new tools in generalised versions of the settings in [Gou08, TK10b]: in cases of vector valued heavy tailed observables for systems modelled by Young towers, and in the presence of general clusters of extremal observations which get collapsed in the same jump of the limiting Lévy process. The limits that we will consider have discontinuous sample paths and therefore live in the space D of càdlàg functions, *i.e.*, right continuous $\mathbb{R}^d$-valued functions with left limits defined on I. However, this space is not sufficiently rich to keep a record of the fluctuations occurring during the clusters of high observations. The height of the limiting jump accounts for the aggregate effect of all the cluster observations. However, the oscillations observed during the cluster may exceed the height of the jump, for example. In order to solve this loss of information, in [Whi02], Whitt proposed a new space that he called E, which decorates each discontinuity of the limit process in D with an excursion corresponding to a connected set describing the maximum and minimum fluctuations observed. In fact, when $d = 1$, the excursion at the discontinuity time t^* of the limiting process $V(t)$ is decorated with an interval bounded by the smallest and largest values achieved by the process during the collapsed cluster, which must contain $V(t^{*-})$ and $V(t^*)$, where $V(t^{*-})$ denotes the lefthand limit of the càdlàg function V at t^*.

Nevertheless, most of the information during the cluster is lost and only the maximum oscillations are recorded in E, while the intermediate fluctuations are completely disregarded. In [Whi02], Whitt also proposed the space F, which keeps track in particular of the ordering in which points are visited within an excursion in E. However, the space F still disregards

information because while it keeps track of all the changes of direction during the excursions, it does not keep record of the intermediate jumps observed in the same direction. One of our main goals is to consider a space where no information collapsed into a jump is lost. For that purpose we introduce a new space that we will call F'. We will endow it with a metric and discuss some of its properties. Then we will use this space to study sums of heavy tailed observables, general extremal processes and records, for which we will obtain enriched functional limit theorems in the setting of non-uniformly hyperbolic systems modelled by Young towers, all of which will be carried out with very minimal loss of information.

Point processes have been used successfully to prove functional limit theorems in D ([Res87]). Here, since we need to keep control of the information during the excursions, we will use a new type of point processes introduced in [BPS18], where the authors considered point processes in non-locally compact spaces, designed to maintain the ordering of the cluster observations collapsed into the same instant of time, and proved their convergence, under some conditions, for stationary jointly regularly varying sequences. In [BPS18], the authors then applied this convergence of point processes to obtain convergence of sums of jointly regularly varying sequences of random variables in the space E. In order, to prove convergence in F', we will generalise their results to more general stationary processes and under weaker conditions, in particular, under a much weaker mixing condition, which is essential to apply to processes arising from dynamical systems. Yet, this generalisation may have an interest on its own, in the more classical probabilistic setting.

A sketch of a reduced version of one of our main theorems (Theorem 2.9) is as follows.

Theorem. *Let (X, T, μ) be a dynamical system as in Section 2.2 and $\phi : X \to [-\infty, \infty]$ a sufficiently regular observable observable with an α-heavy tail for $0 < \alpha < 1$, such that the clustering behaviour of $\{\phi > s\}$ converges as $s \to \infty$. Then there exists a scaling sequence $(a_n)_n$ and $V(t) \in F'$ such that the variable*

$$S_n(t) = \sum_{i=1}^{\lfloor nt \rfloor} \phi \circ T^i$$

converges in distribution to $V(t)$ in F', with excursions determined by the clustering behaviour.

There is also a related result for point processes. Both of these results extend beyond the dynamical and one-dimensional observable settings, and also to $\alpha \in [1, 2)$.

The paper is organised as follows. In Section 2 we outline the theory from the point of view of dynamical systems, though the theory extends beyond that: we give the relevant observables, focussing on the heavy tailed cases, and briefly describe some of the basic dynamical systems models to which the theory applies. We also define out functional spaces, culminating in the new space F', and are then able to state our fist limit theorem on convergence to Lévy processes. We then put the theory into the context of Extreme Value Theory and show convergence to the relevant extremal process. This naturally leads to a theorem on the convergence of record point processes.

In Section 3 we focus on the convergence of point processes for general stochastic processes and rather general observables (not only heavy tailed). Conditions $Д_{q_n}$ and $Д'_{q_n}$ are given and the appropriate sequence spaces to record our exceedances are defined. The tail process is then defined which then leads to the definition of the piling process. With all of these conditions

met and the piling process existing, we then prove complete convergence of the point process to a Poisson point process. We also show the connection to our setting of jointly regularly varying sequences.

In Section 4 we show how the results from the previous section can be applied in the dynamical context and then prove the functional limit theorems. In the appendices we cover some of the required background for the results here: Appendix A contains completely new theory for our space F'; Appendices B and C give some adaptations of classical theory for sequence spaces and point processes; Appendix D gives some adaptations of relatively standard theory in Extreme Value Theory to the context here; and Appendix E contains the remaining arguments to show that the theory in this paper applies to the dynamical systems models claimed.

2. Enriched functional limit theorems for non-uniformly hyperbolic dynamics

In this section we start by introducing the setting and in particular the new space F' and its properties. Then we state the main results regarding the enriched functional limits for sums of vector valued heavy tailed observables, extremal processes and record point processes. We emphasise that the results in Section 3 hold well beyond the dynamical setting.

2.1. **Dynamically defined stochastic processes.** Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, \mu, T)$ be a discrete time dynamical system, where $\mathcal{X}$ is a compact manifold equipped with a norm $\|\cdot\|$, $\mathcal{B}_{\mathcal{X}}$ is its Borel σ-algebra, $T : \mathcal{X} \to \mathcal{X}$ is a measurable map and μ is a T-invariant probability measure, *i.e.*, $\mu(T^{-1}(B)) = \mu(B)$ for all $B \in \mathcal{B}_{\mathcal{X}}$. Let $\Psi : \mathcal{X} \to \mathbb{R}^d$ be an observable (measurable) function and define the stochastic process $\mathbf{X}_0, \mathbf{X}_1, \dots$ given by

$$\mathbf{X}_n = \Psi \circ T^n, \qquad \text{for every } n \in \mathbb{N}_0. \tag{2.1}$$

High values of $\|\Psi(\cdot)\|$ will correspond to entrances in a neighbourhood of a zero measure maximal set $\mathcal{M}$, which we express in the following way. Let $\mathcal{M} \subset \mathcal{X}$ be such that $\mu(\mathcal{M}) = 0$ and let $g : [0, \to \infty) \to \mathbb{R} \cup \{+\infty\}$ be such that 0 is a global maximum, where we allow $g(0) = +\infty$, and g is a strictly decreasing bijection in a neighbourhood of 0. We assume that, on a neighbourhood of $\mathcal{M}$,

$$\|\Psi(x)\| = g(\operatorname{dist}(x, \mathcal{M})), \quad \text{where } \operatorname{dist}(x, \mathcal{M}) = \inf\{\operatorname{dist}(x, \zeta) \colon \ \zeta \in \mathcal{M}\}. \tag{2.2}$$

where g has one of the three types of behaviour:

Type g_1: there exists some strictly positive function[1] $h : W \to \mathbb{R}$ such that for all $y \in \mathbb{R}$

$$\lim_{s \to g_1(0)} \frac{g_1^{-1}(s + yh(s))}{g_1^{-1}(s)} = \mathrm{e}^{-y}; \tag{2.3}$$

Type g_2: $g_2(0) = +\infty$ and there exists $\alpha > 0$ such that for all $y > 0$

$$\lim_{s \to +\infty} \frac{g_2^{-1}(sy)}{g_2^{-1}(s)} = y^{-\alpha}; \tag{2.4}$$

[1] A possible choice for h is given in [LFF+16, Chapter 4.2.1].

Type g_3: $g_3(0) = D < +\infty$ and there exists $\gamma > 0$ such that for all $y > 0$

$$\lim_{s\to 0} \frac{g_3^{-1}(D - sy)}{g_3^{-1}(D - s)} = y^{\gamma}. \tag{2.5}$$

Most of the results regarding hitting times and extreme values for dynamical systems were obtained when $\mathcal{M}$ is reduced to a single point ζ. Recent results have considered $\mathcal{M}$ to be a countable set ([AFFR17]), submanifolds ([FGGV18, CHN19]) and fractal sets ([MP16, FFRS20]). Our results can be applied to general maximal sets $\mathcal{M}$ and general Ψ, under the assumption that the piling process is well defined, which will be verified and illustrated for the more common case where $\mathcal{M} = \{\zeta\}$, for some hyperbolic point $\zeta \in \mathcal{X}$, and where Ψ, on neighbourhood of ζ, can be written as

$$\Psi(x) = g(\mathrm{dist}(x, \zeta)) \frac{\Phi_\zeta(x)}{\|\Phi_\zeta(x)\|} \mathbb{1}_W(x), \tag{2.6}$$

where $\Phi_\zeta : V \to W$ denotes a diffeomorphism, defined on an open ball, V, around ζ in $T_\zeta \mathbf{X}$, the tangent space at ζ, onto a neighbourhood, W, of ζ in $\mathbf{X}$, such that $\Phi_\zeta(E^{s,u} \cap V) = W^{s,u}(\zeta) \cap W$.

Remark 2.1. If g is to type g_2 (for example, $g(x) = x^{-1/\alpha}$), the measure is sufficiently regular and the geometry of the maximal set $\mathcal{M}$ is simple (for example, $\mathcal{M} = \{\zeta\}$, for some $\zeta \in \mathcal{X}$), then the stochastic process $\mathbf{X}_0, \mathbf{X}_1, \ldots$ is jointly regularly varying.

2.2. **Applications to specific systems.** The theory developed in this paper applies to general stochastic processes, but the applications we focus on are dynamical systems. In this subsection we give some preliminary examples of such applications ranging from systems with good exponential mixing properties to some with poor mixing behaviour, leaving further applications to future works. Since some aspects of the proofs of the facts used here require the establishment of some new tools, we postpone these to Section E (see also the discussion following Theorem 4.1).

2.2.1. *Non-uniformly expanding systems.*
Uniformly expanding systems. We first list some well-behaved dynamical systems which are essentially uniformly expanding, though they need not have Markov, or compactness, properties. More details about the required properties are at the beginning of Section 4.

- Uniformly expanding continuous maps of the interval/circle;
- Markov maps;
- Piecewise expanding maps of the interval with countably many branches such as Rychlik maps;
- Saussol's class of higher dimensional expanding maps

Here we are always assuming that μ is an absolutely continuous (with respect to Lebesgue) invariant probability measure or *acip* since these are a very natural class, though the theory extends beyond these.

Non-trivial examples of observables on these systems to which we can apply the theory include those maximised at a repelling periodic point, first studied in [FFT12], but there is huge scope to study further clustering behaviour such as that shown in [AFFR16, AFFR17, FFRS20] for example. We require that the density $\frac{d\mu}{\mathrm{Leb}}$, where Leb is Lebesgue, is bounded at the periodic point so that the conformal properties of T there are reflected in the measure μ as well as

Leb, but this is automatic in the Rychlik case. Note that in the periodic case q_n can be taken as q, the period of the point.

In the heavy-tail applications we will restrict ourselves to $0 < \alpha < 1$ for simplicity, but observe that techniques to prove (2.21), required in the $1 \leq \alpha < 2$ case, are provided in [TK10b]: these immediately apply to some of the simpler cases above. Conditions (2.22) and (2.23) are easy computations in the periodic case.

Finally for this introductory discussion on dynamical applications, we note that there is a condition on $(Q_j)_j$ in Theorem 2.11 which is easily satisfied in all the dynamical examples, see Appendix E.

Remark 2.2. Usually in the context of extreme value theory inducing methods allow us to handle a wider range of systems, those which induce to systems mentioned above. However, the tools we develop here are more fine-tuned than previous ones and are not compatible with inducing. One way to see this is to note that inducing introduces variable time scales, speeding up different parts of the space at different rates. While the speeding up is asymptotically uniform, the short-term non-uniformity does not work well with the way our point processes record exceedances in blocks since the blocks have a fixed time scale.

Nevertheless there are direct methods of proof for some non-uniformly expanding systems, as below.

Benedicks-Carleson quadratic maps. Here we provide a class of maps which are far from uniformly expanding, indeed there are critical points. As in Remark 2.2 we cannot apply the usual inducing techniques.

Here we set $I = [-1, 1]$ and for $a \in (0, 2]$ define $f_a : I \to I$ by $f_a(x) = 1 - ax^2$. This map satisfies the *Benedicks-Carleson* conditions (here a is close to 2) if there exists $c, \gamma > 0$ (c should be close to $\log 2$ and γ is small) such that

$$|Df_a^n(f(0))| \geq e^{cn} \text{ for } n \in \mathbb{N}_0 \text{ and } |f_a^n(0)| \geq e^{-\gamma\sqrt{n}} \text{ for } n \in \mathbb{N}.$$

It is known that there is a positive Lebesgue measure set BC of a such that f_a satisfies these conditions.

When for definiteness we consider that our observable is maximised at a periodic point of f_a, we note that the fact that the density at our periodic point is bounded is more delicate than above, but the existence of suitable periodic points for (a positive Lebesgue measure set of) maps in BC is shown in [FFT13, Section 6].

Manneville-Pomeau maps. Our previous examples all have exponential decay of correlations, but our final interval map example shows that this is not necessary. This is what is often referred to as the Liverani-Saussol-Vaienti version of the Manneville-Pomeau map: for $\gamma \in (0, 1)$, define $T_\gamma : [0, 1] \to [0, 1]$ by

$$T_\gamma(x) := \begin{cases} x(1 + 2^\gamma x^\gamma) & \text{if } x \in [0, 1/2), \\ 2x - 1 & \text{if } x \in [1/2, 1]. \end{cases}$$

This has an acip $\mu = \mu_\gamma$ and this system does not satisfy decay of correlations against $L^1(\mu)$, indeed it has (sharp) polynomial decay of correlations for Hölder observables. To ensure the

required conditions hold here we will assume $\gamma \in (0, 0.289)$, see the adaptation of [FFTV16, Section 4.2.2] in Section E, though we would expect them to hold for all $\gamma \in (0, 1/2)$ at least.

2.2.2. *Invertible hyperbolic systems.* We also consider hyperbolic invertible systems consisting of Anosov linear diffeomorphism, $T : \mathbb{T}^2 \to \mathbb{T}^2$, defined on the flat torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. We can associate T to a 2×2 matrix, L, with integer entries, determinant 1, and without eigenvalues of absolute value 1. As the determinant of L is equal to 1, the Riemannian structure induces a Lebesgue measure on $\mathbb{T}^2$ which is invariant by T. These systems are Bernoulli and have exponential decay of correlations with respect to Hölder observables.

2.3. **The functional spaces.** Let $D = D([0, 1], \mathbb{R}^d)$ be the space of càdlàg functions defined on $[0, 1]$. One can define several metrics in D. The most usual is the so-called J_1 Skorohod's metric, which generalises the uniform metric by allowing a small deformation of the time scale. In this metric a jump in the limit function must be matched by a similar one in the approximating functions. In order to establish limits with unmatched jumps, Skorohod introduced the M_1 and M_2 topologies which use completed graphs of the functions. We will use a metric motivated by M_2, which considers the Hausdorff distance between compact sets. We refer to [Whi02] for precise definitions and properties.

In order to keep some of the information collapsed in limit jumps, and to broaden the class of convergent sequences, Whitt introduced the space $E = E([0, 1], \mathbb{R}^d)$ (see [Whi02, Sections 15.4 and 15.5]), as the space of excursion triples

$$(x, S^x, \{I(s)\}_{s \in S^x}),$$

where $x \in D$, S^x is a countable set containing the discontinuities of x, denoted by $disc(x)$, *i.e.*, $disc(x) \subset S^x$, and, for each $s \in S^x$, $I(s)$ is a compact connected subset of $\mathbb{R}^d$ containing at least $x(s^-)$ and $x(s)$. We may identify each element of E with the set valued function

$$\hat{x}(t) = \begin{cases} I(t) & \text{if } t \in S^x \\ \{x(t)\} & \text{otherwise} \end{cases}, \tag{2.7}$$

and its graph $\Gamma_{\hat{x}} = \{(t, z) \in [0, 1] \times \mathbb{R}^d \colon\ z \in \hat{x}(t)\}$. Letting $p_\ell : \mathbb{R}^d \to \mathbb{R}$ denote the projection onto the ℓ-th coordinate for $\ell = 1, \dots, d$, we define $\hat{x}^\ell(t) = p_\ell(\hat{x}(t))$ and $\Gamma^\ell_{\hat{x}} = \{(t, z) \in [0, 1] \times \mathbb{R} \colon\ z \in \hat{x}^\ell(t)\}$.

We embed D into E, in the following way. For $a, b \in \mathbb{R}^d$, we define the product segment

$$[[a, b]] := [a^1, b^1] \times \cdots \times [a^d, b^d],$$

where the one dimensional segment $[a^\ell, b^\ell]$ coincides with the interval[2] $[a^\ell \wedge b^\ell, a^\ell \vee b^\ell]$. If we have $a^\ell = b^\ell$, then $[a^\ell, b^\ell] = \{a^\ell\} = \{b^\ell\}$. We identify $x \in D$ with the element of E

$$\big(x, disc(x), \{I(s)\}_{s \in disc(x)}\big), \quad \text{where } I(s) = [[x(s^-), x(s)]].$$

We use the Hausdorff metric to define a metric on E. Namely, recall that for compact sets $A, B \subset \mathbb{R}^d$, the Hausdorff distance between A and B is given as

$$m(A, B) = \max \left\{ \sup_{x \in A} \left\{ \inf_{y \in B} \|x - y\| \right\}, \sup_{y \in B} \left\{ \inf_{x \in A} \|x - y\| \right\} \right\}.$$

[2]Recall the notation, used throughout this paper, $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$

For $A \subset \mathbb{R}^d$, let $d(A) = \sup_{x,y\in A}\{\|x-y\|\}$ be the diameter of A. In order to use the Hausdorff metric, we assume that the elements of E satisfy the condition:

$$\text{for all } \epsilon > 0, \text{ there exist finitely many } s \in S^x \text{ such that } d(I(s)) > \epsilon. \tag{2.8}$$

This guarantees that for each element of E, the associated graph $\Gamma_{\hat{x}}$ is a compact set. This way, we endow E with the Hausdorff metric simply by establishing that

$$m_E(\hat{x}, \hat{y}) = \max_{\ell=1,\dots,d} m(\Gamma^\ell_{\hat{x}}, \Gamma^\ell_{\hat{y}}). \tag{2.9}$$

Endowed with this metric E is separable but not complete. Alternatively, we can define the stronger metric, called uniform metric given by

$$m^*_E(\hat{x}, \hat{y}) = \max_{\ell=1,\dots,d} \sup_{t\in[0,1]} m(\hat{x}^\ell(t), \hat{y}^\ell(t)). \tag{2.10}$$

When endowed with the metric m^*_E, the space E is complete but not separable. We refer to [Whi02] for further details about E and its properties.

The space E only records the maximal oscillations when information collapses into a jump in the limit. In order to keep a closer track of the fluctuations during the excursions, Whitt introduced the space F, in [Whi02, Section 15.6], which corresponds to the set of equivalence classes of the set of all the parametric representations of the graphs $\Gamma_{\hat{x}}$ of elements $\hat{x}$ of E, by setting that two parametrisations (continuous functions from $[0,1]$ into $\Gamma_{\hat{x}}$) (r_1, u_1) and (r_2, u_2) are equivalent if there exist continuous nondecreasing onto functions $\lambda_1, \lambda_2 : [0,1] \to [0,1]$ such that $(r_1, u_1) \circ \lambda_1 = (r_2, u_2) \circ \lambda_2$. This means that, in particular, the two functions u_1, u_2 in the two equivalent parametric representations of $\Gamma_{\hat{x}}$ visit all the points of $\Gamma_{\hat{x}}$ in each $I(s)$ the same number of times and in the same order. However, note that F still misses some fluctuations such as intermediate jumps in the same direction which give rise to a big jump in the limit.

2.3.1. *The new functional space F' recording all fluctuations.* In order to keep track of all the fluctuations without missing information we introduce the space $F' = F'([0,1])$. We start by considering $\tilde{D} = \tilde{D}([0,1], \mathbb{R}^d) = D([0,1], \mathbb{R}^d)/\sim$ where $x \sim y$ if there exists a reparametrisation $\lambda : [0,1] \to [0,1]$, i.e., a continuous strictly increasing bijection such that $x \circ \lambda = y$. Denote the equivalence class of x by $[x]$. We define

$$d_{\tilde{D}}([x], [y]) = \inf_\lambda \|x \circ \lambda - y\|,$$

where λ is the set of continuous strictly increasing bijections of $[0,1]$ to itself (this could be thought of as the induced metric from the J_1 metric on $\tilde{D}$).

We will abuse notation within $\tilde{D}$ by writing x to refer to both a representative of its equivalence class $[x]$ and the equivalence class itself.

Now define

$$F' := \{\underline{x} = (x, S^x, \{e^s_x\}_{s\in S^x})\},$$

where $x \in D([0,1], \mathbb{R}^d)$, $S^x \subset [0,1]$ is an at most countable set containing $disc(x)$, the discontinuities of x and $e^s_x \in \tilde{D}([0,1], \mathbb{R}^d)$ is the excursion at $s \in S^x$, which is such that $e^s_x(0) = x(s^-)$ and $e^s_x(1) = x(s)$.

We can embed D into F', by associating to $x \in D([0,1], \mathbb{R}^d)$ the element

$$\underline{x} = \left(x, disc(x), \{e^s_x\}_{s\in disc(x)}\right), \tag{2.11}$$

where $e_x^s \in \tilde{D}([0,1],\mathbb{R}^d)$ is the equivalence class represented, for example, by

$$e_x^s(t) = x(s-) + (x(s) - x(s^-))\mathbb{1}_{[1/2,1]}(t), \quad t \in [0,1].$$

We will project F' into Whitt's space E and into $\tilde{D}$, which will give us a metric and a space with more information than Whitt's space F.

Let $\pi_E(\underline{x}) = \underline{x}^E := (x, S^x, \{I(s)\}_{s\in S^x})$ where

$$I(s) = \left[\inf_{t\in[0,1]} e_x^{s,1}(t), \sup_{t\in[0,1]} e_x^{s,1}(t)\right] \times \cdots \times \left[\inf_{t\in[0,1]} e_x^{s,d}(t), \sup_{t\in[0,1]} e_x^{s,d}(t)\right], \text{ with } e_x^{s,\ell}(t) = p_\ell(e_x^s(t)).$$

We project F' into $\tilde{D}$ as follows. Suppose that S^x is countable and write $S = \{s_i\}_{i=1}^\infty$, since the finite case follows more straightforwardly. Let $0 = a_1 < a_2 < \cdots < 1$ be such that $a_i \to 1$ as $i \to \infty$ (the choice of $(a_n)_n$ really is arbitrary). We insert the intervals $[a_i, a_{i+1}]$ at the points $s \in S^x$. This is a simple idea, but since S may be complicated, we need some notation. Define for $i \in \mathbb{N}$,

$$\bar{a}_i := \sum_{s_j \le s_i} (a_{j+1} - a_j), \quad c_i := s_i + \bar{a}_i - (a_{i+1} - a_i), \quad d_i := s_i + \bar{a}_i, \quad \bar{t} := \sup\{\bar{a}_i : s_i < t\}.$$

Thus $[c_i, d_i]$ is the interval corresponding to s_i and $\bar{t}$ is the accumulated length of the intervals inserted before t, so we can think of t being shunted to $t + \bar{t}$. Note also that our time line is now of length 2, so we will need to rescale back to a length 1 interval. We then define a representative of the equivalence class $\tilde{\pi}(\underline{x})$ by

$$\underline{x}^{\tilde{D}}(t) = \begin{cases} x(2t - \bar{t}) & \text{if } 2t \notin \cup_i [c_i, d_i], \\ e_x^{s_i}\left(\frac{2t - c_i}{d_i - c_i}\right) & \text{if } 2t \in [c_i, d_i]. \end{cases} \tag{2.12}$$

Define

$$d_{F'}(\underline{x}, \underline{y}) = d_E(\pi_E(\underline{x}), \pi_E(\underline{y})) + d_{\tilde{D}}(\tilde{\pi}(\underline{x}), \tilde{\pi}(\underline{y})), \tag{2.13}$$

where d_E denotes either m_E or m_E^* defined in (2.9) and (2.10).

Note that we could also project into Whitt's F space here (using π_E, but with $\{e_x^s\}_{s\in S^x}$ to give the order of the parametrisation in $I(s)$), and that convergence in F' implies convergence in F. In F' we keep the information of the displacement in $\mathbb{R}^d$ of all (including intermediate) jumps in the discontinuities, while F only keeps the order and information on the 'local range', *i.e.*, it only captures local extrema. To see this in a 1d example note that the excursions denoted

$$e_x(t) = \begin{cases} 0 & \text{if } 0 \le t < \frac{1}{3}, \\ \frac{1}{2} & \text{if } \frac{1}{3} \le t < \frac{2}{3}, \\ 1 & \text{if } \frac{2}{3} \le t \le 1, \end{cases} \qquad e_y(t) = \begin{cases} 0 & \text{if } 0 \le t < \frac{1}{3}, \\ 1 & \text{if } \frac{1}{3} \le t \le 1, \end{cases}$$

yield the same representations as part of F: namely the line $[0,1]$ with any parametrisation which is an orientation preserving homeomorphism. However, as components of F' they are distinct with $d_{\tilde{D}}(e_x, e_y) = \frac{1}{2}$. Indeed, if $\underline{x}$ and $\underline{y}$ differ only by having a discontinuity s having e_x and e_y as the corresponding excursion respectively, then $d_{F'}(\underline{x}, \underline{y}) = \frac{1}{2}$.

Remark 2.3. The definition of the space and topology can be generalised trivially to other compact time domains such as $[t_1, t_2]$, with $0 \le t_1 < t_2$. In order to consider a notion of convergence on non-compact domains such as $F'((0,\infty),\mathbb{R}^d)$, we say that $\underline{x}_n \to \underline{x}$, in

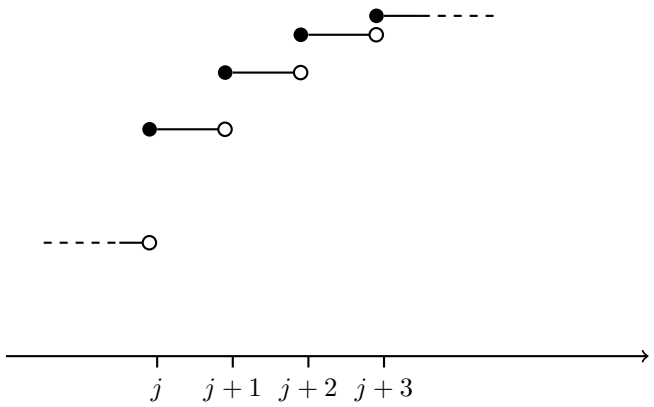

FIGURE 2.1. A piece of $S_n(t)$ (see (2.20)) in the Gouëzel example [Gou08] where there is a close approach to zero at time $1 \leq j \leq n-3$.

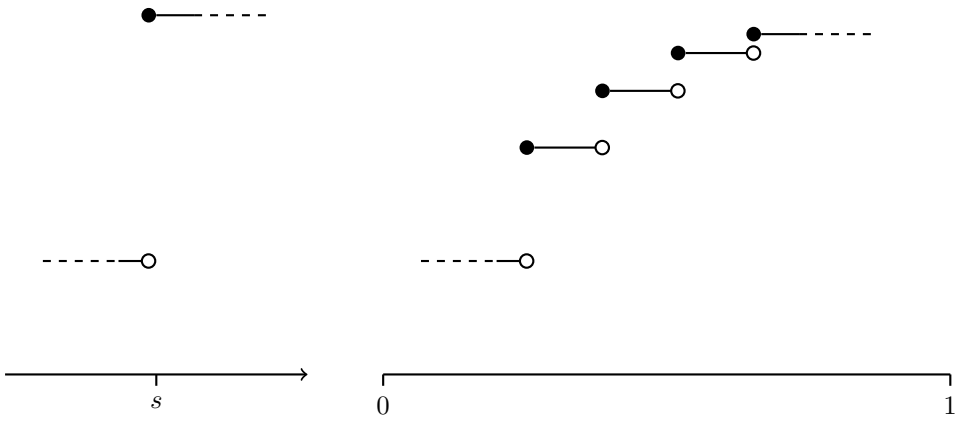

FIGURE 2.2. A sequence of jumps as in the previous figure, after rescaling in time and space, may converge to a part of $(x, S, \{e^s\}_{s\in S}) \in F'$. We sketch the x part on the left and a representative of the corresponding e^s on the right (this belongs to an equivalence class up to time rescaling). Clearly this convergence cannot take place in J_1, and, while convergence does hold in M_1 or M_2 (and indeed in E and F), the height of the individual (rather than cumulative) jumps will not be recorded.

$F'((0,\infty), \mathbb{R}^d)$ if the same holds for the respective restrictions to $F'([t_1, t_2], \mathbb{R}^d)$, for all $0 \leq t_1 < t_2$ such that $t_1, t_2 \notin S^x$.

We illustrate schematic versions of a sequence of elements in F' in Figure 2.1 which converge to the element of F' in Figure 2.2. The jumps stack up to one big jump, but the jumping behaviour is recorded in F'. Figures 2.3, 2.4 below show more complex behaviours in simulations.

We discuss some properties such as completeness and separability of the space F' in Appendix A.

2.4. **Rare events and point processes.** Since [LWZ81, Dav83], it has been known that the behaviour of the mean of heavy tailed processes is determined by the extremal observations. Hence, as in the context of extremal processes and records, we are lead to the study of rare events corresponding to the occurrence of abnormally large observations. In particular, this means that we need to impose some regularity of the tails of the distributions.

2.4.1. *Normalising threshold functions.* We assume that $\mathbf{X}_0, \mathbf{X}_1, \dots$ have proper tails, in the sense that there exists a normalising sequence of threshold functions $(u_n)_{n\in\mathbb{N}} : \mathbb{R}^+ \to \mathbb{R}^+$, where $\mathbb{R}^+ = (0, +\infty)$, satisfying the following properties (see [Hsi87]):

(1) For each n, the function u_n is nonincreasing, left continuous and such that
$$\lim_{\tau_1\to 0, \tau_2\to\infty} \mathbb{P}\left(u_n(\tau_2) < \|\mathbf{X}_0\| < u_n(\tau_1)\right) = 1;$$
(2) For each, $\tau \in \mathbb{R}^+$,
$$\lim_{n\to\infty} n\mathbb{P}\left(\|\mathbf{X}_0\| > u_n(\tau)\right) = \tau. \tag{2.14}$$

We observe that (2.14) requires that the average number of exceedances of $u_n(\tau)$, *i.e.*, events of the type $\|\mathbf{X}_j\| \geq u_n(\tau)$, for $j = 0, \dots, n-1$, is asymptotically constant and equal to $\tau > 0$, which can be interpreted as the asymptotic frequency of exceedances. The nonincreasing nature of u_n reflects the fact that the higher the frequency τ of observed exceedances, the lower the corresponding threshold $u_n(\tau)$ should be.

For every $z \in \mathbb{R}^+$, we define
$$u_n^{-1}(z) = \sup\{\tau > 0\colon\ z \leq u_n(\tau)\} \tag{2.15}$$
Observe that for each value z in the range of the r.v. $\|\mathbf{X}_0\|$, the function u_n^{-1} returns the asymptotic frequency $\tau = u_n^{-1}(z)$ that corresponds to the average number of exceedances of a threshold placed at the value z, among the n i.i.d. observations of the r.v. $\mathbf{X}_0$.

Also note that for all $\tau, s \in \mathbb{R}^+$,
$$u_n^{-1}(s) < \tau \quad \text{if and only if} \quad s > u_n(\tau). \tag{2.16}$$

2.4.2. *Point processes of rare events.* Multidimensional point processes are a powerful tool to record information regarding rare events (see for example [Pic71, Res87]), which can then be used to study record times ([Res75, Res87]), extremal processes [Dwa64, Lam64, Pic71, Res75, HT19], sums of heavy tailed random variables [Dav83, DH95, TK10b, TK10a, BKS12, BPS18]. In particular, they are very useful to keep track of the information within the clusters [Mor77, Hsi87, DH95, Nov02, BKS12, BPS18, FFMa20, PS20]. More specifically, the Mori-Hsing characterisation tells us that in nice situations the limiting process can essentially be described as having two components: a Poisson process determining the occurrences of clusters and an "orthogonal" point process describing the clustering structure. The description of the clustering component can be accomplished in different ways. In [Nov02] a natural generalisation of a compound Poisson point process is used. In [FFMa20], the authors use an outer measure to describe the piling of observations at the base cluster points of the Poisson process, which in the dynamical systems setting can be computed based on the action of the derivative of the map generating the dynamics. In [BS09], the authors introduced the *tail process*, which is a mechanism to describe the clusters and was subsequently used in [BKS12, BPS18], for example.

Here, we are going to use a description provided by the *piling process*, which is an adaptation to a more general setting of the tail process introduced in [BS09]. We will connect this to the outer measure of [FFMa20] and compute it in the context of dynamical systems.

Although we defer the formal definitions of point processes and point processes of rare events designed to keep all clustering information to Section 3.4 and Appendix C, we give here a brief description of the latter, which again has two components. The first is an underlying Poisson process on $\mathbb{R}_0^+ \times \mathbb{R}_0^+$ with intensity measure $\theta\text{Leb} \times \text{Leb}$, which can be represented by

$$M = \sum_{i=1}^{\infty} \delta_{(T_i,U_i)}, \quad \text{where } \delta_x \text{ denotes the Dirac measure at } x \in \mathbb{R}_0^+ \times \mathbb{R}_0^+,$$

so that for any measurable disjoint sets $A, A_1, \dots, A_\ell \subset \mathbb{R}_0^+ \times \mathbb{R}_0^+$, we have that $M(A)$ is a Poisson distributed random variable of intensity $\theta\text{Leb} \times \text{Leb}(A)$ and $M(A_1), \dots, M(A_\ell)$ are mutually independent. The parameter θ is defined formally in Section 3.3.2 and can be thought of as the reciprocal of the average number of exceedances in a cluster. The second component is the angular component of the spectral decomposition of the piling process, which will be defined in Section 3.3.4. It consists of a bi-infinite sequence $(Q_j)_{j\in\mathbb{Z}}$ which will decorate the second coordinate of each mass point of M and is such that $\|Q_j\| \to \infty$, as $|j| \to \infty$, $\min_{j\in\mathbb{Z}} \|Q_j\| = 1$. The sequences $(Q_{i,j})_{i,\in\mathbb{N},j\in\mathbb{Z}}$ associated to each mass point of M are mutually independent and also independent of the sequences $(T_i)_i$ and $(U_i)_i$. The distribution of each piling process, which corresponds to a sequence $(U_iQ_{i,j})_{j\in\mathbb{Z}}$, is designed to capture the behaviour the observations within a cluster of exceedances, which was initiated at time $j = 0$ and whose most severe exceedance has an asymptotic frequency given by U_i.

Remark 2.4. In order to have some intuition regarding the sequence $(Q_j)_{j\in\mathbb{Z}}$, we mention that in the case Ψ has the form (2.6), with $\mathcal{M}$ reduced to a repelling fixed point ζ where the invariant density is sufficiently regular, then, in the non-invertible case, Q_0 has a uniform distribution on $\mathbb{S}^{d-1}$, the unit sphere in $\mathbb{R}^d$, and for all $j \in \mathbb{N}$,

$$Q_j = \|(DT_\zeta)^j(Q_0)\|^d \frac{(DT_\zeta)^j(Q_0)}{\|(DT_\zeta)^j(Q_0)\|}, \tag{2.17}$$

where DT_ζ denotes the derivative of T at ζ and $(DT_\zeta)^j$ its j-fold product. For all negative j we have $Q_j = \infty$ a.s.

2.5. **Functional limit theorems for heavy tailed dynamical sums.** Throughout this section we assume that the process $\mathbf{X}_0, \mathbf{X}_1, \dots$ is obtained from a system as described in (2.1) and (2.2), where g is of type g_2, which together with some regular behaviour of the invariant measure μ in the vicinity of the maximal set $\mathcal{M}$, guarantees that there exists a sequence of positive real numbers $(a_n)_{n\in\mathbb{N}}$, such that

$$\lim_{n\to\infty} n\mathbb{P}(\|\mathbf{X}_0\| > ya_n) = y^{-\alpha}. \tag{2.18}$$

Hence, taking $\tau = x^{-\alpha}$ and $u_n(\tau) = a_n\tau^{-\frac{1}{\alpha}}$, equation (2.14) holds.

We are also going to assume that the piling process given in Definition 3.8 exists and is well defined, which in particular means that properties (1)–(4) given at the beginning of Section 3.3.4 hold. Section 2.2 provides examples of systems satisfying all our requirements.

The Lévy-Itô representation gives a nice way to describe the Lévy process as a functional of Poisson point process, whose intensity measure gives the Lévy measure that determines the

process (see [Sat13]). In the case of an α-stable Lévy process, it is usually identified through a limit of a Poisson integral of a Poisson point process $M_\alpha = \sum_{i=1}^\infty \delta_{(T_i,P_i)}$ with intensity measure $\mathrm{Leb}\times\nu_\alpha$, where the Lévy measure ν_α is such that $\nu_\alpha(\{x : \|x\| > y\}) = y^{-\alpha}$. Namely, when there is no clustering, for example, the limiting Lévy process can be written as

$$V(t) = \lim_{\varepsilon\to 0}\left(\sum_{T_i\le t} P_i \mathbb{1}_{\{\|P_i\|>\varepsilon\}} - \int_{\varepsilon<\|x\|\le 1} x\,\mathrm{d}\nu_\alpha(x)\right).$$

Hence, in this case, as further elaborated in Section 3.5, we consider a transformed version of the general rare events point processes by the map

$$\begin{aligned}\xi\colon \left(\mathbb{R}^d\cup\{\infty\}\right)\setminus\{0\} &\longrightarrow \mathbb{R}^d\\ x &\mapsto \begin{cases}(\|x\|)^{-\frac{1}{\alpha}}\frac{x}{\|x\|}, & \text{if } x\neq\infty\\ 0, & \text{otherwise}\end{cases},\end{aligned} \tag{2.19}$$

and by application of the continuous mapping theorem one obtains a limiting process, whose Poisson component can be written as $M_\alpha = \sum_{i=1}^\infty \delta_{(T_i,P_i)} = \sum_{i=1}^\infty \delta_{(T_i,U_i^{-\alpha})}$, with intensity measure $\theta\mathrm{Leb}\times\nu_\alpha$, with $\nu_\alpha(y) = \mathrm{d}(-y^{-\alpha})$, while the piling process is $(\mathcal{Q}_j)_{j\in\mathbb{Z}}$, with $\mathcal{Q}_j = \xi(Q_j)$.

Consider the partial sum process in $D([0,1],\mathbb{R}^d)$ defined by:

$$S_n(t) = \sum_{i=0}^{\lfloor nt\rfloor-1}\frac{1}{a_n}\mathbf{X}_i - tc_n, \qquad t\in[0,1], \tag{2.20}$$

where the sequence $(c_n)_{n\in\mathbb{N}}$ is such that $c_n = 0$ if $0<\alpha<1$ and

$$c_n = \frac{n}{a_n}\mathbb{E}\left(\mathbf{X}_0\mathbb{1}_{\|\mathbf{X}_0\|\le a_n}\right), \qquad \text{for } 1\le\alpha<2.$$

Our main goal in this section is to establish an invariance principle for S_n which keeps record of all the fluctuations during a cluster of high values which are responsible for a jump of S_n. As usual, for $1\le\alpha<2$, we need that the small contributions for the sum are close to the respective expectation, namely, for all $\delta>0$

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\mathbb{P}\left(\max_{1\le k\le n}\left\|\sum_{j=1}^k\left(\mathbf{X}_j\mathbb{1}_{\|\mathbf{X}_j\|\le\varepsilon a_n}\right)-\mathbb{E}\left(\mathbf{X}_j\mathbb{1}_{\|\mathbf{X}_j\|\le\varepsilon a_n}\right)\right\|\ge\delta a_n\right)=0. \tag{2.21}$$

In order to describe the limit, we assume the existence of the piling process, as in Definition 3.8. For $1<\alpha<2$, we will also need to assume that the sequence $(Q_j)_{j\in\mathbb{Z}}$, obtained from the spectral decomposition of the piling process, satisfies the assumption

$$\mathbb{E}\left(\left(\sum_{j\in\mathbb{Z}}\|\mathcal{Q}_j)\|\right)^\alpha\right)<\infty, \tag{2.22}$$

which in the case $\alpha=1$ should be replaced by

$$\mathbb{E}\left(\sum_{j\in\mathbb{Z}}\|\mathcal{Q}_j\|\log\left(\|\mathcal{Q}_j\|^{-1}\sum_{i\in\mathbb{Z}}\|\mathcal{Q}_j\|\right)\right)<\infty. \tag{2.23}$$

We can now state our main theorem regarding the behaviour of sums of heavy tailed observables.

Theorem 2.5. *Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, T, \mu)$ be a dynamical system as described in Section 2.2. Let $\mathbf{X}_0, \mathbf{X}_1, \dots$ be obtained from such a system as described in (2.1) and (2.2), where g is of type g_2 and condition (2.18) holds. Assume also that the piling process given in Definition 3.8 is well defined. Consider the continuous time process S_n given by (2.20). For $1 \le \alpha < 2$ assume further that condition (2.21) holds and, for $1 < \alpha < 2$, assume also (2.22), while for $\alpha = 1$ assume (2.23), instead. Then S_n converges in F' to $\underline{V} := (V, disc(V), \{e_V^s\}_{s \in disc(V)})$, where V is an α-stable Lévy process on $[0,1]$ which can be written as*

$$V(t) = \sum_{T_i \le t} \sum_{j \in \mathbb{Z}} U_i^{-\frac{1}{\alpha}} \mathcal{Q}_{i,j},$$

for $1 \le \alpha < 2$ and

$$V(t) = \lim_{\varepsilon \to 0} \left(\sum_{T_i \le t} \sum_{j \in \mathbb{Z}} U_i^{-\alpha^{-1}} \mathcal{Q}_{i,j} \mathbb{1}_{\{\|U_i^{-\alpha^{-1}} \mathcal{Q}_{i,j}\| > \varepsilon\}} - t\theta \int_0^{+\infty} \mathbb{E}\left(y \sum_{j \in \mathbb{Z}} \mathcal{Q}_j \mathbb{1}_{\{\varepsilon < y \|\mathcal{Q}_j\| \le 1\}} \right) d(-y^{-\alpha}) \right)$$

for $0 < \alpha < 1$; and the excursions can be represented by

$$e_V^s(t) = V(T_i^-) + U_i^{-\frac{1}{\alpha}} \sum_{j \le \lfloor \tan(\pi(t - \frac{1}{2})) \rfloor} \mathcal{Q}_{i,j}, \qquad t \in [0,1],$$

where $(T_i)_{i \in \mathbb{N}}$, $(U_i)_{i \in \mathbb{N}}$ are as described above (see also (3.33)), $(\mathcal{Q}_{i,j})_{i \in \mathbb{N}, j \in \mathbb{Z}}$ is such that $\mathcal{Q}_{i,j} = \xi(Q_{i,j})$, where $(Q_{i,j})_{i \in \mathbb{N}, j \in \mathbb{Z}}$ are as in (3.23).

We illustrate the theorem in Figures 2.3 and 2.4. The latter blow-up shows how the stacking of jumps can occur S_n, and the requirement for convergence in F' in the limit.

Remark 2.6. Recall that when Ψ is as in (2.6) and $\mathcal{M}$ is reduced to an hyperbolic periodic point, the piling process is well defined and the $(Q_{i,j})_{i \in \mathbb{N}, j \in \mathbb{Z}}$ are as in Remark 2.4. Also note that when $\mathcal{M}$ is reduced to a generic point, we have no clustering and then the result holds with the trivial $Q_j = \infty$, for all $j \in \mathbb{Z} \setminus \{0\}$ and Q_0 as before.

2.6. **Enriched extremal process dynamics in the presence of clustering.** Extremal processes are a very useful tool to study the stochastic behaviour of maxima and records (see [Res87]). We define the partial maxima associated to the sequence $\mathbf{X}_0, \mathbf{X}_1, \dots$ by

$$M_n := \max\{\|\mathbf{X}_0\|, \dots, \|\mathbf{X}_{n-1}\|\} = \bigvee_{i=0}^{n-1} \|\mathbf{X}_i\|. \tag{2.24}$$

Finding a distributional limit for M_n is one of the first goals in Extreme Value Theory.

Definition 2.7. We say that we have an *Extreme Value Law* (EVL) for M_n if there is a non-degenerate d.f. $H : \mathbb{R}_0^+ \to [0,1]$ with $H(0) = 0$ and, for every $\tau > 0$, there exists a sequence of thresholds $u_n(\tau)$, $n = 1, 2, \dots$, satisfying equation (2.14) and for which the following holds:

$$\mathbb{P}(M_n \le u_n(\tau)) \to \bar{H}(\tau), \quad \text{as } n \to \infty. \tag{2.25}$$

where $\bar{H}(\tau) := 1 - H(\tau)$ and the convergence is meant at the continuity points of $H(\tau)$.

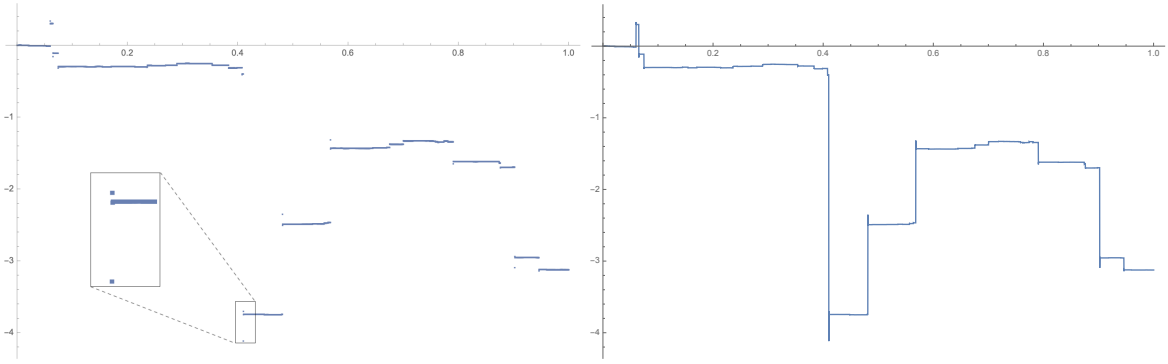

FIGURE 2.3. Plots of a finite sample simulation of $S_n(t)$ with $n = 5000$, where $X_j = \psi \circ T^j(x)$, where $T(x) = 3x \mod 1$, $\psi(x) = |x - 1/8|^{-2} - |x - 3/8|^{-2}$. The left plot includes a zoom at one of the jumps and the right plot is the same simulation, but the completed graph version, to give a different view of the jumps. Note that $\zeta = 1/8$ is a periodic point of period 2 and $T(\zeta) = 3/8$. Here $DT_\zeta = DT_{T(\zeta)} = 3$.

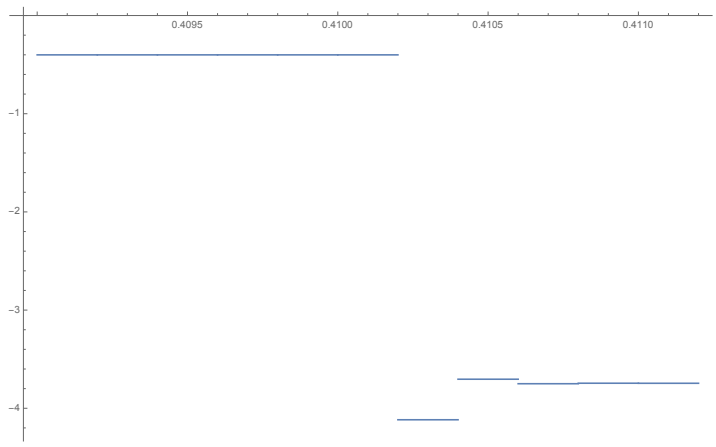

FIGURE 2.4. Blowup of the previous graph at the jump observed near 0.4: asymptotically the four jumps seen here will happen instantaneously, necessitating an appropriate space for convergence.

It turns out that the limit H allows us to describe the functional limit for associated extremal processes.

In this context, we now consider the continuous time process $\{Z_n(t) : 0 \le t \le 1\}$ defined by

$$Z_n(t) := u_n^{-1}(M_{\lfloor nt \rfloor}), \qquad t \in (0, \infty) \tag{2.26}$$

Recall that $u_n^{-1}(z)$ gives the asymptotic frequency of exceedances of a threshold placed at z and therefore Z_n is non-increasing.

For each $n \ge 1$, $Z_n(t)$ is a random graph with values in $D((0, \infty), \mathbb{R}))$, which can be embedded into F', as in (2.11). The process Z_n will be shown to converge, in F', to the process $\underline{Z}$, whose

first component in D is Z_H, which can be described by the finite dimensional distributions:

$$\mathbb{P}(Z_H(t_1) \geq y_1, \dots, Z_H(t_k) \geq y_k) = \bar{H}^{t_1}\left(\bigvee_{i=1}^{k}\{y_i\}\right)\bar{H}^{t_2-t_1}\left(\bigvee_{i=2}^{k}\{y_i\}\right)\cdots \bar{H}^{t_k-t_{k-1}}(y_k), \qquad (2.27)$$

with $0 \leq t_1 < t_2 < \cdots < t_k \leq 1$. By the Kolmogorov extension theorem such a process is well defined and we call it an *extremal process*, although, strictly speaking, this is a transformed version of the original extremal processes studied by Resnick [Res87]. The relation between the two is obtained through the connection between the levels $u_n(\tau)$ and the more classical linear normalising sequences $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ and $(b_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ such that we can write $u_n = y/a_n + b_n$ and

$$n\mathbb{P}(X_0 > u_n) = n\mathbb{P}(a_n(X_0 - b_n) > y) \to \tau$$

with $\tau = f(y)$ for some homeomorphism f, then, as in (2.24),

$$\mathbb{P}(M_n \leq u_n) = \mathbb{P}(a_n(M_n - b_n) \leq y) \to G(y)$$

where $G = \bar{H} \circ f$. Then, if $Y_n(t) = a_n(M_{\lfloor nt\rfloor+1} - b_n)$ and Y_G denotes the respective extremal process obtained in [Res87], we have that $Z_n(t) = f(Y_n(t))$ and $Z_H(t) = f(Y_G(t))$.

Remark 2.8. Depending on the type of limit law that applies, $f(y)$ is of one of the following three types: $f_1(y) = \mathrm{e}^{-y}$ for $y \in \mathbb{R}$, $f_2(y) = y^{-\alpha}$ for $y > 0$, and $f_3(y) = (-y)^{\alpha}$ for $y \leq 0$.

Theorem 2.9. *Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, T, \mu)$ be a dynamical system as described in Section 2.2. Let $\mathbf{X}_0, \mathbf{X}_1, \dots$ be obtained from such a system as described in (2.1) and assume that the piling process given in Definition 3.8 is well defined. Consider the continuous time process Z_n defined by (2.26). Then Z_n converges in F' to $(Z_H, disc(Z_H), \{e^s_{Z_H}\}_{s\in disc(Z_H)})$, where Z_H is defined as in (2.27), with $\bar{H}(\tau) = \mathrm{e}^{-\theta\tau}$, and the excursions can be represented by*

$$e^s_{Z_H}(t) = \min\left\{Z_H(s-), \inf_{j\leq\lfloor\tan(\pi(t-\frac{1}{2}))\rfloor} Z_H(s)\cdot Q^s_j\right\}, \quad t \in [0,1]$$

where each sequence $(Q^s_j)_{j\in\mathbb{N}}$ is independent of $Z_H(s^-)$ and with common distribution given by (3.23). Moreover, Z_H can be seen as a Markov jump process with

$$\mathbb{P}(Z_H(t+s) \geq y \mid Z_H(s) = z) = \begin{cases} \mathrm{e}^{-\theta t y} & \text{if } y < z \\ 0 & \text{if } y \geq z \end{cases}, \qquad \text{for } t, s > 0.$$

The parameter of the exponential holding time in state z is θz and given that a jump is due to occur the process jumps from z to $[0, y)$ with probability

$$\Pi(z, [0, y)) = \begin{cases} \frac{y}{z} & \text{if } y < z \\ 1 & \text{if } y \geq z \end{cases}.$$

Remark 2.10. Alternatively and similarly to extremal processes that can be described as a projection into D of a point process (see [Res87, Equation (4.20)]), we can describe the limit of Z_n as the projection into F' of the point process $N = \sum_{i=1}^{\infty}\delta_{(T_i, U_i\tilde{\mathbf{Q}}_i)}$, given in (3.33). Namely, $Z_H(t) = \inf\{U_i\colon\ T_i \leq t\}$, $disc(Z_H) = \{T_i\colon\ :\ i \in \mathbb{N}\}$ and

$$e^{T_i}_{Z_H}(t) = \min\left\{\inf\{U_j\colon\ T_j < T_i\}, \inf_{j\leq\lfloor\tan(\pi(t-\frac{1}{2}))\rfloor} U_i\cdot Q_{i,j}\right\}, \quad t \in [0,1].$$

2.7. **Record point processes.** The study of record times of observational data has important applications in the study of natural phenomena. Consider the original sequences $(\mathbf{X}_n)_{n\in\mathbb{N}_0}$ and $(M_n)_{n\in\mathbb{N}}$ given in (2.24), and let $t_1 = 0$. Define the strictly increasing sequence $(t_k)_{k\in\mathbb{N}}$:

$$t_k := \inf\{j > t_{k-1} : \|\mathbf{X}_j\| > M_j\}. \tag{2.28}$$

This sequence $(t_k)_k$ corresponds to the *record times* associated to M_n, namely the times where M_n jumps. In the presence of clustering, record times may collapse in the limit and it is important to keep track of possible increments on the number of records occurring during the clusters. Two possible approaches are to consider the enriched limits in F' of the extremal processes Z_H or simply to project directly from the point processes N, as done in [BPS18]. Since following the former approach, to be able to handle the possibility of having vanishing jump points in the limit we would need to consider more restrictive subspaces to obtain continuity and then apply the Continuous Mapping Theorem (CMT), we will use the latter approach so that we can benefit from the work already carried in [BPS18] and reduce the length of the exposition.

We follow [BPS18, Section 5] closely, although we make some adjustments due to the fact that the spatial component of our point processes lives in l_∞ instead of l_0, see Section 3.3.3 for the definitions of these sequence spaces as well as $\tilde{l}_\infty$. The main advantage here is that we obtain the convergence of the record times point processes for stationary vector valued sequences with much more general distributions rather than regularly varying sequences (or sequences that could be monotonically transformed into regularly varying ones) as in [BPS18].

In order to count the number of records during the clusters we define, for $y \geq 0$ and $\mathbf{x} = (x_j)_j \in l_\infty$,

$$R^{\mathbf{x}}(y) = \sum_{j\in\mathbb{Z}} \mathbb{1}_{\{\|x_j\| < (y \wedge \inf_{i<j} \|x_i\|)\}},$$

which gives the number of record asymptotic frequencies corresponding to the smallest observations in $\mathbf{x}$ that have beaten the frequency y, which must be finite because $\lim_{|j|\to\infty} \|x_j\| = \infty$. For $\gamma = \sum_{i=1}^\infty \delta_{(t_i,\tilde{\mathbf{x}}^i)} \in \mathcal{N}^{\#}_{\mathbb{R}_0^+\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}}$, let $h(\gamma) \in D((0,\infty),\mathbb{R})$ be defined as in (4.1) and define the record point process in $\mathcal{N}^{\#}_{(0,+\infty)}$ by

$$R_\gamma = \sum_{i\in\mathbb{N}} \delta_{t_i} R^{\mathbf{x}^i}(h(\gamma)(t_i^-)), \quad \text{where } \mathbf{x}^i \in l_\infty \text{ is any chosen representative of } \tilde{\mathbf{x}}^i.$$

In order to count the number of records at finite time we consider the empirical process:

$$N_n'' := \sum_{i=0}^\infty \delta_{\left(\frac{i}{n}, u_n^{-1}(\|\mathbf{X}_i\|)\frac{\mathbf{X}_i}{\|\mathbf{X}_i\|}\right)},$$

which we consider as defined in $\mathcal{N}^{\#}_{\mathbb{R}_0^+\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}}$ by embedding the second coordinate in $\mathbb{R}^d$ into l_∞ by adding a sequence of ∞ before and after that entrance as in (3.19).

Note that $R_{N_n''}$ indeed counts the number of records of the process $\mathbf{X}_0, \mathbf{X}_1, \ldots$, namely,

$$R_{N_n''} = \sum_{i=0}^\infty \delta_{\frac{i}{n}} \mathbb{1}_{\{\|\mathbf{X}_i\| > M_i\}}.$$

Theorem 2.11. *Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, T, \mu)$ be a dynamical system as described in Section 2.2. Let $\mathbf{X}_0, \mathbf{X}_1, \ldots$ be obtained from such a system as described in (2.1) and assume that the piling process given in Definition 3.8 is well defined and that $\mathbb{P}(\text{all finite } Q_j\text{'s are mutually different}) = 1$. Then $R_{N_n''}$ converges weakly to R_N, in $\mathcal{N}^{\#}_{(0,+\infty)}$. The limiting process is a compound Poisson process which can be represented as*

$$R_N = \sum_{i \in \mathbb{Z}} \delta_{\tau_i} \kappa_i \,,$$

where $\sum_{i\in\mathbb{Z}} \delta_{\tau_i}$ is a Poisson point process on $(0, \infty)$ with intensity measure $x^{-1}\, dx$ and $(\kappa_i)_i$ is a sequence of i.i.d. random variables independent of it with the same distribution as the integer valued random variable $R^{\mathbf{Q}}(U^{-1})$ where U is a uniformly distributed random variable independent of $\mathbf{Q} = (Q_j)_j$.

Remark 2.12. In the periodic point case, condition $\mathbb{P}(\text{all finite } Q_j\text{'s are mutually different}) = 1$. is trivially satisfied.

3. General complete convergence of multi-dimensional cluster point processes

In this section we present our technical tools and various results in the context of general stochastic processes: the shift σ marries this with a dynamical point of view, but the work here holds in wide generality.

Let $\mathcal{V} = \mathbb{R}^d$, for some $d \in \mathbb{N}$, where we consider a norm which we denote by $\|\cdot\|$. For definiteness, we may consider the usual Euclidean norm. We will be considering the spaces of one-sided and two-sided $\mathcal{V}$-valued sequences, which we will denote, respectively, by $\mathcal{V}^{\mathbb{N}_0}$ and $\mathcal{V}^{\mathbb{Z}}$, where we consider the one-sided and two-sided shift operators defined by $\sigma : \mathcal{V}^{\mathbb{N}_0,\mathbb{Z}} \to \mathcal{V}^{\mathbb{N}_0,\mathbb{Z}}$, where

$$\sigma((x)_i) = ((x)_{i+1}). \tag{3.1}$$

Consider a stationary sequence of random vectors $\mathbf{X}_0, \mathbf{X}_1, \ldots$, taking values on $\mathcal{V} = \mathbb{R}^d$, which we will identify with the respective coordinate-variable process on $(\mathcal{V}^{\mathbb{N}_0}, \mathcal{B}^{\mathbb{N}_0}, \mathbb{P})$, given by Kolmogorov's existence theorem, where $\mathcal{B}^{\mathbb{N}_0}$ is the σ-field generated by the coordinate functions $Z_n : \mathcal{V}^{\mathbb{N}_0} \to \mathcal{V}$, with $Z_n(x_0, x_1, \ldots) = x_n$, for $n \in \mathbb{N}_0$.

Note that, under these identifications, we can write:

$$Z_{i-1} \circ \sigma = Z_i, \quad \text{for all } i \in \mathbb{N}.$$

Since, we assume that the process is stationary, then $\mathbb{P}$ is σ-invariant. Note that $Z_i = Z_0 \circ \sigma^i$, for all $i \in \mathbb{N}_0$, where σ^i denotes the i-fold composition of σ, with the convention that σ^0 denotes the identity map on $\mathcal{V}^{\mathbb{N}_0,\mathbb{Z}}$.

In what follows, for every $A \in \mathcal{B}^{\mathbb{N}_0,\mathbb{Z}}$, we denote the complement of A as $A^c := \mathcal{V}^{\mathbb{N}_0,\mathbb{Z}} \setminus A$.

3.1. Identifying clusters.

Our goal is to study the impact of clustering on the convergence of general multidimensional point processes. As mentioned earlier, information regarding the observations within the same cluster gets collapsed at the same time point, which makes it hard to recover it from the limiting process. In order to keep track of that information, we need to start by identifying clusters.

There are two main approaches to identify clusters, which are commonly referred to as declustering procedures. One is the blocking method, which involves dividing observations into blocks and stipulating that all abnormal observations (typically with very high norm) within one such block belong to the same cluster. The second is the runs declustering procedure, which determines that a cluster ends when a run of consecutive non-abnormal observations occurs. Here we will impose a condition on the dependence structure of the stochastic processes, in which case both procedures are essentially equivalent.

We start with the blocking procedure, which, in fact, serves two purposes. Namely, not only will it separate clusters, it will also introduce time gaps between the blocks in order to restore some independence between them. The size of the blocks must be sensitively tuned so that the blocks are neither too long so that they do not separate clusters, nor too small so that they do not split clusters apart. The same applies to time gaps between the blocks, which are created by disregarding observations. Hence, one must not disregard too much information, but on the other hand the time gap must be sufficiently large in order to create sufficient loss of memory. Following the classical scheme [LLR83], considering a finite sample of size n, we split the data into $k_n \in \mathbb{N}$ blocks of size $r_n := \lfloor n/k_n \rfloor$ and take time gaps of size $t_n \in \mathbb{N}$. This way, we define sequences $(k_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$, which we assume to be such that

$$k_n, r_n, t_n \xrightarrow[n\to\infty]{} \infty \quad \text{and} \quad k_n t_n = o(n). \tag{3.2}$$

Regarding the runs declustering, for a finite sample of size n, we set the run length $q_n \in \mathbb{N}$ with the aim that all abnormal observations occurring within a time difference of at most q_n units between each other belong to the same cluster. The sequence $(q_n)_{n\in\mathbb{N}}$ must be chosen so that

$$q_n = o(r_n), \tag{3.3}$$

and also so that it satisfies conditions $Д_{q_n}$, $Д'_{q_n}$, below. Note that $q_n = q$ for all $n \in \mathbb{N}$ and some $q \in \mathbb{N}$ is a possibility here.

3.2. **Dependence structure.** In order to prove the main convergence results we need to introduce some conditions on the dependence structure of the stationary processes and therefore introduce the the following objects. We use follow more or less the notation used in [FFMa18, FFMa20].

Let $A \in \mathcal{B}^{\mathbb{N}_0,\mathbb{Z}}$ be an event, let J be an interval contained in $[0,\infty)$. We define

$$\mathscr{W}_J(A) := \bigcap_{i\in J\cap\mathbb{N}_0} \sigma^{-i}(A^c), \quad \mathscr{W}_J^c(A) := (\mathscr{W}_J(A))^c = \bigcup_{i\in J\cap\mathbb{N}_0} \sigma^{-i}(A). \tag{3.4}$$

Let $i, j \in \mathbb{N}$ and set

$$\mathscr{W}_j(A) := \mathscr{W}_{[0,j)}(A), \quad \mathscr{W}_{i,j}(A) := \mathscr{W}_{[i,j)}(A), \text{ when } i \le j-1, \quad \mathscr{W}_{i,j}(A) := \mathcal{V}^{\mathbb{N}_0,\mathbb{Z}} \text{ when } i \ge j.$$

Moreover, $\mathscr{W}_j^c(A) = (\mathscr{W}_j(A))^c$ and $\mathscr{W}_{i,j}^c(A) = (\mathscr{W}_{i,j}(A))^c$.

For the event $A \in \mathcal{B}^{\mathbb{N}_0,\mathbb{Z}}$ and $j \in \mathbb{N}$,

$$A^{(j)} := A \cap \sigma^{-1}(A^c) \cap \cdots \cap \sigma^{-j}(A^c), \tag{3.5}$$

and for $j = 0$ we simply define $A^{(0)} := A$.

In what follows, for some $a > 0$, $y \in \mathbb{R}^d$ and a set $A \subset \mathbb{R}^d$, we set $aA = \{ax : x \in A\}$ and $A + y = \{x + y : x \in A\}$. We define the class of sets

$$\mathscr{F} = \Big\{ \{(x_j)_j \in \mathcal{V}^{\mathbb{N}_0,\mathbb{Z}} \colon x_j \in H_j,\ j = 0, \dots, m\} \colon\ H_j \in \mathcal{B}_{\mathcal{V}},\ j = 0, \dots, m,\ m \in \mathbb{N} \Big\}, \tag{3.6}$$

where $\mathcal{B}_{\mathcal{V}}$ is the Borel σ-algebra of $\mathcal{V}$. Note that $\mathscr{F}$ is a field.

For each $\ell = 1, \dots, m \in \mathbb{N}$ suppose

$$A_\ell \in \mathscr{F} \text{ and } J_\ell = [a_\ell, b_\ell), \tag{3.7}$$

where $0 \le a_1 < b_1 \le a_2 < b_2 \le \dots \le a_m < b_m \le 1$. Then for each $n \in \mathbb{N}$ define

$$\tilde{J}_{k_n,\ell} := k_n J_\ell, \quad J_{n,\ell} := \big[(\lceil k_n a_\ell \rceil - 1) r_n, (\lfloor k_n b_\ell \rfloor + 1) r_n\big), \tag{3.8}$$

$$A_{n,\ell} = \left\{ \left(u_n^{-1}(\|\mathbf{X}_j\|) \frac{\mathbf{X}_j}{\|\mathbf{X}_j\|} \right)_j \in \mathcal{V}^{\mathbb{N}_0} \colon\ \left(u_n^{-1}(\|\mathbf{X}_j\|) \frac{\mathbf{X}_j}{\|\mathbf{X}_j\|} \right)_j \in A_\ell \right\}, \tag{3.9}$$

where u_n is defined as in (2.14), k_n is as in (3.2). We discuss this normalisation further in (3.13) and (3.14).

We introduce a mixing condition which is specially designed for the application to the dynamical setting.

Condition ($Д_{q_n}$)**.** We say that $Д_{q_n}$ holds for the sequence $\mathbf{X}_0, \mathbf{X}_1, \dots$ if there exist sequences $(k_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$ and $(q_n)_{n\in\mathbb{N}}$ satisfying (3.2) and (3.3), such that for every $m, t, n \in \mathbb{N}$ and every J_ℓ and A_ℓ, with $\ell = 1, \dots, m$, chosen as in (3.7), we have

$$\left| \mathbb{P}\left(A_{n,\ell}^{(q_n)} \cap \bigcap_{i=\ell}^{m} \mathscr{W}_{J_{n,i}}\left(A_{n,i}^{(q_n)}\right) \right) - \mathbb{P}\left(A_{n,\ell}^{(q_n)}\right) \mathbb{P}\left(\bigcap_{i=\ell}^{m} \mathscr{W}_{J_{n,i}}\left(A_{n,i}^{(q_n)}\right) \right) \right| \le \gamma(n,t), \tag{3.10}$$

where $\min\{J_{n,\ell} \cap \mathbb{N}_0\} \ge t$ and $\gamma(n,t)$ is decreasing in t for each n and $\lim_{n\to\infty} n\gamma(n, t_n) = 0$, where $A_{n,\ell}^{(q_n)}$ is given by (3.9) and (3.5).

This mixing condition is much milder than similar conditions used in the literature and is particularly suited for applications to dynamical systems, since it is easily verified for systems with sufficiently fast decay of correlations, see for example the discussion in Section 2.2.

Condition ($Д'_{q_n}$)**.** We say that $Д'_{q_n}$ holds for the sequence $\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \dots$ if there exist sequences $(k_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$ and $(q_n)_{n\in\mathbb{N}}$ satisfying (3.2) and (3.3), such that for every $A_1 \in \mathcal{F}$, we have

$$\lim_{n\to\infty} n\mathbb{P}\left(A_{n,1}^{(q_n)} \cap \mathscr{W}^c_{[q_n+1, r_n)}\left(A_{n,1}\right) \right) = 0,$$

where $A_{n,1}^{(q_n)}$ is given by (3.9) and (3.5).

Remark 3.1. Condition $Д'_{q_n}$ forbids the appearance of new abnormal observations (the occurrence of $A_{n,\ell}$), within the same block, once a run of q_n consecutive non-abnormal observations has been realised. This means that $Д'_{q_n}$ establishes a connection between the two declustering procedures and, in particular, requires that no more than one cluster should occur within one block.

Remark 3.2. Suppose $(q_n)_{n\in\mathbb{N}}$ and $(\tilde{q}_n)_{n\in\mathbb{N}}$ are two sequences satisfying (3.3) and $q_n \le \tilde{q}_n$, for all $n \in \mathbb{N}$. Then $Д'_{q_n}$ implies $Д'_{\tilde{q}_n}$. As in Section 2.2, for nice dynamical systems when

clustering is created by a periodic point of prime period q, then condition $Д'_q$ holds, in the sense that $q_n = q$, for all n.

Next we state a stronger version of $Д'_{q_n}$, which, in some cases may be easier to check. We set

$$U(\tau) = \left\{(x_j)_j \in \mathcal{V}^{\mathbb{N}_0,\mathbb{Z}} \colon x_0 \in B(0,\tau)\right\} \in \mathscr{F}, \tag{3.11}$$

where, for $x \in \mathcal{V}$ and $\epsilon > 0$, we denote by $B(x,\epsilon)$,the ball centred at x of radius ϵ. Then, following (3.9) and (3.5), we define:

$$\begin{aligned} U_n(\tau) &= \{\|\mathbf{X}_0\| > u_n(\tau)\} \\ U_n^{(q_n)}(\tau) &= \{\|\mathbf{X}_0\| > u_n(\tau), \|\mathbf{X}_1\| \le u_n(\tau), \dots, \|\mathbf{X}_{q_n}\| \le u_n(\tau)\}. \end{aligned}$$

Condition ($\tilde{Д}'_{q_n}$)**.** We say that $\tilde{Д}'_{q_n}$ holds for the sequence $\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \dots$ if there exist sequences $(k_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$ and $(q_n)_{n\in\mathbb{N}}$ satisfying (3.2) and (3.3), such that for every $\tau > 0$, we have

$$\lim_{n\to\infty} n\mathbb{P}\left(U_n(\tau) \cap \mathscr{W}^c_{[q_n+1,r_n)}\left(U_n(\tau)\right)\right) = 0. \tag{3.12}$$

For A_ℓ, as in (3.7), let $\tau^* > \sup\{\|x\| \colon\ x \in H_0\}$. Then, by (2.16), we have

$$\left\| u_n^{-1}(\|\mathbf{X}_0\|)\frac{\mathbf{X}_0}{\|\mathbf{X}_0\|}\right\| < \tau^* \Leftrightarrow \|\mathbf{X}_0\| > u_n(\tau^*).$$

Therefore, it is clear that $A_{n,\ell} \subset U_n(\tau^*)$, so if $\tilde{Д}'_{q_n}$ holds then so does $Д'_{q_n}$.

Remark 3.3. Condition $\tilde{Д}'_{q_n}$ is already weaker than [BPS18, Assumption 1.1], which had been used in previous papers (see, for example, [DH95, Equation (2.8)], [Seg05, Equation (3)], [BS09, Condition 4.1], [BKS12, Condition 2.1]) and was introduced in [Smi92]. We also remark that all these conditions allow for the appearance of clustering which already makes them weaker than conditions D' from [Dav83] or LD(ϕ_0) from [TK10a], which imply $Д'_{q_n}$ with $q_n = 1$ for all $n \in \mathbb{N}$.

3.3. **Bookkeeping of clusters.** In this section, we introduce a device called the *piling process*, which is designed to keep track of the clustering oscillations. It is an adaptation of the *tail process*, introduced in [BS09], to a more general setting. In [BS09] and subsequent papers (for example, [BKS12, BPS18]), the tail process was always defined under the assumption that the original process $\mathbf{X}_0, \mathbf{X}_1, \dots$ is *jointly regularly varying* (see Definition 3.19), which is not natural to assume *a priori* in the dynamical systems setting. This assumption (jointly regularly variation) together with an assumption on the dependence structure stronger than $\tilde{Д}'_{q_n}$ ([BS09, Condition 4.1]) allowed the authors there to prove the existence of the tail process and several very useful properties about it. Motivated by the applications to dynamical systems, here, we will work in a more general setting (without assuming, *a priori*, jointly regular variation) and since $Д'_{q_n}$ is even weaker than $\tilde{Д}'_{q_n}$, some of the properties of the tail process will be required as adapted assumptions in the definition of the piling process. One of the advantages is that we obtain very general enriched functional limits for extremal processes and record point processes, for example, without assuming regularly varying tails and much less the restrictive structure imposed by jointly regular variation.

3.3.1. *The normalisation of the blocks.* We recall that under assumption $Д'_{q_n}$ the information regarding to the structure of the clusters is kept in each block of size r_n, which we are going to normalise in the following way, by defining for each $i < j \in \{0, \dots, n\}$

$$\mathbb{X}_n^{i,j} = \left(u_n^{-1}(\|\mathbf{X}_i\|) \frac{\mathbf{X}_i}{\|\mathbf{X}_i\|}, \dots, u_n^{-1}(\|\mathbf{X}_{j-1}\|) \frac{\mathbf{X}_{j-1}}{\|\mathbf{X}_{j-1}\|} \right), \quad \mathbb{X}_{n,i} := \mathbb{X}_n^{(i-1)r_n, ir_n} \tag{3.13}$$

so that $\mathbb{X}_{n,i}$ denotes the i-th normalised block. We also use the notation $\mathbb{X}_n^j = \mathbb{X}_n^{j,j+1}$, for all $j = 0, \dots, n$.

Note that the normalisation used is such that the norm of each normalised variable in each block is equal to the asymptotic frequency $\tau = u_n^{-1}(\|\mathbf{X}_j\|)$ corresponding to the mean number of exceedances of a threshold placed at the value $\|\mathbf{X}_j\|$, among the first n observations of the process.

In particular, observe that by (2.16), for all $\tau > 0$, we have

$$\left\|\mathbb{X}_n^j\right\| = \left\| u_n^{-1}(\|\mathbf{X}_j\|) \frac{\mathbf{X}_j}{\|\mathbf{X}_j\|} \right\| < \tau \Leftrightarrow u_n^{-1}(\|\mathbf{X}_j\|) < \tau \Leftrightarrow \|\mathbf{X}_j\| > u_n(\tau). \tag{3.14}$$

3.3.2. *The Extremal Index.* Before we characterise the piling process, we define the Extremal Index (EI), denoted by $\theta \in [0, 1]$, which was formally introduced by Leadbetter in [Lea83] and measures the degree of clustering of exceedances. When $\theta = 1$ we have no clustering and a small θ means intense clustering. A common interpretation for the EI is that it is reciprocal of the average cluster size (see [AFF20]). We define the EI following O'Brien's formula ([O'B87]) and assume that, for all $\tau > 0$, we have

$$\theta = \lim_{n \to \infty} \frac{\mathbb{P}(U_n^{(q_n)}(\tau))}{\mathbb{P}(U_n(\tau))}. \tag{3.15}$$

3.3.3. *Underlying spaces.* Let $\dot{\mathcal{V}} = \overline{\mathbb{R}^d} \setminus \{0\} = \left(\mathbb{R}^d \cup \{\infty\}\right) \setminus \{0\}$, $\mathcal{V} = \mathbb{R}^d$ and define

$$l_\infty = \left\{ \mathbf{x} = (x_j)_j \in \dot{\mathcal{V}}^{\mathbb{Z}} \colon \lim_{|j| \to \infty} \|x_j\| = \infty \right\} \quad \text{and} \quad l_0 = \left\{ \mathbf{x} = (x_j)_j \in \mathcal{V}^{\mathbb{Z}} \colon \lim_{|j| \to \infty} \|x_j\| = 0 \right\}.$$

The piling process will be defined to take values in l_∞, while the tail process lives in l_0. The space l_∞ will borrow the metric structure of l_0 by means of the map $P \colon l_\infty \longrightarrow l_0$ given by $P((x_j)_j) = (p(x_j))_j$, where

$$\begin{aligned} p \colon \dot{\mathcal{V}} &\longrightarrow \mathcal{V} \\ x &\mapsto \begin{cases} \frac{x}{\|x\|^2}, & \text{if } x \neq \infty \\ 0, & \text{otherwise} \end{cases}. \end{aligned}$$

Lemma 3.4. *The map P is invertible.*

Proof. By definition of P, we only need to show that p is invertible. First note that $p(x) = 0$ if and only if $x = \infty$. Let $x, y \in \dot{\mathcal{V}} \setminus \{\infty\}$. Then $p(x) = p(y)$ implies that

$$x = \frac{\|x\|^2}{\|y\|^2}\, y. \tag{3.16}$$

Let $c = \frac{\|x\|^2}{\|y\|^2}$. Then $x = cy$ which means that $\|x\| = c\|y\|$. Substituting back in (3.16) we obtain $x = c^2 y$. Hence, we must have $cy = c^2 y$, which implies that $c = 1$ and, therefore, $x = y$. Hence P is one-to-one.

To see that P is onto, we let $\mathbf{y} \in l_0$ and show that there exists $\mathbf{x} \in l_\infty$ such that $P(\mathbf{x}) = \mathbf{y}$. For all the $j \in \mathbb{Z}$ such that $y_j = 0$, we set $x_j = \infty$. For all the other $j \in \mathbb{Z}$, we require $x_j/\|x_j\|^2 = y_j$, which implies that $\|y_j\| = 1/\|x_j\|$. But then, $x_j = \|x_j\|^2 y_j = y_j/\|y_j\|^2$, which means that by setting $x_j = p(y_j)$, for all such $j \in \mathbb{Z}$, we have defined the desired $\mathbf{x} \in l_\infty$. □

As in [BPS18], in l_0, we consider the supremum norm given by

$$\|\mathbf{x}\|_\infty = \sup_{j\in\mathbb{Z}} \|x_j\|,$$

and the complete metric defined on $l_0 \setminus \{\mathbf{0}\}$, where $\mathbf{0} = (0)_j$,

$$d'(\mathbf{x}, \mathbf{y}) = (\|\mathbf{x} - \mathbf{y}\|_\infty \wedge 1) \vee \left| \frac{1}{\|\mathbf{x}\|_\infty} - \frac{1}{\|\mathbf{y}\|_\infty} \right|. \tag{3.17}$$

Now, we consider the metric d defined on $l_\infty \setminus \{(\infty)_j\}$ given by

$$d(\mathbf{x}, \mathbf{y}) = d'(P(\mathbf{x}), P(\mathbf{y})). \tag{3.18}$$

Recall that $l_0 \setminus \{\mathbf{0}\}$ equipped with the metric d' is a complete separable metric space. Since P is invertible and componentwise continuous, one can easily show that $l_\infty \setminus \{(\infty)_j\}$ equipped with the metric d is a complete separable metric space.

Note that we can embed $\cup_{n\in\mathbb{N}} \dot{\mathcal{V}}^n$ ($\cup_{n\in\mathbb{N}} \mathcal{V}^n$) into l_∞ (l_0) simply by adding a sequence of ∞ (0) before and after the n entrances of any element of $\dot{\mathcal{V}}^n$ ($\mathcal{V}^n$). For example, $\mathbb{X}_n^{i,j}$ can be seen as an element of l_∞ by identifying it with

$$\left(\ldots, \infty, \infty, u_n^{-1}(\|\mathbf{X}_i\|) \frac{\mathbf{X}_i}{\|X_i\|}, \ldots, u_n^{-1}(\|\mathbf{X}_{j-1}\|) \frac{\mathbf{X}_{j-1}}{\|\mathbf{X}_{j-1}\|}, \infty, \infty, \ldots \right). \tag{3.19}$$

We define the quotient spaces $\tilde{l}_\infty = l_\infty/\sim$ and $\tilde{l}_0 = l_0/\sim$, where $\sim$ is the equivalence relation defined on both l_∞ and l_0 by $\mathbf{x} \sim \mathbf{y}$ if and only if there exists $k \in \mathbb{Z}$ such that $\sigma^k(\mathbf{x}) = \mathbf{y}$, where σ is the shift operator defined in (3.1). Also let $\tilde{\pi}$ denote the natural projection from l_∞ (l_0) to $\tilde{l}_\infty$ ($\tilde{l}_0$), which assigns to each element $\mathbf{x}$ of l_∞ (l_0) the corresponding equivalence class $\tilde{\pi}(\mathbf{x}) = \tilde{\mathbf{x}}$ in $\tilde{l}_\infty$ ($\tilde{l}_0$).

Observe that, since P is invertible, we may define $\tilde{P} : \tilde{l}_\infty \to \tilde{l}_0$ so that $\tilde{P}(\tilde{\pi}(\mathbf{x})) = \tilde{\pi}(P(\mathbf{x}))$.

Consider the metric $\tilde{d}'$ in $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ given by

$$\tilde{d}'(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \inf\{d'(\mathbf{x}', \mathbf{y}') \colon\ \mathbf{x}' \in \pi^{-1}(\tilde{\mathbf{x}}),\ \mathbf{y}' \in \pi^{-1}(\tilde{\mathbf{y}})\} = \inf\{d'(\sigma^k(\mathbf{x}), \sigma^m(\mathbf{y})) \colon\ k, m \in \mathbb{Z}\}. \tag{3.20}$$

This metric makes $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ a complete separable metric space. (See Lemma 2.1 and Lemma 6.1 of [BPS18]). Accordingly, on $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$, where $\tilde{\infty} = (\infty)_j$, we define the metric

$$\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \tilde{d}'(\tilde{P}(\tilde{\mathbf{x}}), \tilde{P}(\tilde{\mathbf{y}})),$$

which also gives a complete separable metric space.

Remark 3.5. The choice of the metric implies that a set $A \subset \tilde{l}_\infty \setminus \{\tilde{\infty}\}$ is bounded if and only if there exists $\varepsilon > 0$, such that for all $\tilde{\mathbf{x}} \in A$ we have $\|\tilde{P}(\tilde{\mathbf{x}})\|_\infty > \varepsilon$ or, equivalently, that $\inf_{j\in\mathbb{Z}} \|x_j\| < 1/\varepsilon$.

For $A \in \mathscr{F}$, as in (3.6), we define

$$\tilde{A} = \{\tilde{\mathbf{x}} \in \tilde{l}_\infty \colon\ \tilde{\pi}^{-1}(\tilde{\mathbf{x}}) \cap A \neq \emptyset\}; \qquad \tilde{\mathscr{J}} = \{\tilde{A} \colon\ A \in \mathscr{F}\}. \tag{3.21}$$

Using that $\mathscr{F}$ is a field, one can show that the class of subsets $\tilde{\mathscr{J}}$ is closed for unions. Indeed, this follows easily by observing that

$$\tilde{A} \cup \tilde{B} = \{\tilde{\mathbf{x}} \in \tilde{l}_\infty \colon\ \tilde{\pi}^{-1}(\tilde{\mathbf{x}}) \cap A \neq \emptyset \text{ or } \tilde{\pi}^{-1}(\tilde{\mathbf{x}}) \cap B \neq \emptyset\} = \{\tilde{\mathbf{x}} \in \tilde{l}_\infty \colon\ \tilde{\pi}^{-1}(\tilde{\mathbf{x}}) \cap (A \cup B) \neq \emptyset\}.$$

Let $\tilde{\mathscr{R}}$ denote the class of subsets of $\tilde{l}_\infty, \tilde{l}_0$ corresponding to the ring generated by $\tilde{\mathscr{J}}$, which is actually a field because by definition of $\mathscr{F}$, we have that $l_\infty, l_0 \subset \mathscr{F}$.

For $A \in \mathscr{F}$, let

$$\mathbb{A} = \{\mathbf{x} \in l_\infty \colon \tilde{\pi}(\mathbf{x}) \in \tilde{A}\} = \left\{\mathbf{x} \in l_\infty, l_0 \colon \mathbf{x} \in \bigcup_{j\in\mathbb{Z}} \sigma^{-j}(A)\right\}, \tag{3.22}$$

$$\mathscr{J} = \{\mathbb{A} \colon\ A \in \mathscr{F}\} \text{ and } \mathscr{R} \text{ be the ring generated by } \mathscr{J}.$$

If we start with some $A_\ell \in \mathcal{F}$ here, we correspondingly write $\mathbb{A}_\ell$. Note that $\sigma^{-1}(\mathbb{A}) = \mathbb{A}$, which means that both $\mathscr{J}$ and $\mathscr{R}$ are σ-invariant classes of subsets of l_∞, l_0. Also observe that $\mathscr{R} = \tilde{\pi}^{-1}(\tilde{\mathscr{R}})$ and $\mathscr{J} = \tilde{\pi}^{-1}(\tilde{\mathscr{J}})$, which is also closed for unions.

3.3.4. *The piling process.* We can now define the piling process, which presupposes the existence of a process $(Y_j)_{j\in\mathbb{Z}} \in l_\infty$ satisfying the following assumptions:

(1) $\mathcal{L}\left(\frac{1}{\tau}\mathbb{X}_n^{r_n+s,r_n+t} \,\middle|\, \|\mathbf{X}_{r_n}\| > u_n(\tau)\right) \xrightarrow[n\to\infty]{} \mathcal{L}\left((Y_j)_{j=s,\dots,t}\right)$, for all $s < t \in \mathbb{Z}$ and all $\tau > 0$;
(2) the process $(\Theta_j)_{j\in\mathbb{Z}}$ given by $\Theta_j = \frac{Y_j}{\|Y_0\|}$ is independent of $\|Y_0\|$;
(3) $\lim_{|j|\to\infty} \|Y_j\| = \infty$ a.s.;
(4) $\mathbb{P}\left(\inf_{j\le -1} \|Y_j\| \ge 1\right) > 0$.

Remark 3.6. In the setting of a jointly regularly varying sequence $(\mathbf{X}_j)_{j\in\mathbb{Z}}$, the process $(Y_j)_{j\in\mathbb{Z}}$ defined here (which lives in l_∞) is a transformed version of the tail process introduced in [BS09], which takes values in l_0 and was used later in [BKS12, BPS18], for example. Lemma 3.9 is one reason we choose to base our piling process in l_∞. Observe that joint regular variation of $(\mathbf{X}_j)_{j\in\mathbb{Z}}$ implies the existence of the sequence $(Y_j)_{j\in\mathbb{Z}}$ assumed in (1) (see [BS09, Theorem 2.1]) and also the independence of the respective polar decomposition assumed in (2) (see [BS09, Theorem 3.1]).

Remark 3.7. If, instead of $Д'_{q_n}$, we assume the much stronger conditions considered in [BPS18, Assumption 1.1] and other previous works, going back to [Smi92], one can show that both (3) and (4) hold. See [BS09, Proposition 4.2].

Definition 3.8. Assuming the existence of a sequence $(Y_j)_{j\in\mathbb{Z}}$ satisfying conditions (1)–(4), we define the *piling process* $(Z_j)_{j\in\mathbb{Z}}$ as a sequence of random vectors satisfying

$$\mathcal{L}\left((Z_j)_{j\in\mathbb{Z}}\right) = \mathcal{L}\left((Y_j)_{j\in\mathbb{Z}} \,\middle|\, \inf_{j\le -1} \|Y_j\| \ge 1\right).$$

We consider a polar decomposition of the piling process by defining the random variable L_Z and the process $(Q_j)_{j\in\mathbb{Z}}$ by

$$L_Z = \inf_{j\in\mathbb{Z}} \|Z_j\| \qquad Q_j = \frac{Z_j}{L_Z}. \tag{3.23}$$

We carry this polar decomposition to $\tilde{l}_\infty$ by letting $\mathbb{S} = \{\tilde{\mathbf{x}} \in \tilde{l}_\infty \colon \|\tilde{P}(\tilde{\mathbf{x}})\|_\infty = 1\}$ and defining the map

$$\begin{aligned} \psi \colon \tilde{l}_\infty \setminus \{\tilde{\infty}\} &\longrightarrow \mathbb{R}^+ \times \mathbb{S} \\ \tilde{\mathbf{x}} &\mapsto \left(\frac{1}{\|\tilde{P}(\tilde{\mathbf{x}})\|_\infty}, \frac{\tilde{x}}{(\|\tilde{P}(\tilde{\mathbf{x}})\|_\infty)^{-1}} \right). \end{aligned} \tag{3.24}$$

We define $\tilde{\mathbf{Q}} := \tilde{\pi}((Q_j)_{j\in\mathbb{Z}})$ and observe that $\psi(\tilde{\pi}((Z_j)_{j\in\mathbb{Z}})) = (L_Z, \tilde{\mathbf{Q}})$.

In order to understand how the piling process contains the details of the cluster structure, we start by discussing some properties of the objects just introduced.

Observe that by (3.14), $\left\|\frac{1}{\tau}\mathbb{X}_n^{r_n+j}\right\| < 1 \Leftrightarrow \|\mathbf{X}_{r_n+j}\| > u_n(\tau)$, which, in particular, implies that $\|Y_0\| < 1$ a.s. Hence, an exceedance of $u_n(\tau)$ at time $r_n + j$, given that such an exceedance occurred at time r_n, is associated to the event $\|Y_j\| < 1$. Therefore, loosely speaking, the sequence $(Y_j)_{j\in\mathbb{Z}}$ keeps the information about what happens around the time of an abnormal observation corresponding to an exceedance of an high threshold. Note that, in this case, we just know that an exceedance detected by $\|Y_0\| < 1$ has occurred but we do not know where the corresponding cluster was initiated. On the other hand, the piling process $(Z_j)_{j\in\mathbb{Z}}$ keeps the information around the time of an abnormal observation, given that this exceedance, associated with $\|Y_0\| < 1$, actually *initiated* a cluster because $\inf_{j\le -1} \|Y_j\| \ge 1$ has occurred and therefore we are certain that such exceedance does not belong to a cluster initiated earlier.

Note that due to assumption (3) both $(Y_j)_{j\in\mathbb{Z}}$ and the piling process $(Z_j)_{j\in\mathbb{Z}}$ take values in l_∞.

3.3.5. *Properties of the piling process.* As noted in Remark 3.6 the following lemma is a nice consequence for $\|Y_0\|$ of our piling process being based in l_∞.

Lemma 3.9. *The random variable $\|Y_0\|$ is uniformly distributed.*

Proof. Using stationarity (2.14) and (2.16), it follows that for all $v \in [0,1]$,

$$\begin{aligned} \mathbb{P}(\|Y_0\| < v) &= \lim_{n\to\infty} \mathbb{P}\left(\left\| \frac{u_n^{-1}(\|\mathbf{X}_{r_n}\|)}{\tau} \frac{\mathbf{X}_{r_n}}{\|\mathbf{X}_{r_n}\|} \right\| < v \;\middle|\; \|\mathbf{X}_{r_n}\| > u_n(\tau) \right) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\|\mathbf{X}_{r_n}\| > u_n(\tau v) \mid \|\mathbf{X}_{r_n}\| > u_n(\tau) \right) = \lim_{n\to\infty} \frac{n\mathbb{P}\left(\|\mathbf{X}_0\| > u_n(\tau v)\right)}{n\mathbb{P}(\|\mathbf{X}_0\| > u_n(\tau))} = \frac{\tau v}{\tau} = v. \end{aligned}$$

□

Next we show a relation that will provide a connection between the piling process and the outer measure used in [FFMa20]: this describes clustering by splitting the events into annuli of different cluster lengths.

Proposition 3.10. *Let $A_1 \in \mathscr{F}$. Under Д$'_{q_n}$,*

$$\lim_{n\to\infty} \left| k_n \mathbb{P}(\mathbb{X}_{n,1} \in \mathbb{A}_1) - n\mathbb{P}(A_{n,1}^{(q_n)}) \right| = \lim_{n\to\infty} \left| k_n \mathbb{P}\left(\mathscr{W}_{r_n}^c(A_{n,1})\right) - n\mathbb{P}(A_{n,1}^{(q_n)}) \right| = 0,$$

where $A_{n,1}$ is defined in (3.9) *and $\mathbb{A}_1$ is from* (3.22) *applied to A_1.*

Proof. We start by estimating $\mathbb{P}(\mathbb{X}_{n,1} \in \mathbb{A}_1) = \mathbb{P}(\mathscr{W}_{r_n}^c(A_{n,1}))$, which we do by decomposing $\mathscr{W}_{r_n}^c(A_{n,1})$ according to the last occurrence of event $A_{n,1}$. Namely,

$$\mathbb{P}(\mathscr{W}_{r_n}^c(A_{n,1})) = \sum_{j=0}^{r_n-1} \mathbb{P}(\sigma^{-j}(A_{n,1})) - \sum_{j=0}^{r_n-1} \mathbb{P}(\sigma^{-j}(A_{n,1}), \mathscr{W}_{j+1,r_n}^c(A_{n,1}))).$$

It follows that

$$\left| \mathbb{P}(\mathscr{W}_{r_n}^c(A_{n,1})) - \left(\sum_{j=0}^{r_n-q_n-2} \mathbb{P}(\sigma^{-j}(A_{n,1})) - \sum_{j=0}^{r_n-q_n-2} \mathbb{P}(\sigma^{-j}(A_{n,1}), \mathscr{W}_{j+1,r_n}^c(A_{n,1}))) \right) \right|$$
$$\leq \sum_{j=r_n-q_n-1}^{r_n-1} \mathbb{P}(\sigma^{-j}(A_{n,1}), \mathscr{W}_{j+1,r_n}(A_{n,1}))) \leq (q_n+1)\mathbb{P}(A_{n,1}) \tag{3.25}$$

Using stationarity,

$$\sum_{j=0}^{r_n-q_n-2} \mathbb{P}(\sigma^{-j}(A_{n,1}), \mathscr{W}_{j+1,r_n}^c(A_{n,1})) = \sum_{s=q_n+2}^{r_n} \mathbb{P}(A_{n,1}, \mathscr{W}_{1,s}^c(A_{n,1})).$$

For $s \geq q_n + 2$,

$$\mathbb{P}(A_{n,1}, \mathscr{W}_{1,s}^c(A_{n,1})) - \mathbb{P}(A_{n,1}, \mathscr{W}_{1,q_n+1}^c(A_{n,1})) = \mathbb{P}(A_{n,1}^{(q_n)}, \mathscr{W}_{q_n+1,s}^c(A_{n,1}))$$

and therefore

$$\left| \sum_{s=q_n+2}^{r_n} \mathbb{P}(A_{n,1}, \mathscr{W}_{1,s}^c(A_{n,1})) - \sum_{s=q_n+2}^{r_n} \mathbb{P}(A_{n,1}, \mathscr{W}_{1,q_n+1}^c(A_{n,1})) \right| \leq \sum_{s=q_n+2}^{r_n} \mathbb{P}(A_{n,1}^{(q_n)}, \mathscr{W}_{q_n+1,s}^c(A_{n,1}))$$
$$\leq r_n \mathbb{P}(A_{n,1}^{(q_n)}, \mathscr{W}_{q_n+1,r_n}^c(A_{n,1})). \tag{3.26}$$

Combining (3.25) and (3.26) and using stationarity, we obtain

$$\left| \mathbb{P}(\mathscr{W}_{r_n}^c(A_{n,1})) - \left(\sum_{s=q_n+2}^{r_n} \mathbb{P}(A_{n,1}) - \sum_{s=q_n+2}^{r_n} \mathbb{P}(A_{n,1}, \mathscr{W}_{1,q_n+1}^c(A_{n,1})) \right) \right|$$
$$\leq (q_n+1)\mathbb{P}(A_{n,1}) + r_n \mathbb{P}(A_{n,1}^{(q_n)}, \mathscr{W}_{q_n+1,r_n}^c(A_{n,1}))$$

Noting that the term between big brackets is equal to $(r_n - q_n + 1)\mathbb{P}(A_{n,1}, \mathscr{W}_{1,q_n+1}(A_{n,1})) = (r_n - q_n + 1)\mathbb{P}(A_{n,1}^{(q_n)})$ then

$$\left| \mathbb{P}(\mathscr{W}_{r_n}^c(A_{n,1})) - r_n \mathbb{P}(A_{n,1}^{(q_n)}) \right| \leq 2q_n \mathbb{P}(A_{n,1}) + r_n \mathbb{P}(A_{n,1}^{(q_n)}, \mathscr{W}_{q_n+1,r_n}^c(A_{n,1})).$$

Multiplying by k_n we obtain

$$\left| k_n \mathbb{P}(\mathbb{X}_{n,1} \in \mathbb{A}_1) - n\mathbb{P}(A_{n,1}^{(q_n)}) \right| \leq 2q_n k_n \mathbb{P}(A_{n,1}^{(q_n)}) + n\mathbb{P}(A_{n,1}^{(q_n)}, \mathscr{W}_{q_n+1,r_n}^c(A_{n,1})).$$

The second term on the right vanishes by $Д'_{q_n}$. Since by definition of $A_{n,1}$, we have that $A_{n,1} \subset \{\|X_0\| > u_n(h_0)\}$, where $h_0 = \inf\{\|x\| \colon x \in H_0\} \geq 0$, then $n\mathbb{P}(A_{n,1}) \leq n\mathbb{P}(\|X_0\| > u_n(h_0)) \xrightarrow[n\to\infty]{} h_0$. Recalling that $q_n = o(r_n)$, it follows that the first term on right also vanishes. □

Corollary 3.11. *Under* $Д'_{q_n}$,

$$\lim_{n\to\infty} k_n \mathbb{P}(\mathscr{W}^c_{r_n}(U_n(\tau)) = \theta\tau \qquad \textit{and} \qquad \lim_{n\to\infty} \frac{\mathbb{P}(\mathscr{W}^c_{r_n}(U_n(\tau))}{r_n \mathbb{P}(U_n(\tau))} = \theta.$$

Proof. By Proposition 3.10, (2.14) and (3.15),

$$\lim_{n\to\infty} k_n \mathbb{P}(\mathscr{W}^c_{r_n}(U_n(\tau)) = \lim_{n\to\infty} n\mathbb{P}(U_n^{q_n}(\tau)) = \lim_{n\to\infty} n\mathbb{P}(U_n(\tau)) \frac{\mathbb{P}(U_n^{q_n}(\tau))}{\mathbb{P}(U_n(\tau))} = \tau\theta.$$

It follows that

$$\lim_{n\to\infty} \frac{\mathbb{P}(\mathscr{W}^c_{r_n}(U_n(\tau))}{r_n \mathbb{P}(U_n(\tau))} = \frac{k_n \mathbb{P}(\mathscr{W}^c_{r_n}(U_n(\tau))}{n\mathbb{P}(U_n(\tau))} = \frac{\tau\theta}{\tau} = \theta.$$

□

Lemma 3.12. *Recalling the definition of the EI given in* (3.15),

$$\theta = \mathbb{P}\left(\inf_{j\geq 1} \|Y_j\| \geq 1\right) = \mathbb{P}\left(\inf_{j\leq -1} \|Y_j\| \geq 1\right).$$

Proof. By (3.15), stationarity and definition of the process $(Y_j)_{j\in\mathbb{Z}}$, given in assumption (1), we may write

$$\begin{aligned}\theta &= \lim_{n\to\infty} \mathbb{P}\left(U_n^{(q_n)}(\tau) \,\middle|\, U_n(\tau)\right) = \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\tau}\mathbb{X}_n^{r_n, r_n+q_n} \in \mathscr{W}_{r_n+1, r_n+q_n+1}(U(1)) \,\middle|\, \|\mathbf{X}_{r_n}\| > u_n(\tau)\right) \\ &= \mathbb{P}\left(\inf_{j\geq 1} \|Y_j\| \geq 1\right).\end{aligned}$$

By stationarity,

$$\begin{aligned}&\mathbb{P}\left(\mathscr{W}_{r_n+1, r_n+q_n+1}(U_n(\tau)) \cap \sigma^{-r_n}(U_n(\tau)\right) = \mathbb{P}\left(\mathscr{W}_{r_n+1, r_n+q_n+1}(U_n(\tau))\right) - \mathbb{P}\left(\mathscr{W}_{r_n, r_n+q_n+1}(U_n(\tau))\right) \\ &\quad = \mathbb{P}\left(\mathscr{W}_{r_n-q_n, r_n}(U_n(\tau))\right) - \mathbb{P}\left(\mathscr{W}_{r_n-q_n, r_n+1}(U_n(\tau))\right) = \mathbb{P}\left(\mathscr{W}_{r_n-q_n, r_n}(U_n(\tau)) \cap \sigma^{-r_n}((U_n(\tau))\right) \\ &\quad = \mathbb{P}\left(\mathbb{X}_n^{r_n-q_n, r_n+1} \in \mathscr{W}_{r_n-q_n, r_n}(U(\tau)) \cap \sigma^{-r_n}((U(\tau))\right).\end{aligned}$$

The second equality in the statement of the lemma follows now easily. □

The next result is instrumental because it shows how the piling process can be used to encode the information regarding the clustering. Essentially, it says that the joint distribution of the random variables in a block where an exceedance is observed (which makes it a cluster) is given by the piling process. Recall that by (3.19), $\mathbb{X}_{n,i}$ can be thought of as lying in l_∞.

Proposition 3.13. *Under the assumptions used to define the piling process and condition* $Д'_{q_n}$, *for every* $\tau > 0$,

$$\mathcal{L}\left(\tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right) \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}(U(\tau))\right) \longrightarrow \mathcal{L}\left(\tilde{\pi}\left((Z_j)_{j\in\mathbb{Z}}\right)\right).$$

Proof. In what follows we write U^τ for $U(\tau)$. As in Appendix B, $\tilde{\mathscr{J}}$ is a convergence determining class and therefore we need to show that for all $A \in \mathscr{F}$ and corresponding $\tilde{A} \in \tilde{\mathscr{J}}$, such that $\mathbb{P}\left(\tilde{\pi}\left((Z_j)_{j\in\mathbb{Z}}\right) \in \partial\tilde{A}\right) = 0$, we have

$$\mathbb{P}\left(\tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right) \in \tilde{A} \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U^\tau\right)\right) \longrightarrow \mathbb{P}\left(\tilde{\pi}\left((Z_j)_{j\in\mathbb{Z}}\right) \in \tilde{A}\right),$$

which will follow if we show that

$$\mathbb{P}\left(\frac{1}{\tau}\mathbb{X}_{n,1} \in \mathbb{A} \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U^\tau\right)\right) \longrightarrow \mathbb{P}\left((Y_j)_{j\in\mathbb{Z}} \in \mathbb{A} \,\middle|\, \inf_{j\le -1}\|Y_j\| \ge 1\right),$$

for all $\mathbb{A} \in \mathscr{J}$, such that $\mathbb{P}((Y_j)_{j\in\mathbb{Z}} \in \partial\mathbb{A}) = 0$. Recall that $\mathbb{A}$ is σ-invariant, *i.e.*, $\sigma^{-1}(\mathbb{A}) = \mathbb{A}$. We start estimating $\mathbb{P}(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}^c_{r_n}\left(U^\tau\right))$ by decomposing the event on the right (which essentially says that at least one exceedance of $u_n(\tau)$ has occurred up to time $r_n - 1$) with respect to the first time, $i = 0, \dots, r_n - 1$, when that exceedance occurs, *i.e.*, $\{\mathbb{X}_{n,1} \in \sigma^{-i}(U^\tau\} = \{\|\mathbf{X}_i\| > u_n(\tau)\}$:

$$\mathbb{P}\left(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}^c_{r_n}\left(U^\tau\right)\right) = \sum_{i=0}^{r_n-1} \mathbb{P}\left(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}_i\left(U^\tau\right) \cap \sigma^{-i}(U^\tau)\right) =: \sum_{i=0}^{r_n-1} B_{n,i}$$

Since $B_{n,i} \le \mathbb{P}(\|\mathbf{X}_i\| > u_n(\tau))$,

$$\left|\mathbb{P}\left(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}^c_{r_n}\left(U^\tau\right)\right) - \sum_{i=q_n}^{r_n-1} B_{n,i}\right| \le q_n\mathbb{P}(\|\mathbf{X}_i\| > u_n(\tau)) = q_n\mathbb{P}(U_n(\tau)) =: I(n) \quad (3.27)$$

For $i \ge q_n$, we now estimate $B_{n,i}$ by $D_{n,i} := \mathbb{P}\left(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}_{i-q_n,i}\left(U^\tau\right) \cap \sigma^{-i}(U^\tau)\right)$. Namely,

$$\begin{aligned}
|B_{n,i} - D_{n,i}| &= \mathbb{P}\left(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}^c_{0,i-q_n}\left(U^\tau\right) \cap \mathscr{W}_{i-q_n,i}\left(U^\tau\right) \cap \sigma^{-i}(U^\tau)\right) \\
&\le \mathbb{P}\left(\mathbb{X}_{n,1} \in \mathscr{W}^c_{0,i-q_n}\left(U^\tau\right) \cap \mathscr{W}_{i-q_n,i}\left(U^\tau\right) \cap \sigma^{-i}(U^\tau)\right) \\
&= \sum_{t=q_n+1}^{i} \mathbb{P}\left(\mathbb{X}_{n,1} \in U^\tau \cap \mathscr{W}_{1,t}\left(U^\tau\right) \cap \sigma^{-t}(U^\tau)\right) \\
&= \mathbb{P}\left(\mathbb{X}_{n,1} \in U^\tau \cap \mathscr{W}_{1,q_n+1}\left(U^\tau\right) \cap \mathscr{W}^c_{q_n+1,i+1}\left(U^\tau\right)\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left|\sum_{i=q_n}^{r_n-1} B_{n,i} - \sum_{i=q_n}^{r_n-1} D_{n,i}\right| &\le \sum_{i=q_n}^{r_n-1} \mathbb{P}\left(\mathbb{X}_{n,1} \in U^\tau \cap \mathscr{W}_{1,q_n+1}\left(U^\tau\right) \cap \mathscr{W}^c_{q_n+1,i+1}\left(U^\tau\right)\right) \\
&\le (r_n - q_n)\mathbb{P}\left(U_n^{(q_n)}(\tau) \cap \mathscr{W}^c_{q_n+1,r_n}\left(U_n(\tau)\right)\right) =: II(n) \qquad (3.28)
\end{aligned}$$

By stationarity and because $\mathbb{A}$ is σ-invariant, for all $i = q_n, \dots, r_n - 1$,

$$\begin{aligned}
D_{n,i} &= \mathbb{P}\left(\mathbb{X}_{n,2} \in \tau\mathbb{A} \cap \mathscr{W}_{r_n+i-q_n,r_n+i}\left(U^\tau\right) \cap \sigma^{-(r_n+i)}(U^\tau)\right) \\
&= \mathbb{P}\left(\mathbb{X}_n^{r_n-q_n,2r_n} \in \tau\mathbb{A} \cap \mathscr{W}_{r_n-q_n,r_n}\left(U^\tau\right) \cap \sigma^{-r_n}(U^\tau)\right) =: D_n
\end{aligned}$$

Then using estimates (3.27) and (3.28), we obtain

$$\left|\mathbb{P}\left(\mathbb{X}_{n,1} \in \tau\mathbb{A} \cap \mathscr{W}^c_{r_n}\left(U^\tau\right)\right) - (r_n - q_n)\mathbb{P}\left(D_n\right)\right| \le I(n) + II(n)$$

and therefore

$$\left|\mathbb{P}\left(\mathbb{X}_{n,1}\in\tau\mathbb{A}\cap\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)-r_n\mathbb{P}\left(D_n\right)\right|\leq 2I(n)+I\!I(n).$$

Hence, we can write

$$\left|\mathbb{P}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\in\mathbb{A}\;\middle|\;\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)-\frac{r_n\mathbb{P}(U_n(\tau))}{\mathbb{P}\left(\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)}\mathbb{P}\left(D_n\mid\|\mathbf{X}_{r_n}\|>u_n(\tau)\right)\right|$$
$$\leq(2I(n)+I\!I(n))\frac{1}{\mathbb{P}\left(\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)}=:E(n).$$

Note that we can write

$$E(n)=(2I(n)+I\!I(n))\frac{r_n\mathbb{P}(U_n(\tau))}{\mathbb{P}\left(\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)}\frac{k_n}{k_nr_n\mathbb{P}(U_n(\tau))}.$$

Since, by Corollary 3.11, $\lim_{n\to\infty}\frac{r_n\mathbb{P}(U_n(\tau))}{\mathbb{P}\left(\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)}=\theta^{-1}$ and, by definition of $U_n(\tau)$ we have $\lim_{n\to\infty}k_nr_n\mathbb{P}(U_n(\tau))=\lim_{n\to\infty}n\mathbb{P}(U_n(\tau))=\tau$, it follows that

$$E(n)\leq C(k_nI(n)+k_nI\!I(n))$$

for some $C>0$. By definition of the sequences $(k_n)_n$ and $(q_n)_n$, we have that

$$\lim_{n\to\infty}k_nq_n\mathbb{P}(U_n(\tau))=0,$$

which means that $\lim_{n\to\infty}k_nI(n)=0$.

Observe also that $Д'_{q_n}$ implies that

$$\lim_{n\to\infty}k_nr_n\mathbb{P}\left(U_n^{(q_n)}(\tau)\cap\mathscr{W}_{q_n+1,r_n}^c\left(U_n(\tau)\right)\right)=0,$$

which means that $\lim_{n\to\infty}k_nI\!I(n)=0$ and therefore $\lim_{n\to\infty}E(n)=0$.

In order to get the result we need to check that

$$\lim_{n\to\infty}\frac{r_n\mathbb{P}(U_n(\tau))}{\mathbb{P}\left(\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)}\mathbb{P}\left(D_n\mid\|\mathbf{X}_{r_n}\|>u_n(\tau)\right)=\mathbb{P}\left((Y_j)_{j\in\mathbb{N}}\in\mathbb{A}\;\middle|\;\inf_{j\leq-1}\|Y_j\|\geq1\right).$$

Since by Corollary 3.11 and Lemma 3.12, we have $\lim_{n\to\infty}\frac{\mathbb{P}\left(\mathbb{X}_{n,1}\in\mathscr{W}_{r_n}^c\left(U^\tau\right)\right)}{r_n\mathbb{P}(U_n(\tau))}=\theta=\mathbb{P}(\inf_{j\leq-1}\|Y_j\|\geq 1)$, then we need to show that

$$\lim_{n\to\infty}\mathbb{P}\left(\mathbb{X}_n^{r_n-q_n,2r_n}\in\tau\mathbb{A}\cap\mathscr{W}_{r_n-q_n,r_n}\left(U^\tau\right)\;\middle|\;\|\mathbf{X}_{r_n}\|>u_n(\tau)\right)=\mathbb{P}\left((Y_j)_{j\in\mathbb{N}}\in\mathbb{A},\;\inf_{j\leq-1}\|Y_j\|\geq1\right),$$

which we can rewrite as

$$\lim_{n\to\infty}\mathbb{P}\left(\frac{1}{\tau}\mathbb{X}_n^{r_n-q_n,2r_n}\in\mathbb{A}\cap\mathscr{W}_{-\infty,r_n}\left(U^1\right)\;\middle|\;\|\mathbf{X}_{r_n}\|>u_n(\tau)\right)=\mathbb{P}\left((Y_j)_{j\in\mathbb{N}}\in\mathbb{A}\cap\mathscr{W}_{-\infty,0}\left(U^1\right)\right),\tag{3.29}$$

Note that $\{(Y_j)_{j\in\mathbb{N}}\in\partial\mathscr{W}_{-\infty,0}\left(U^1\right)\}\subset\cup_{j\leq1}\{\|Y_j\|=1\}$. Since, by assumption (2), for every $j\in\mathbb{Z}$, we have $Y_j=\|Y_0\|\Theta_j$, with Θ_j independent of $\|Y_0\|$ and since the latter is uniformly distributed by Lemma 3.9, then $\mathbb{P}(\|Y_j\|=1)=0$, for all $j\in\mathbb{Z}$. Therefore, $\mathbb{P}((Y_j)_{j\in\mathbb{N}}\in\partial\mathscr{W}_{-\infty,0}\left(U^1\right))=0$. Since $\mathbb{P}((Y_j)_{j\in\mathbb{N}}\in\partial\mathbb{A})=0$ and $\partial\left(\mathbb{A}\cap\mathscr{W}_{-\infty,0}\left(U^1\right)\right)\subset\partial\mathbb{A}\cup\partial\mathscr{W}_{-\infty,0}\left(U^1\right)$ it follows that $\mathbb{P}\left((Y_j)_{j\in\mathbb{N}}\in\partial\left(\mathbb{A}\cap\mathscr{W}_{-\infty,0}\left(U^1\right)\right)\right)=0$. Then (3.29) follows by definition of the sequence $(Y_j)_{j\in\mathbb{N}}$ given in assumption (1). □

Corollary 3.14. *Under the assumptions of Proposition* 3.13*, we have that* L_Z *and the process* $(Q_j)_{j\in\mathbb{Z}}$ *defined in* (3.23) *satisfy*

(1) L_Z *is uniformly distributed on* $[0,1]$*;*
(2) L_Z *and* $\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right)$ *are independent.*

Proof. For part (1), note that $L_Z \le \|Z_0\| \le \|Y_0\| \le 1$ a.s. Let $0 \le v \le 1$. By Proposition 3.13, (3.14) and Corollary 3.11, we may write

$$\begin{aligned}\mathbb{P}(L_Z < v) &= \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{\tau}\mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U(v)\right) \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U(\tau)\right)\right)\\ &= \lim_{n\to\infty} \mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau v)\right) \mid \mathscr{W}^c_{r_n}\left(U_n(\tau)\right)\right) = \lim_{n\to\infty} \frac{\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau v)\right) \cap \mathscr{W}^c_{r_n}\left(U_n(\tau)\right)\right)}{\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau)\right)\right)}\\ &= \lim_{n\to\infty} \frac{\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau v)\right)\right)}{\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau)\right)\right)} = \lim_{n\to\infty} \frac{k_n\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau v)\right)\right)}{k_n\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau)\right)\right)} = \frac{\tau v\theta}{\tau\theta} = v.\end{aligned}$$

To prove (2), we start by observing that $\frac{1}{\|P(\mathbf{x})\|_\infty} = \inf_{j\in\mathbb{Z}} \|x_j\|$ and that the map $\tilde{\mathbf{x}} \mapsto \left(\tilde{\mathbf{x}}, \frac{1}{\|\tilde{P}(\tilde{\mathbf{x}})\|_\infty}\right)$ is continuous on $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$. Then, by Proposition 3.13 and the CMT

$$\mathcal{L}\left(\tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right), \frac{1}{\|\tilde{P}\left(\tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right)\right)\|_\infty} \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U(\tau)\right)\right) \longrightarrow \mathcal{L}\left(\tilde{\pi}\left((Z_j)_{j\in\mathbb{Z}}\right), L_Z\right)$$

Since the map $(\tilde{\mathbf{x}}, a) \mapsto \left(\frac{\tilde{\mathbf{x}}}{b}, a\right)$ is continuous on $\tilde{l}_\infty \setminus \{\tilde{\infty}\} \times (0,\infty)$, then

$$\mathcal{L}\left(\left\|\tilde{P}\left(\tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right)\right)\right\|_\infty \tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right), \frac{1}{\|\tilde{P}\left(\tilde{\pi}\left(\frac{1}{\tau}\mathbb{X}_{n,1}\right)\right)\|_\infty} \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U(\tau)\right)\right) \longrightarrow \mathcal{L}\left(\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right), L_Z\right) \quad (3.30)$$

Since $\tilde{\mathscr{J}}$ is a convergence determining class (see Appendix B), the result will follow if we show that for all $A \in \mathscr{F}$ and corresponding $\tilde{A} \in \tilde{\mathscr{J}}$, such that $\mathbb{P}(\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right) \in \partial\tilde{A}) = 0$, and all $v \in [0,1]$, we have

$$\mathbb{P}\left(\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right) \in \tilde{A}, L_Z < v\right) = \mathbb{P}\left(\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right) \in \tilde{A}\right) \cdot \mathbb{P}(L_Z < v) \quad (3.31)$$

Letting $\mathbb{A} \in \mathscr{J}$ be such that $\mathbb{A} = \tilde{\pi}^{-1}(\tilde{A})$ and $m_{r_n} = \min\{\|\mathbb{X}^0_n\|, \dots, \|\mathbb{X}^{r_n-1}_n\|\}$, by (3.30), we can write that

$$\begin{aligned}\mathbb{P}\Big((Q_j)_{j\in\mathbb{Z}} \in \mathbb{A}, L_Z < v\Big) &= \lim_{n\to\infty} \mathbb{P}\left(\frac{1}{m_{r_n}}\mathbb{X}_{n,1} \in \mathbb{A} \,\middle|\, \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}\left(U(\tau v)\right)\right) \frac{k_n\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau v)\right)\right)}{k_n\mathbb{P}\left(\mathscr{W}^c_{r_n}\left(U_n(\tau)\right)\right)}\\ &= \mathbb{P}\left(\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right) \in \tilde{A}\right) \cdot v = \mathbb{P}\left(\tilde{\pi}\left((Q_j)_{j\in\mathbb{Z}}\right) \in \tilde{A}\right) \cdot \mathbb{P}(L_Z < v).\end{aligned}$$

□

The next result, an analogue of [BPS18, Lemma 3.3] in our setting, formally establishes the convergence of the intensity measures of the cluster point processes we introduce later. We refer to Appendix C for the definitions of weak$^\#$ convergence and boundedly finite measures.

Corollary 3.15. *Under the assumptions used to define the piling process and condition* $Д'_{q_n}$*, the sequence of boundedly finite measures* $\mu_n = k_n \mathbb{P}(\tilde{\pi}(\mathbb{X}_{n,1}) \in \cdot)$ *in* $\mathcal{M}^{\#}_{\tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ *converges in the* $w^{\#}$ *topology to* $\mu = \theta(Leb \times \mathbb{P}_{\tilde{\mathbf{Q}}}) \circ \psi$*, where* $\mathbb{P}_{\tilde{\mathbf{Q}}}$ *is the distribution of* $\tilde{\pi}((Q_j)_{j\in\mathbb{Z}})$*.*

Proof. By Lemma B.4, we only need to check the convergence for all bounded $\tilde{A} \in \tilde{\mathscr{J}}$ such that $\mu(\partial \tilde{A}) = 0$. By Remark 3.5, since $\tilde{A}$ is bounded, there exists $\tau > 0$ such that for all $\tilde{\mathbf{x}} \in \tilde{A}$, we have $\inf_{j\in\mathbb{Z}} \|x_j\| < \tau$. Hence, by 3.14, if $\tilde{\pi}(\mathbb{X}_{n,1}) \in \tilde{A}$, then $\|\mathbf{X}_j\| > u_n(\tau)$ for $j = 0, \dots, r_n - 1$ and hence $\mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}(U(\tau))$. Let, as before, $\mathbb{A} = \tilde{\pi}^{-1}(\tilde{A})$. Then

$$\begin{aligned}
k_n \mathbb{P}(\tilde{\pi}(\mathbb{X}_{n,1}) \in \tilde{A}) &= k_n \mathbb{P}(\mathbb{X}_{n,1} \in \mathbb{A}) = k_n \mathbb{P}\left(\mathbb{X}_{n,1} \in \mathbb{A} \cap \mathscr{W}^c_{r_n}(U(\tau))\right) \\
&= \frac{n}{r_n} \frac{\mathbb{P}(\|\mathbf{X}_0\| > u_n(\tau))}{\mathbb{P}(\|\mathbf{X}_0\| > u_n(\tau))} \mathbb{P}\left(\mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}(U(\tau))\right) \mathbb{P}\left(\mathbb{X}_{n,1} \in \mathbb{A} \;\middle|\; \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}(U(\tau))\right) \\
&= n\mathbb{P}(\|\mathbf{X}_0\| > u_n(\tau)) \frac{\mathbb{P}\left(\mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}(U(\tau))\right)}{r_n \mathbb{P}(\|\mathbf{X}_0\| > u_n(\tau))} \mathbb{P}\left(\frac{\mathbb{X}_{n,1}}{\tau} \in \tau^{-1}\mathbb{A} \;\middle|\; \mathbb{X}_{n,1} \in \mathscr{W}^c_{r_n}(U(\tau))\right)
\end{aligned}$$

By (2.14) and Corollary 3.11, we have that the first term converges to τ and the second to θ, as $n \to \infty$. By Proposition 3.10, the third term goes to $\mathbb{P}((Z_j)_{j\in\mathbb{Z}} \in \tau^{-1}\mathbb{A})$. We now use Corollary 3.14 in order to finish the proof.

$$\begin{aligned}
\mathbb{P}((Z_j)_{j\in\mathbb{Z}} \in \tau^{-1}\mathbb{A}) = \mathbb{P}(\tau(Z_j)_{j\in\mathbb{Z}} \in \mathbb{A}) &= \int_0^1 \mathbb{P}\left(\tau(Z_j)_{j\in\mathbb{Z}} \in \mathbb{A} \mid L_Z = v\right) dv \\
&= \int_0^1 \mathbb{P}\left(\tau v(Q_j)_{j\in\mathbb{Z}} \in \mathbb{A} \mid L_Z = v\right) dv = \frac{1}{\tau}\int_0^\tau \mathbb{P}\left(s(Q_j)_{j\in\mathbb{Z}} \in \mathbb{A}\right) ds = \frac{\mu(\tilde{A})}{\tau\theta}.
\end{aligned}$$

□

3.4. **Complete convergence of point processes.** We define and prove the weak convergence of the point processes that keep all the cluster information. We refer to Appendix C for the precise definition of point processes and their weak convergence. Essentially, we consider a random element on the space $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ of boundedly finite point measures on $\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}$. Namely, similarly to [BPS18], we define the point point processes of clusters by

$$N_n = \sum_{i=1}^{\infty} \delta_{(i/k_n, \tilde{\pi}(\mathbb{X}_{n,i}))}. \tag{3.32}$$

Now, we define the point process that will appear as the limit of the cluster point process. Let $(T_i)_{i\in\mathbb{N}}$ and $(U_i)_{i\in\mathbb{N}}$ be such that $\sum_{i=1}^{\infty} \delta_{(T_i, U_i)}$ is a bidimensional Poisson point process on $\mathbb{R}_0^+ \times \mathbb{R}_0^+$ with intensity measure θLeb $\times$ Leb. Also let $(\tilde{\mathbf{Q}}_i)_{i\in\mathbb{N}}$ be an i.i.d. sequence of random elements in $\mathbb{S}$ such that each $\tilde{\mathbf{Q}}_i$ has a distribution given by (3.23). We assume that the sequences $(T_i)_{i\in\mathbb{N}}$, $(U_i)_{i\in\mathbb{N}}$ and $(\tilde{\mathbf{Q}}_i)_{i\in\mathbb{N}}$ are mutually independent. We define

$$N = \sum_{i=1}^{\infty} \delta_{(T_i, U_i \tilde{\mathbf{Q}}_i)}. \tag{3.33}$$

Note that N just defined is a Poisson point process on $\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}$ with intensity Leb $\times\, \mu$, where μ is as in Corollary 3.15.

We are now ready to state a general complete convergence result.

Theorem 3.16. *Let $\mathbf{X}_0, \mathbf{X}_1, \dots$ be a stationary process of random vectors in $\mathbb{R}^d$ with have proper tails, in the sense given in Section* 2.4.1. *Assume that the piling process given in Definition* 3.8 *is well defined and conditions $Д_{q_n}$ and $Д'_{q_n}$ hold. Then point process N_n, given in* (3.32) *converges weakly in $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ to the Poisson point process N given by* (3.33).

In order to prove the theorem, we need essentially to show two things. The first is a sort of *independent increments* property together with convergence in distribution of joint random variables using the avoidance function, *i.e.*, by computing the probability of having no extremal occurrences (see the definition of the avoidance function in Appendix C). This is done in Proposition 3.17, whose complete proof is rather lengthy because we are using the very weak mixing assumption $Д_{q_n}$, though a lot of the work for that has been done in previous papers by the authors. Then, in the second step, we need to show that the intensity measures of the processes converge to the right intensity measure. This has actually already been done in Corollary 3.15. Finally we need to join the pieces using the theory of weak$^{\#}$ convergence on non locally compact spaces developed in [DVJ03, DVJ08]. In fact, we needed to redo one of the results to correct a typo and improve it in order to be able to use the convergence of the intensity measures (see Appendix C).

In [FFMa20] the existence of a σ-finite outer measure ν on $\mathcal{V}^{\mathbb{N}_0,\mathbb{Z}}$ was assumed, so that the following limit exists

$$\lim_{n\to\infty} n\mathbb{P}(A_{n,\ell}^{(q_n)}) = \nu(A_\ell), \tag{3.34}$$

for all $A_\ell \in \mathscr{F}$ and $A_{n,\ell}^{(q_n)}$ given by equation (3.9). This outer measure described the piling of points on the multidimensional point processes created by clustering in [FFMa20]. Note that if we associate $\tilde{A}_\ell$, $\mathbb{A}_\ell$ to $A_\ell \in \mathscr{F}$ as in (3.21) and (3.22), respectively, then using Proposition 3.10 and Corollary 3.15, it follows that when we have the existence of a piling process and condition $Д'_{q_n}$ then (3.34) holds and

$$\nu(A_\ell) = \mu(\tilde{A}_\ell). \tag{3.35}$$

We state now the main result that provides independence of disjoint time pieces and convergence of joint distributions by use of the avoidance function.

Proposition 3.17. *Let $m \in \mathbb{N}$ and for each $\ell = 1, \dots, m$ let J_ℓ, A_ℓ be given, as in* (3.7). *For $n \in \mathbb{N}$, consider the respective versions $J_{n,\ell}$, $A_{n,\ell}$ given in* (3.8) *and* (3.9). *Assume that $Д_{q_n}$ and $Д'_{q_n}$ hold. Also assume that there exists a σ-finite outer measure ν such that* (3.34) *holds. Then*

$$\lim_{n\to\infty} \mathbb{P}\left(\bigcap_{\ell=1}^{m} \mathscr{W}_{J_{n,\ell}}(A_{n,\ell})\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{\ell=1}^{m} \mathscr{W}_{J_{n.\ell}}\left(A_{n,\ell}^{(q_n)}\right)\right) = \prod_{\ell=1}^{m} \mathrm{e}^{-\nu(A_\ell)|J_\ell|}.$$

The following proposition gives the first equality in Proposition 3.17 from the definition of $A_{n,\ell}$ and since $q_n = o(n)$. The idea is that the non-occurrence of the asymptotically rare event $A_{n,\ell}$ can be replaced by the non-occurrence of the event $A_{n,\ell}^{(q_n)}$, up to an asymptotically negligible error. This goes back to [FFT12, Proposition 1] and was further developed in [FFT15, Proposition 2.7]. We refer to [FFMa20, Proposition 3.2] for a proof.

Proposition 3.18.

$$\left|\mathbb{P}\left(\bigcap_{\ell=1}^{m}\mathscr{W}_{J_{n,\ell}}\left(A_{n,\ell}^{(q_n)}\right)\right)-\mathbb{P}\left(\bigcap_{\ell=1}^{m}\mathscr{W}_{J_{n,\ell}}\left(A_{n,\ell}\right)\right)\right|\leq q_n\sum_{\ell=1}^{m}\mathbb{P}(A_{n,\ell}).$$

The second equality in Proposition 3.17 follows essentially the same argument used to prove [FFMa20, Theorem 3.3], with the natural adjustments to compensate for the fact that the version of condition Д′ used here is weaker. For completeness, we include it in Appendix D.

We are now ready to prove the weak convergence of the cluster point processes.

Proof of Theorem 3.16. By Proposition C.2 and Lemma B.2, we first need to check that for all bounded sets $B\in\mathcal{I}$, defined in (C.6) we have $\lim_{n\to\infty}\mathbb{P}(N_n(B)=0)=\mathbb{P}(N(B)=0)$. Let $B=\cup_{\ell=1}^{m}J_\ell\times\tilde{A}_\ell$, where for each $\ell=1,\dots,m$, we have $J_\ell=[a_\ell,b_\ell)$ and $\tilde{A}_\ell$ is associated to some $A_\ell\in\mathscr{F}$ as in (3.21).

Since $\tilde{\pi}(\mathbb{X}_{n,i})\in\tilde{A}_\ell$ corresponds to the event $\mathbb{X}_{n,i}\in\mathbb{A}_\ell=\cup_{i\in\mathbb{Z}}\sigma^{-i}(A_\ell)$, then, by definition of the sets $J_{n,\ell}$ and $A_{n,\ell}$ given in (3.8) and (3.9), we have

$$\mathbb{P}(N_n(B)=0)=\mathbb{P}\left(\bigcap_{\ell=1}^{m}\{N_n(J_\ell\times\tilde{A}_\ell)=0\}\right)=\mathbb{P}\left(\bigcap_{\ell=1}^{m}\mathscr{W}_{J_{n,\ell}}\left(A_{n,\ell}\right)\right).$$

Now, by definition of the Poisson point process N, we have $\mathbb{P}(N(B)=0)=\prod_{\ell=1}^{m}\mathrm{e}^{-|J_\ell|\mu(\tilde{A}_\ell)}$. Hence, condition (i) of Proposition C.2 follows from Proposition 3.17 and (3.35).

In order to check condition (II) of Proposition C.2, we observe that by stationarity and Corollary 3.15 we have

$$\begin{aligned}\mathbb{E}(N_n(B))&=\mathbb{E}\left(\sum_{\ell=1}^{m}\sum_{i=\lceil k_na_\ell\rceil}^{\lceil k_nb_\ell\rceil-1}\mathbb{1}_{\{\tilde{\pi}(\mathbb{X}_{n,i})\in\tilde{A}_\ell\}}\right)\sim\sum_{\ell=1}^{m}|J_\ell|k_n\mathbb{P}\left(\tilde{\pi}(\mathbb{X}_{n,i})\in\tilde{A}_\ell\right)\\&\xrightarrow[n\to\infty]{}\sum_{\ell=1}^{m}|J_\ell|\mu(\tilde{A}_\ell)=\mathbb{E}(N(B)).\end{aligned}$$

□

3.5. **Application to jointly regularly varying sequences.**

Definition 3.19. A k-dimensional random vector $\mathbf{X}$ is said to be *jointly regularly varying*, with index $\alpha>0$, if there exists a sequence of constants $(a_n)_{n\in\mathbb{N}}$ and a random vector $\mathbf{\Theta}$ with $\mathbb{P}(\|\mathbf{\Theta}\|=1)=1$, such that

$$n\mathbb{P}(\|\mathbf{X}\|>xa_n,\ \mathbf{X}/\|\mathbf{X}\|\in\cdot)\xrightarrow[n\to\infty]{w}x^{-\alpha}\mathbb{P}(\mathbf{\Theta}\in\cdot).$$

where we are considering weak convergence of measures on $\mathbb{S}^{k-1}$, the unit sphere in $\mathbb{R}^k$. An $\mathbb{R}^d$-valued sequence $\mathbf{X}_0,\mathbf{X}_1,\dots$ is said to be *jointly regularly varying*, with index $\alpha>0$, if all the finite dimensional vectors $(\mathbf{X}_k,\dots,\mathbf{X}_\ell)$, $k\leq\ell\in\mathbb{N}_0$, are jointly regularly varying, with index $\alpha>0$.

For a jointly regularly varying sequence $\mathbf{X}_0, \mathbf{X}_1, \ldots$, in particular, condition (2.18) holds. Hence, taking $\tau = x^{-\alpha}$ and $u_n(\tau) = a_n \tau^{-\frac{1}{\alpha}}$, equation (2.14) holds. We also have that $u_n^{-1}(z) = z^{-\alpha} a_n^{\alpha}$. We define $\Xi\colon l_\infty \longrightarrow l_0$ given by $\Xi((x_j)_j) = (\xi(x_j))_j$, where ξ is defined as in (2.19). Observe that, as P, the function Ξ is invertible and we may define $\tilde{\Xi} : \tilde{l}_\infty \to \tilde{l}_0$ so that $\tilde{\Xi}(\tilde{\pi}(\mathbf{x})) = \tilde{\pi}(\Xi(\mathbf{x}))$. Observe that

$$\xi\left(u_n^{-1}(\|\mathbf{X}_j\|)\frac{\mathbf{X}_j}{\|X_j\|}\right) = \frac{\mathbf{X}_j}{a_n}. \tag{3.36}$$

Corollary 3.20. *Let $\mathbf{X}_0, \mathbf{X}_1, \ldots$ be a stationary $\mathbb{R}^d$-valued jointly regularly varying sequence, with tail index $\alpha > 0$, satisfying conditions Д$_{q_n}$ and Д$'_{q_n}$. Then the piling process given in Definition* 3.8 *is well defined and the point process*

$$N'_n = \sum_{i=1}^{\infty} \delta_{(i/k_n, \tilde{\Xi}(\tilde{\pi}(\mathbb{X}_{n,i})))} = \sum_{i=1}^{\infty} \delta_{\left(\frac{i}{k_n}, \tilde{\pi}\left(\frac{\mathbf{X}_{(i-1)r_n}}{a_n}, \ldots, \frac{\mathbf{X}_{ir_n-1}}{a_n}\right)\right)}. \tag{3.37}$$

converges weakly in $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_0 \setminus \{\mathbf{0}\}}$ to the Poisson point process N' given by

$$N' = \sum_{i=1}^{\infty} \delta_{(T_i, U_i^{-\frac{1}{\alpha}} \tilde{\Xi}(\tilde{\mathbf{Q}}_i))}, \tag{3.38}$$

where $(T_i)_{i\in\mathbb{N}}$, $(U_i)_{i\in\mathbb{N}}$ and $(\tilde{\mathbf{Q}}_i)_{i\in\mathbb{N}}$ are as in (3.33).

This corollary of Theorem 3.16 follows from a direct application of the CMT for the map $\Xi^{\#} : \mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}} \to \mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_0 \setminus \{\mathbf{0}\}}$ defined by

$$\Xi^{\#}\left(\sum_{i=1}^{\infty} \delta_{(t_i, \tilde{\mathbf{x}}_i)}\right) = \sum_{i=1}^{\infty} \delta_{(t_i, \tilde{\Xi}(\tilde{\mathbf{x}}_i))}.$$

Note that the existence of piling process and the respective properties are guaranteed by the joint regular variation of the process $\mathbf{X}_0, \mathbf{X}_1, \ldots$ (See [BS09]).

4. Proofs of the dynamical enriched functional limit theorems

The major step to obtain the invariance principles stated in Theorems 2.5, 2.9 and 2.11 for dynamically defined stochastic processes is the the complete convergence of point processes stated as follows

Theorem 4.1. *Let $(\mathcal{X}, \mathcal{B}_{\mathcal{X}}, T, \mu)$ be a dynamical system as described in Section* 2.2 *and $\mathbf{X}_0, \mathbf{X}_1, \ldots$ be obtained from such a system as described in* (2.1) *and assume that the piling process given in Definition* 3.8 *is well defined. Consider the point process N_n defined as in* (3.32)*. Then N_n converges weakly in $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ to the Poisson point process N given by* (3.33).

The systems and the observables that we consider give rise to stochastic processes for which Д$_{q_n}$ and Д$'_{q_n}$ hold and then the conclusion follows immediately by Theorem 3.16. A property which can be used to prove these conditions is in the next definition

Definition 4.2. Let $\mathcal{C}_1$ and C_2 be Banach spaces of real-valued measurable functions on $\mathcal{X}$. Define the *correlation* of non-zero functions $\phi \in \mathcal{C}_1$ and $\psi \in \mathcal{C}_2$ with respect to μ at time $n \in \mathbb{N}$ by

$$\mathrm{Cor}_\mu(\phi, \psi, n) := \frac{1}{\|\phi\|_{\mathcal{C}_1}\|\psi\|_{\mathcal{C}_2}} \left| \int \phi(\psi \circ T^n)\, d\mu - \int \phi\, d\mu \int \phi\, d\mu \right|.$$

Then say that the system has *decay of correlations*, with respect to μ, for observables in $\mathcal{C}_1$ *against* observables in $\mathcal{C}_2$ if there exists a rate function $\rho : \mathbb{N} \to [0, \infty)$ with

$$\lim_{n\to\infty} \rho(n) = 0,$$

and for every $\phi \in \mathcal{C}_1$, $\psi \in \mathcal{C}_2$,

$$\mathrm{Cor}_\mu(\phi, \psi, n) \le \rho(n).$$

The systems in Section 2.2.1 have decay of correlations against $L^1(\mu)$, i.e., where $\mathcal{C}_2 = L^1(\mu)$. In that setting we also require a suitable space $\mathcal{C}_1$, a key feature being that the characteristic functions on our sets of interest, like the annuli $A^{(j)}$ in (3.5) do not have large norm. In fact in the interval setting $\mathcal{C}_2$ will be the BV norm which we recall here. If $\psi : I \to \mathbb{R}$ is a measurable function on an interval I then its *variation* is defined as

$$\mathrm{Var}(\psi) := \sup \left\{ \sum_{i=0}^{n-1} |\psi(x_{i+1}) - \psi(x_i)| \right\},$$

where the supremum is taken over all finite ordered sequences $(x_i)_{i=0}^{n-1}$ in I. The BV norm is $\|\psi\|_{BV} := \sup|\psi| + \mathrm{Var}(\phi)$ and $BV := \{\psi : I \to \mathbb{R} : |\psi\|_{BV} < \infty\}$. In Saussol's class of higher dimensional expanding maps $\mathcal{C}_1$ is a quasi-Hölder norm.

Remark 4.3. While our conditions on the *rate* of decay of correlations here may appear very weak, in fact summable decay of correlations against $L^1(\mu)$ implies exponential decay of correlations for L^∞ observables against Hölder observables, as in [AFLV11].

We make a brief list of references where one can find the arguments to prove $Д_{q_n}$ and $Д'_{q_n}$ for the systems mentioned in Section 2.2. For non-invertible systems admitting decay of correlations against L^1 and observables with maximal sets $\mathcal{M}$ consisting on periodic points or a countable number of points in the same orbit, we note that our conditions on the system and the observable can be expressed as requiring, for $A_{n,\ell}^{(q_n)}$, as in (3.9) and (3.5),

(1) $\lim_{n\to\infty} \|\mathbb{1}_{A_{n,\ell}^{(q_n)}}\| n\rho(t_n) = 0$ for some sequence $(t_n)_n$ with $t_n = o(n)$,
(2) $\lim_{n\to\infty} \|\mathbb{1}_{A_{n,\ell}^{(q_n)}}\| \sum_{j=q_n}^{n} \rho(j) = 0$.

These conditions can be found in, for example [AFFR17, FFRS20]. Along with decay of correlations against $L^1(\mu)$, they imply $Д_{q_n}$, $Д'_{q_n}$: for proofs see those references. See also [FFMa20, Section 4.3].

Regarding the Benedicks-Carleson maps equipped with observables maximised at periodic points, the required estimates to satisfy $Д_{q_n}$ can be found in [FF08] and [FFT13, Sections 5,6].

As for the slowly mixing systems of the Manneville-Pomeau type and observables maximised at periodic points distinct from the indifferent fixed point, we defer the proof that this is sufficient to satisfy $Д_{q_n}$ and $Д'_{q_n}$ to the Section E, though most of the ideas of the proof are in [FFTV16].

For Anosov linear diffeomorphisms on the torus and observables maximised at periodic points, we refer to [CFF+15].

4.1. **Proof of Theorem 2.5.** As a consequence of Theorem 4.1, since in this case we are assuming that g appearing in (2.2) is of type g_2 and condition (2.18) holds, we obtain, by direct application of the CMT for the map $\Xi^{\#}$ given in Section 3.5, that the convergence of point processes stated in Corollary 3.20 holds. Therefore, we are left to show that such point process convergence implies the convergence in F' of the continuous time process S_n to $\underline{V}$.

Proposition 4.4. *Let $\mathbf{X}_0, \mathbf{X}_1, \dots$ be a stationary $\mathbb{R}^d$-valued process for which condition* (2.18) *holds and moreover N'_n defined in* (3.37) *converges weakly$^{\#}$ in $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_0 \setminus \{\mathbf{0}\}}$ to N' given in* (3.38) *then, under the same assumptions of Theorem* 2.5*, the conclusion regarding the convergence of S_n to $\underline{V}$ holds, in F'.*

Proof. Recall that the convergence in $F'([0,1],\mathbb{R}^d)$ consists of showing that the respective projections into $E([0,1],\mathbb{R}^d)$ and $\tilde{D}([0,1],\mathbb{R}^d)$ converge. The choice of metric in $E([0,1],\mathbb{R}^d)$ (see (2.9)) implies that the convergence in this space will follow from the convergence of the coordinate projections, in $E([0,1],\mathbb{R})$, which follows immediately from [BPS18, Theorem 4.5]. Hence, we are left to check the convergence of the $\tilde{D}([0,1],\mathbb{R}^d)$ counterparts, which we prove by splitting the argument into the same steps considered in [BPS18, Theorem 4.5] so that we can keep track of the required adjustments.

We start by defining a projection $\Upsilon\colon \mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_0 \setminus \{\mathbf{0}\}} \to F'([0,1],\mathbb{R}^d)$. Suppose we are given $\gamma = \sum_{i=1}^{\infty} \delta_{(t_i,\tilde{x}^i)} \in \mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_0 \setminus \{\mathbf{0}\}}$. At time t_i we have $\tilde{x}^i \in \tilde{l}_0$: let $(\dots, x_{-1}, x_0, x_1, \dots)$ be a representative of this in l_0 and define

$$e_x^{t_i}(t) = \sum_{i=-\infty}^{\lfloor \tan(\pi(t-\frac{1}{2}))\rfloor} x_i, \quad x(t) = \sum_{T_i<t} e_x^{T_i}(1), \quad S^x = \{t_i\}_i.$$

Note that since $e_x^{t_i}$ is really an element of an equivalence class, the particular representative of $\tilde{x}$ chosen does not matter. Finally, let $\Upsilon(\gamma) = (x, S^x, \{e_x^{t_i}\}_i)$.

We start assuming $\alpha \in (0,1)$.

Step 1. For $\varepsilon > 0$, we define an ε-truncated projection Υ^{ε}. We do this as follows. For $\tilde{x} \in \tilde{l}_0$ set

$$u^{\varepsilon}(\tilde{x})(t) = \sum_{j=-\infty}^{\lfloor \tan(\pi(t-\frac{1}{2}))\rfloor} x_j \mathbb{1}_{\{|x_j|>\varepsilon\}}.$$

Then for $\gamma = \sum_{i=1}^{\infty} \delta_{T_i,\tilde{x}^i} \in \mathcal{N}^{\#}_{\mathbf{X}}$, define $S^{\varepsilon} = \{T_i : \|\tilde{x}_i\|_{\infty} > \varepsilon\}$ and

$$\Upsilon^{\varepsilon}(\gamma) = \left(\left\{ \sum_{T_i \le t} u^{\varepsilon}(\tilde{x}^i)(1) \right\}_{t\in[0,1]}, S^{\varepsilon}, \{u^{\varepsilon}(\tilde{x}^s)\}_{s\in S^{\varepsilon}} \right)$$

(here we understand $\tilde{x}^s$ for $s = T_i \in S^\varepsilon$ as being $\tilde{x}^i$).

To show that given $\gamma_n \to_{w\#} \gamma$, we have $\tilde{\pi}(\Upsilon^\varepsilon(\gamma_n)) \to \tilde{\pi}(\Upsilon^\varepsilon(\gamma))$, we first define a simpler space which is not essential here, but is intended to help the reader's visualisation of the situation.

Given an element $z \in \tilde{D}$, let $(t_i)_i$ be the set $disc(z)$ of discontinuities of a representative of z, which we abuse notation and also call z. Let $v_i(z) = \sup_{t_i \le s_1, s_2 < t_{i+1}} \|z(s_1) - z(s_2)\|$ and $j(t_i) = \|z(t_i^-) - z(t_i)\|$. Then for $\varepsilon > 0$ set

$$\tilde{D}_\varepsilon := \{z \in \tilde{D} : \#disc(z) < \infty, v_i(z) = 0 \text{ and } j_i(z) > \varepsilon \text{ for all } i\}.$$

Now notice that since we are based in $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_0 \setminus \{\mathbf{0}\}}$, which only allows a finite number of jumps of norm higher than ε, each element of $\{u^\varepsilon(\tilde{x}^s)\}_{s \in S^\varepsilon}$ lies in $\tilde{D}_\varepsilon$. Moreover for $\underline{x} = \Upsilon^\varepsilon(\gamma)$, $\tilde{\pi}(\underline{x}) = \underline{x}^{\tilde{D}} \in \tilde{D}_\varepsilon$. We can easily see in this simpler space that if $\gamma_n \to_{w\#} \gamma$ then for $\underline{x}_n = \Upsilon^\varepsilon(\gamma_n)$ and $\underline{x} = \Upsilon^\varepsilon(\gamma)$, we have $\underline{x}_n^{\tilde{D}} \to \underline{x}^{\tilde{D}}$.

These arguments together complete the proof of Step 1.

Step 2. The aim here is to show that $\tilde{\pi}(\Upsilon^\varepsilon(N')) = \underline{x}_\varepsilon^{\tilde{D}} \to \tilde{\pi}(\Upsilon(N')) = \underline{x}^{\tilde{D}}$ as $\varepsilon \to 0$. It is easy to see that

$$d_{\tilde{D}}(\underline{x}_\varepsilon^{\tilde{D}}, \underline{x}^{\tilde{D}}) \le \sum_{i=1}^{\infty} \sum_{j \in \mathbb{Z}} U_i^{-\frac{1}{\alpha}} \|H(Q_{ij})\| \cdot \mathbb{1}_{\{U_i^{-\frac{1}{\alpha}} \|H(Q_{ij})\| \le \varepsilon\}}$$

And this is shown to converge to zero almost surely because condition (2.22) implies that $W_i = \sum_{j \in \mathbb{Z}} \|H(Q_{ij})\|$ is a.s. finite. Moreover, as noted in [BPS18, Remark 4.6], we have

$$\sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{Z}} U_i^{-\frac{1}{\alpha}} \|H(Q_{i,j})\| \mathbb{1}_{U_i^{-\frac{1}{\alpha}} \|H(Q_{i,j})\| \le \varepsilon} \to 0, \qquad \text{a.s. as } \varepsilon \to 0.$$

Note that condition (2.22), which is assumed by hypothesis for $1 < \alpha < 2$, holds for $0 < \alpha < 1$, as a byproduct of the convergence of the point processes stated in Theorem 3.16 as observed in [DH95, Theorem 2.6]. For $\alpha = 1$, a similar argument holds by making use of assumption (2.23).

Step 3. This step looks to compare the projection into $\tilde{D}$ of empirical process $\Upsilon(N_n')$ with its ε-truncated version $\Upsilon(N_n')$. Namely, we set $\underline{x}_{n,\varepsilon}^{\tilde{D}} := \tilde{\pi}(\Upsilon^\varepsilon(N_n'))$, $\underline{x}_n^{\tilde{D}} := \tilde{\pi}(\Upsilon(N_n'))$ and observe that

$$d_{\tilde{D}}(\underline{x}_{n,\varepsilon}^{\tilde{D}}, \underline{x}_n^{\tilde{D}}) \le \sum_{j=1}^{k_n r_n} \frac{\|\mathbf{X}_j\|}{a_n} \mathbb{1}_{\{\|\mathbf{X}_j\| \le a_n \varepsilon\}}.$$

Then, as in [BPS18, Proof of Theorem 4.5], it follows by Markov's inequality and Karamata's Theorem that

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}(d_{\tilde{D}}(\underline{x}_{n,\varepsilon}^{\tilde{D}}, \underline{x}_n^{\tilde{D}}) > \delta) &\le \limsup_{n\to\infty} \frac{k_n r_n}{\delta a_n} \mathbb{E}(\|\mathbf{X}_1\| \mathbb{1}_{\{\|\mathbf{X}_1\| \le a_n \varepsilon\}}) \\ &\le \lim_{n\to\infty} \frac{n}{\delta a_n} \frac{\alpha a_n \varepsilon \mathbb{P}(\|\mathbf{X}_1\| > a_n \varepsilon)}{1-\alpha} = \frac{\alpha}{\delta(1-\alpha)} \varepsilon^{1-\alpha} \xrightarrow[\varepsilon \to 0]{} 0,\end{aligned}$$

since we are assuming $\alpha < 1$.

Step 4. Here we wish to compare the $\tilde{D}$ component of the embedding of S_n into F' with the respective $\tilde{D}$ projection of $\Upsilon(N'_n)$. But we need only notice that the representations of each of these elements is in fact identical so there is nothing to do: in the $\Upsilon(N'_n)$ case all the jumps in a block are compressed into a single sequence, but the $\tilde{D}$ component undoes this compression, capturing all the jumps as an element of $D([0,1])$, which is equivalent to that obtained from S_n.

The proof of the $1 \leq \alpha < 2$ case does not require further arguments, though we do need to employ condition (2.21) in order to bound the error term in the equivalent to step 3 above. □

4.2. **Proof of Theorem 2.9.** We split the proof of Theorem 2.9 into several steps given in a series of lemmas.

We start by defining a map $\Upsilon\colon \mathcal{N}^{\#}_{\mathbb{R}_0^+\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}} \to F'((0,\infty),\mathbb{R})$, which to each $\gamma \in \mathcal{N}^{\#}_{\mathbb{R}_0^+\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}}$, such that $\gamma = \sum_{i=1}^{\infty}\delta_{(t_i,\tilde{\mathbf{x}}^i)}$, assigns

$$\Upsilon(\gamma) = \left(h(\gamma), \{t_i\colon\ i\in\mathbb{N}\}, e^{t_i}_{\gamma}\right),$$

where $h(\gamma)$ is defined for all $t>0$ by

$$h(\gamma)(t) := \begin{cases} \inf\{\|\tilde{P}(\tilde{\mathbf{x}}^i)\|_\infty^{-1}\colon\ t_i \leq t\} & \text{if } t > \underline{t} \\ \overline{y} & \text{if } t \leq \underline{t} \end{cases} \tag{4.1}$$

with $\underline{t} = \inf\{(t_i)_{i=1}^{\infty}\}$, $\overline{y} = \sup\{\inf\{\|\tilde{P}(\tilde{\mathbf{x}}^i)\|_\infty^{-1}\colon\ t_i \leq t\}\colon\ t > \underline{t}\}$ and, for $t\in[0,1]$

$$e^{t_i}_{\gamma}(t) = \min\{h(\gamma)(t_i^-), \inf_{j\leq\lfloor\tan(\pi(t-\frac{1}{2}))\rfloor}\|x^i_j\|\}.$$

Note that since we are using P here, we are looking for our observations to go *down* in value, so the objects here are all non-increasing functions.

Let $\Lambda = \left\{\gamma = \sum_{i=1}^{\infty}\delta_{(t_i,\tilde{\mathbf{x}}^i)}\colon\ t_i \neq t_j,\ \forall i\neq j;\ \gamma((0,a)\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}) > 0,\ \forall a>0\right\} \subset \mathcal{N}^{\#}_{\mathbb{R}_0^+\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}}$.

Lemma 4.5. *The map* $\Upsilon\colon \mathcal{N}^{\#}_{\mathbb{R}_0^+\times\tilde{l}_\infty\setminus\{\tilde{\infty}\}} \to F'((0,\infty),\mathbb{R})$ *is continuous on* Λ.

Proof. Assume that $\gamma_n \to_{w^\#} \gamma = \sum_{i=1}^{\infty}\delta_{(t_i,\tilde{\mathbf{x}}^i)} \in \Lambda$. Let $0<a<b\in\{t_i\colon i\in\mathbb{N}\}^c$, we need to verify that

$$m(\Gamma_{a,b,\underline{x}^E_n}, \Gamma_{a,b,\underline{x}^E}) + d_{\tilde{D}}\left(\underline{x}^{\tilde{D}}_{na,b}, \underline{x}^{\tilde{D}}_{a,b}\right) \to 0,$$

where $\Gamma_{a,b,\underline{x}^E}$ and $\underline{x}^{\tilde{D}}_{a,b}$ denote the restrictions to the time interval $[a,b]$ of the respective objects. Let $\alpha = h(\gamma)(a) \in (0,\infty)$ and consider the bounded set $\tilde{B}_\alpha = \{\tilde{\mathbf{x}}\in\tilde{l}_\infty\colon \|\tilde{P}(\tilde{\mathbf{x}})\|_\infty^{-1} < \alpha\}$ of $\tilde{l}_\infty\setminus\{\tilde{\infty}\}$. There are finitely many t_i such that $(t_i,\tilde{\mathbf{x}}^i)\in[a,b]\times\tilde{B}$. Let $(t_{i_1},\tilde{\mathbf{x}}^{i_1}),\ldots,(t_{i_k},\tilde{\mathbf{x}}^{i_k})$ be an enumeration of those points. Then, as in [Res87, Proposition 3.13], one can show that for n sufficiently large there exists a shadow enumeration $\left(t^{(n)}_{i_1},\tilde{\mathbf{x}}^{(n),i_1}\right),\ldots,\left(t^{(n)}_{i_k},\tilde{\mathbf{x}}^{(n),i_k}\right)$ of the mass points of γ_n in $[a,b]\times\tilde{B}_\alpha$ so that

$$t^{(n)}_{i_j} \to t_{i_j} \quad\text{and}\quad \tilde{\mathbf{x}}^{(n),i_j} \to \tilde{\mathbf{x}}^{i_j}, \quad \text{as } n\to\infty, \text{ for all } j=1,\ldots,k \tag{4.2}$$

Let t_i be such that $h(\gamma)(a) = \|\tilde{P}(\tilde{\mathbf{x}}^i)\|_\infty^{-1}$. Then applying the same argument to the bounded set $[t_i - \epsilon, a] \times B_{\alpha+\epsilon}$ where ϵ is chosen so that no mass point of γ lies on the border of this set, one obtains that $h(\gamma_n)(a) \to h(\gamma)(a)$, as $n \to \infty$. It follows that

$$h(\gamma_n)(t) \xrightarrow[n\to\infty]{} h(\gamma)(t), \quad \text{for all } t \in [a,b] \setminus \left(\bigcup_{n\in\mathbb{N}} \left\{ t_{i_1}^{(n)}, \dots, t_{i_k}^{(n)} \right\} \cup \{t_{i_1}, \dots, t_{i_k}\} \right). \tag{4.3}$$

The convergence stated in (4.2) regarding the quotient space $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$ means that we can take representatives for which the convergence is obtained componentwise, which allows us to choose corresponding representatives on the space of excursions $\tilde{D}([0,1],\mathbb{R})$ for which the following holds

$$e_{\gamma_n}^{t_i^{(n)}}(t) \to e_\gamma^{t_i}(t) \quad \text{as } n \to \infty, \text{ for all } t \in [0,1], \text{ which, in turn, implies} \tag{4.4}$$

$$\inf\left\{ e_{\gamma_n}^{t_i^{(n)}}(t) \colon t \in [0,1] \right\} \xrightarrow[n\to\infty]{} \inf\{e_\gamma^{t_i}(t) \colon t \in [0,1]\}, \tag{4.5}$$

From (4.2), (4.5) and (4.3), it follows that $m(\Gamma_{a,b,\underline{x}_n^E}, \Gamma_{a,b,\underline{x}^E}) \xrightarrow[n\to\infty]{} 0$. In a similar way, (4.2), (4.4) and (4.3) imply that $d_{\tilde{D}}\left(\underline{x}_{n\,a,b}^{\tilde{D}}, \underline{x}_{a,b}^{\tilde{D}}\right) \xrightarrow[n\to\infty]{} 0$. □

Lemma 4.6. *We have that $d_{F',\infty}(Z_n, \Upsilon(N_n)) \to 0$ as $n \to \infty$, in probability, where N_n is defined in* (3.32)

Proof. The processes Z_n and $\Upsilon(N_n))$ have different internal clocks, but $h(\Upsilon(N_n)))(i/k_n) = Z_n(ir_n/n)$, for all $i \in \mathbb{N}_0$. Observe that $h(\Upsilon(N_n)))$ is constant between i/k_n and $(i+1)/k_n$, having a possible jump at $(i+1)/k_n$ if $\inf\{u_n^{-1}(\mathbf{X}_j) \colon\ j = ir_n/n, \dots, (i+1)r_n/n - 1\} < h(\Upsilon(N_n)))(i/k_n)$. Note that when such a jump occurs, Z_n is not necessarily constant between ir_n/n and $(i+1)r_n/n$, namely, it may jump several times corresponding to the several moments $j = ir_n/n, \dots, (i+1)r_n/n$ at which $u_n^{-1}(\mathbf{X}_j)$ beats the running minimum until the threshold $\inf\{u_n^{-1}(\mathbf{X}_j) \colon\ j = ir_n/n, \dots, (i+1)r_n/n - 1\}$ is reached. Nevertheless, these oscillations are recorded by the excursion decorating $\Upsilon(N_n)$ at the discontinuity point $(i+1)/k_n$.

Note that by definition of r_n,

$$\frac{i}{k_n} \geq \frac{ir_n}{n} \geq \frac{i}{k_n} - \frac{i}{n} \quad \Rightarrow \quad \frac{i}{n} \geq \frac{i}{k_n} - \frac{ir_n}{n} \geq 0 \quad \text{and similarly} \quad \frac{i+1}{n} \geq \frac{i+1}{k_n} - \frac{(i+1)r_n}{n} \geq 0. \tag{4.6}$$

Hence, the maximum distance between one instant point of the time interval $[i/k_n, (i+1)/k_n)$ and another from $[ir_n/n, (i+1)r_n/n)$ is at most $\frac{i+1}{n}$.

Let $0 < a < b \in \mathbb{R}$, $i^- = \min\{j \in \mathbb{N}_0 \colon j/k_n \geq a\}$ and $i^+ = \max\{j \in \mathbb{N}_0 \colon jr_n/n < b\}$. Note that the range of the graphs $\Gamma_{i^- jr_n/n, i^+ jr_n/n, Z_n^E}$ and $\Gamma_{i^- jr_n/n, i^+ jr_n/n, \Upsilon(N_n)^E}$ is exactly the same, which means that the Hausdorff distance between the graphs is given by the time component, *i.e.*,

$$m\left(\Gamma_{i^- jr_n/n, i^+ jr_n/n, Z_n^E}, \Gamma_{i^- jr_n/n, i^+ jr_n/n, \Upsilon(N_n)^E}\right) \leq \frac{i+1}{n} \xrightarrow[n\to\infty]{} 0.$$

We only need to worry with the observations corresponding to $j \in \mathbb{N}_0$ such that $a \leq j/n < i^-/k_n$ where the range of the process Z_n may differ from the value of $h(\Upsilon(N_n))(a) =: \tau$, by either being above or below. Note that there are at most k_n such observations. This

means that for $|\epsilon| < \tau$, we have that $m\left(\Gamma_{a,i^-/k_n,Z_n^E}, \Gamma_{a,i^-/k_n,\Upsilon(N_n)^E}\right) > |\epsilon|$ implies that there must be at least one of these possible k_n observations such that $\|\mathbf{X}_j\| > u_n(\tau+\epsilon)$. Since $k_n\mathbb{P}(\|\mathbf{X}_j\| > u_n(\tau+\epsilon)) \to 0$, we conclude that $m\left(\Gamma_{a,i^-/k_n,Z_n^E}, \Gamma_{a,i^-/k_n,\Upsilon(N_n)^E}\right) \to 0$ in probability. A similar argument applies for the time frame between i^+jr_n/n and b.

Hence, $m\left(\Gamma_{a,b,Z_n^E}, \Gamma_{a,b,\Upsilon(N_n)^E}\right) \xrightarrow[n\to\infty]{} 0$, in probability. This takes care of the E component of the distance in F' and we need now to conclude that the same applies to the projections into $\tilde{D}$.

Let $i^- < i \leq i^+$ and assume that there is a jump of $h(\Upsilon(N_n))$ at i/k_n. Recall that while $h(\Upsilon(N_n)))$ is constant between $(i-1)/k_n$ and i/k_n, the process Z_n may have oscillations in the corresponding interval $[(i-1)r_n/n, ir_n/n]$. But, by construction, the excursion $e^{i/k_n}_{\Upsilon(N_n)}$ exactly mimics these oscillations of Z_n. Observe now that the projection $\Upsilon(N_n)^{\tilde{D}}$ reconstructs the behaviour of Z_n by incorporating in the time frame the excursions. Then, since the metric $\tilde{D}$ allows for time deformations, it is clear that the distance between the projections $\Upsilon(N_n)^{\tilde{D}}$ and $Z_n^{\tilde{D}}$, when restricted to $[i^-r_n/n, ir_n/n]$ is actually equal to 0. Again, we are left to analyse the time periods $[a, i^-r_n/n)$ and $[i^+r_n/n, b]$, where some pieces of information may be missing. However, as observed with the projections into E, the missing information corresponds to at most $\max\{k_n, r_n\}$ random variables and the probability of these producing oscillations that will not be mimicked is bounded by $\max\{k_n, r_n\}\mathbb{P}(\|\mathbf{X}_j\| > u_n(\tau+\epsilon)) \to 0$. Again, we conclude that $d_{\tilde{D}}\left(\Upsilon(N_n)^{\tilde{D}}_{a,b}, Z_{n\,a,b}^{\tilde{D}}\right) \xrightarrow[n\to\infty]{} 0$, in probability. □

Proof of Theorem 2.9. By Lemma 4.5, we may apply the CMT and conclude from Theorem 4.1 that $\Upsilon(N_n)$ converges weakly to $\Upsilon(N)$, in $F'((0,+\infty))$. By Lemma 4.6 and a Slutsky argument we conclude that Z_n converges weakly to $\Upsilon(N)$, in $F'((0,+\infty))$. Therefore, we are only left to show that $h(\Upsilon(N))$ has the prescribed finite-dimensional distributions of Z_H.

For each $\alpha > 0$, let $\tilde{B}_\alpha = \{\tilde{\mathbf{x}} \in\colon \|\tilde{P}(\tilde{\mathbf{x}})\|_\infty^{-1} < \alpha\} \subset \tilde{l}_\infty \setminus \{\tilde{\infty}\}$. For the unidimensional distribution, with $t, y \geq 0$ and $\bar{H}(\tau) = \mathrm{e}^{-\theta\tau}$, we have

$$\mathbb{P}(h(\Upsilon(N))(t) \geq y) = \mathbb{P}(N([0,t] \times \tilde{B}_y) = 0) = \mathrm{e}^{-\theta ty} = \bar{H}^t(y).$$

For the bidimensional distribution, with $0 \leq t_1 < t_2$ and $y_1 \geq y_2 \geq 0$,

$$\begin{aligned}&\mathbb{P}\left(h(\Upsilon(N))(t_1) \geq y_1, h(\Upsilon(N))(t_2) \geq y_2\right) = \\ &\quad = \mathbb{P}\left(N([0,t_1] \times B_{y_1}) = 0, N((t_1,t_2] \times B_{y_2}) = 0\right) = \bar{H}^{t_1}(y_1)\bar{H}^{t_2-t_1}(y_2).\end{aligned}$$

In case $0 \leq y_1 < y_2$,

$$\mathbb{P}\left(h(\Upsilon(N))(t_1) \geq y_1, h(\Upsilon(N))(t_2) \geq y_2\right) = \mathbb{P}(h(\Upsilon(N))(t_2) \geq y_2) = \bar{H}^{t_2}(y_2),$$

so in general

$$\mathbb{P}(h(\Upsilon(N))(t_1) \geq y_1, h(\Upsilon(N))(t_2) \geq y_2) = \bar{H}^{t_1}(y_1 \vee y_2)\bar{H}^{t_2-t_1}(y_2).$$

By induction we get for the k-dimensional distribution

$$\begin{aligned}\mathbb{P}(h(\Upsilon(N))(t_1) \geq y_1, \Upsilon(N))(t_2) \geq y_2 \ldots, h(\Upsilon(N))(t_k) \geq y_k) = \\ = \bar{H}^{t_1}\left(\bigvee_{i=1}^{k}\{y_i\}\right)\bar{H}^{t_2-t_1}\left(\bigvee_{i=2}^{k}\{y_i\}\right)\cdots\bar{H}^{t_k-t_{k-1}}(y_k) \\ = \mathbb{P}(Z_H(t_1) \geq y_1, \ldots, Z_H(t_k) \geq y_k).\end{aligned}$$

Regarding the statements about the Markov structure of the process Z_H, we give brief indications how to prove them and refer to [Res87, Chapter 4.1] for some useful properties of the theory of Markov processes. The fact that Z_H is Markov with the given transition probability follows from the form of the finite dimensional distributions. Namely, for $0 < y < z$,

$$\begin{aligned}\mathbb{P}(Z_H(t+s) \geq y \mid Z_H(s) = z) &= \mathbb{P}(h(\Upsilon(N))(t+s)) \geq y \mid h(\Upsilon(N))(s)) = z) \\ &= \mathbb{P}(N((s, s+t] \times B_y) = 0) = \mathrm{e}^{-\theta t y}. \end{aligned} \tag{4.7}$$

The exponential holding time parameter at state z, which is denoted by $\lambda(z)$, can be derived from

$$\mathbb{P}(Z_H(t+s) = Z_H(s) \mid Z_H(s) = z) = \mathbb{P}(N((s, s+t] \times B_z) = 0) = \mathrm{e}^{-\theta t z} \Rightarrow \lambda(z) = \theta z.$$

The jump distribution $\Pi(z, [0, y))$ can be computed from the equation

$$\lim_{t\to 0} t^{-1}\mathbb{P}(Z_H(t+s) < y) \mid Z_H(s) = z) = \lambda(z)\Pi(z, [0, y)).$$

But from (4.7), for $y < z$, we have $t^{-1}\mathbb{P}(Z_H(t+s) < y) \mid Z_H(s) = z) = t^{-1}(1 - \mathrm{e}^{-\theta t y}) \xrightarrow[t\to 0]{} \theta y$ and therefore $\Pi(z, [0, y)) = y/z$. □

4.3. **Proof of Theorem 2.11.** Consider the subset of $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ defined by

$$\Lambda = \left\{\gamma = \sum_{i=1}^{\infty} \delta_{(t_i, \tilde{\mathbf{x}}^i)} \colon\ t_j \neq t_\ell \text{ and } \|x_j^i\| = \|x_\ell^i\| \Rightarrow j = \ell \text{ or } \|x_j^i\| = \infty\right\}.$$

Since $(0, \infty)$ is locally compact and separable, then the weak$^{\#}$ topology in $\mathcal{N}^{\#}_{(0,+\infty)}$ coincides with the vague topology and by a trivial adjustment of [BPS18, Lemma 5.1] one obtains that the map $\gamma \to R_\gamma$ from $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ to $\mathcal{N}^{\#}_{(0,+\infty)}$ is continuous at every $\gamma \in \Lambda$.

Proposition 4.7. *Let $\mathbf{X}_0, \mathbf{X}_1, \ldots$ be a stationary $\mathbb{R}^d$-valued sequence, with tail index $\alpha > 0$, for which the point process N_n defined in* (3.32) *converges weakly in $\mathcal{N}^{\#}_{\mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}}$ to the Poisson point process N given by* (3.33)*. Then, under the same assumptions of Theorem* 2.11*, then the conclusion regarding the convergence of $R_{N_n''}$ to R_N holds, in $\mathcal{N}^{\#}_{(0,+\infty)}$, and the limit process has the representation given there.*

Together with Theorem 3.16, this concludes the proof of Theorem 2.11.

We elaborate a bit further on the representation of the limit record point process, which in this case is given in a more natural way through the use of the variable U with a uniform distribution. The rest of the proof of this proposition follows from trivial adjustments to the proof of [BPS18, Theorem 5.2].

Let $M = \sum_{i=1}^{\infty} \delta_{(T_i,U_i)}$ be the Poisson point process on $\mathbb{R}_0^+ \times \mathbb{R}_0^+$ defining the point process N in (3.33). Recall that $\|Q_j\| \geq 1$ for all j, but there exists one index j for which the equality holds. Let $(\tau_n)_{n\in\mathbb{Z}}$ denote the ordered sequence of record times of M counted by R_M. For definiteness, fix a certain $s > 0$ and let τ_1 denote the first record time larger than s, *i.e.*, $\tau_1 = \inf\{\tau_j \colon \tau_j < s\}$, and now denote the respective record frequencies by $Y_n = \inf_{T_i \leq \tau_n} U_i$, with $Y_0 = \inf_{T_i \leq s} U_i$. The fact that $\sum_{n\in\mathbb{Z}} \delta_{\tau_n}$ is a Poisson process with intensity $x^{-1}\,dx$ follows for example from [FFMa20, Theorem 5.7]. Note that for all n we have $Y_n/Y_{n-1} < 1$ and, in fact, from the jump distribution Π computed in Theorem 2.9 we obtain that $(Y_n/Y_{n-1})_{n\in\mathbb{N}}$ is a sequence of independent and uniformly distributed random variables.

The number of records of R_N observed at $\tau_n = T_i$, say, corresponds to the number of j's for which $Y_n\|Q_{i,j}\| < (Y_{n-1} \wedge \inf_{\ell<j}\{Y_n\|Q_{i,\ell}\|\}) \Leftrightarrow \|Q_{i,j}\| < (Y_{n-1}/Y_n \wedge \inf_{\ell<j}\{\|Q_{i,\ell}\|\})$, *i.e.*, is equal to $R^{\mathbf{Q}_i}(Y_{n-1}/Y_n)$, where $\mathbf{Q}_i = (Q_{i,j})_j$ is independent of the sequence $(Y_n/Y_{n-1})_{n\in\mathbb{N}}$ for all i. Since s was arbitrary, the given representation of R_N holds.

Appendix A. Completeness and separability of the space F'

As in Whitt's space E, there are two natural metrics to use for the E component of our space F'. These are denoted m_E and m_E^* (see (2.9) and (2.10)). We keep the same metric on the $\tilde{D}$ part.

Lemma A.1. *D with the sup norm is complete.*

Proof. Let $x_n \in D$ define a Cauchy sequence w.r.t. $\|\cdot\|_\infty$. Then set $x(t) = \lim_{n\to\infty} x_n(t)$. Note that $\cup_n disc(x_n)$ is an at most countable set. Observe also that in this metric (unlike M_1 say), each discontinuity t of x must correspond to a limit of discontinuities $(t_n)_n$ of $(x_n)_n$ (indeed we can take a sequence $(t_n)_n$ which is eventually constant). So x must have at most countably many discontinuities $disc(x)$. The fact that x is continuous on the right of the discontinuities follow from the continuity of x_n on the right. Hence $x \in D$. □

Lemma A.2. *$\tilde{D}$ is separable and complete.*

Proof. Completeness follows since D with the sup norm is complete: if $[x_n] \in \tilde{D}$ gives a Cauchy sequence then there are $y_n \in [x_n]$ such that $(y_n)_n$ is Cauchy in D with the sup norm. Then the limit y, which exists as in Lemma A.1, defines a class in $\tilde{D}$.

Separability follows since the set of piecewise constant $\mathbb{Q}$-valued functions in D with rational discontinuity points (the discontinuity points can be any countable set here) is countable and defines a set of equivalence classes which is dense in $\tilde{D}$. □

As shown in [Whi02, Theorem 15.4.3], (E, m_E) is separable, and in [Whi02, Example 15.4.2] this space is not complete. On the other hand for (E, m_E^*) is complete, but not separable.

Assume we are dealing with real valued functions rather than $\mathbb{R}^d$-valued for $d \geq 2$.

Proposition A.3. *(a) F' with the m_E-metric on the E-component is not complete, but is separable.*
(b) F' with the m_E^*-metric on the E-component is complete, but is not separable.

Proof. **(a)** [Whi02, Example 15.4.2] would not give a convergent sequence in F' since we would not be able to get convergent excursions: there are large jumps in adjacent terms of the sequence which cannot be matched up. Hence we need an alternative example for non-completeness.

For each n we define

$$\phi_n(x) = \begin{cases} nx & \text{if } x \in [0, \frac{1}{n}) \\ 1 & \text{if } x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

Also let, for each $k \in \mathbb{N}$, $x_{k,1}, \ldots, x_{k,n+k}$ be the set of $n+k$ points in $[\frac{1}{2^k}, \frac{1}{2^{k-1}}]$ so that these points, along with the boundary points, are equidistributed through this interval. Define $\varepsilon_{n,k} = \frac{1}{n2^{k+2}}$. Then define

$$x_n = \sum_{k=1}^{\infty} \sum_{i=1}^{n+k} \phi_n \cdot \mathbb{1}_{\left[x_{k,i}-\varepsilon_{n,k}, x_{k,i}+\varepsilon_{n,k}\right)}.$$

Then $S_n = \{x_{k,i} \pm \varepsilon_{n,k}\}_{k,n}$ and the excursions are for example $\phi_n(x_{k,i} - \varepsilon_{n,k}) \cdot \mathbb{1}_{[1/2,1]}$ at the points $x_{k,i} - \varepsilon_{n,k}$ and $\phi_n(x_{k,i} + \varepsilon_{n,k}) \cdot \mathbb{1}_{[0,1/2)}$ at the points $x_{k,i} + \varepsilon_{n,k}$, producing an element $\underline{x}_n \in F'$ (note the damping effect of ϕ_n ensuring that (2.8) holds). Then $(\underline{x}_n)_n$ is Cauchy with the m-metric, but this sequence does not have a limit in F' (eg the graph of the limit would have to be $[0,1] \times [0,1]$ as in [Whi02, Example 15.4.2]). Hence the space is not complete.

To show separability, we use the usual approximation by objects with rational coordinates. As shown in [Whi02, Theorem 15.4.3], the E-part of the space is separable. We then use Lemma A.2 to give separability of the excursion part.

(b) Completeness follows from the E-component being complete as in [Whi02, Section 15.4], which is the same type of argument as completeness in D (see Lemma A.1) and $\tilde{D}$ since this is uniform convergence. The latter facts give convergence to a limit in $\tilde{D}$ for the excursions also.

Non-separability of E (with m^*) implies the non-separability of F': in fact this is the same type of argument as for D. Note that for non-separability it is sufficient, for a given positive distance, to find an uncountable collection of elements of the set all at least that distance apart. So for example for D, $\{\mathbb{1}_{[a,1]}\}_{a\in[0,1)}$ are all distance 1 apart in the m^* metric. □

Remark A.4. [Whi02, Example 15.4.2] would not give a convergent sequence in F' since we would not be able to get convergent excursions. Hence we needed an alternative example above for non-completeness.

Appendix B. Weak and weak$^{\#}$ convergence

The purpose of this section is to review notions of convergence of measures on the metric spaces $\tilde{l}_\infty$ and $\tilde{l}_0$. We are particularly interested in weak convergence of probability measures and, since we want to consider point processes which are random elements corresponding to boundedly finite measures, we are also interested in weak$^{\#}$ convergence. Recall that $\tilde{l}_\infty$ and $\tilde{l}_0$ are complete, separable, metric spaces, which are not locally compact and therefore vague convergence is not useful here. We remark that weak$^{\#}$ convergence is equivalent to vague convergence when the ambient space is locally compact (see [DVJ03, Appendix A2.6]).

The Portmanteau Theorem (see for example [Bil99, Theorem 2.1]) is the following.

Theorem B.1. *Given a metric space, the following conditions are equivalent to weak convergence of probability measures.*

(1) $\lim_{n\to\infty}\int f d\mathbb{P}_n = \int f d\mathbb{P}$*, for all bounded, uniformly continuous real* f*;*
(2) $\limsup_{n\to\infty}\mathbb{P}_n(F) \le \mathbb{P}(F)$ *for all closed set* F*;*
(3) $\liminf_{n\to\infty}\mathbb{P}_n(E) \ge \mathbb{P}(E)$ *for all open set* E*;*
(4) $\lim_{n\to\infty}\mathbb{P}_n(A) = \mathbb{P}(A)$ *for all set* A *such that* $\mathbb{P}(\partial A) = 0$*.*

A set A such that $\mathbb{P}(\partial A) = 0$ is called a $\mathbb{P}$-continuity set. A class $\mathscr{U}$ of sets is called *a convergence determining class* if the convergence $\lim_{n\to\infty}\mathbb{P}_n(A) = \mathbb{P}(A)$ for all $\mathbb{P}$-continuity sets $A \in \mathscr{U}$ implies the weak convergence of $\mathbb{P}_n$ to $\mathbb{P}$.

Lemma B.2. *The class of sets* $\tilde{\mathscr{J}}$ *is a convergence determining class for weak convergence on the metric space* $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$*.*

Proof. We start by proving a claim.

Claim. *For all* $\tilde{\mathbf{x}} \in \tilde{l}_\infty \setminus \{\tilde{\infty}\}$ *and all* $\varepsilon > 0$*, there exists a* $\mathbb{P}$*-continuity set* $\tilde{A} \in \tilde{\mathscr{J}}$ *such that* $\tilde{\mathbf{x}} \in \mathring{\tilde{A}} \subset \tilde{A} \subset B(\tilde{\mathbf{x}}, \varepsilon)$*.*

Proof of Claim. Let $\mathbf{z}, \mathbf{w} \in l_0 \setminus \{\mathbf{0}\}$. Since $|\|\mathbf{z}\|_\infty - \|\mathbf{w}\|_\infty| \le \|\mathbf{z} - \mathbf{w}\|_\infty$ and $h(x) = \frac{1}{x}$ is continuous on $\mathbb{R} \setminus \{0\}$, then there exists $\delta > 0$ such that $\|\mathbf{z} - \mathbf{w}\|_\infty < \delta$ implies that $d'(\mathbf{z}, \mathbf{w}) < \varepsilon$, where d' is the metric on $l_0 \setminus \{\mathbf{0}\}$ defined in (3.17).

Let $\mathbf{z} = P(\mathbf{x}) \in l_0 \setminus \{\mathbf{0}\}$ and let $k \in \mathbb{N}$ be such that $\|z_j\| < \delta$ for all $|j| \ge k$. Define

$$B_\rho = \bigcap_{|j|<k} \sigma^{-j}(B(z_j, \rho)),$$

for some $\rho \le \delta$. Let $A_\rho = P^{-1}(B_\rho)$ and let $\tilde{A}_\rho \in \tilde{\mathscr{J}}$ be associated to A_ρ as in (3.21). Then $\tilde{A}_\rho$ will be the open set in the statement of the claim. We choose ρ so that $\tilde{A}_\rho$ is a $\mathbb{P}$-continuity set. We can always choose such ρ because $\partial\tilde{A}_\rho \subset \tilde{\pi}\left(\cup_{i\in\mathbb{Z}}(\sigma^{-i}(\cup_{|j|<k}\{\mathbf{w} : \|p(w_j) - z_j\|_\infty = \rho\}))\right)$, which means that each $\partial\tilde{A}_\rho$ can intersect at most countably many other such sets. Therefore, there is an uncountable number of disjoint sets $\partial\tilde{A}_\rho$, for $0 < \rho < \delta$, and since there cannot be an uncountable number of them with positive probability, at least one of them must have 0 probability, which means one of the $\tilde{A}_\rho$ is a $\mathbb{P}$-continuity set.

We claim that $\tilde{A}_\rho \subset B(\tilde{\mathbf{x}}, \varepsilon)$. To see this, let $\mathbf{w} \in l_\infty$ be such that $\tilde{\mathbf{w}} = \tilde{\pi}(\mathbf{w}) \in \tilde{A}_\rho$. By definition of $\tilde{A}_\rho$ we must have $\sigma^\ell(\mathbf{w}) \in A$, for some $\ell \in \mathbb{Z}$. Hence, $P(\sigma^\ell(\mathbf{w})) = \sigma^\ell(P(\mathbf{w})) \in B_\rho$, which means that $\|P(\mathbf{x}) - \sigma^\ell(P(\mathbf{w}))\|_\infty < \delta$ and therefore $d'(P(\mathbf{x}), \sigma^\ell(P(\mathbf{w}))) < \varepsilon$. It follows that, by definition of the metric $\tilde{d}'$ on $\tilde{l}_0$, given in (3.20), we have $\tilde{d}'(\tilde{P}(\tilde{\mathbf{x}}), \tilde{P}(\tilde{\mathbf{w}})) < \varepsilon$ and hence $\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) < \varepsilon$. □

Let E be an open set. By the claim, for each $\tilde{\mathbf{x}} \in E$ there exists a $\mathbb{P}$-continuity set $\tilde{A} \in \tilde{\mathscr{J}}$ such that $\tilde{\mathbf{x}} \in \mathring{\tilde{A}} \subset \tilde{A} \subset E$. This means we have an open cover of E by sets $\mathring{\tilde{A}}$, where $\tilde{A} \in \tilde{\mathscr{J}}$ is a $\mathbb{P}$-continuity set. Since on a separable metric space (such as $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$) each open cover admits a countable subcover (see [Bil99, Appendix M3], for example)then there exists a sequence

$(\tilde{A}_i)_{i\in\mathbb{N}}$ such that each $\tilde{A}_i \in \tilde{\mathscr{J}}$ is a $\mathbb{P}$-continuity set, $\tilde{A}_i \subset E$ and, moreover, $E \subset \cup_{i\in\mathbb{N}} \mathring{\tilde{A}}_i$, which means that $E = \cup_{i\in\mathbb{N}} \tilde{A}_i$.

We are assuming that $\lim_{n\to\infty} \mathbb{P}_n(\tilde{A}) = \mathbb{P}(\tilde{A})$ for all $\mathbb{P}$-continuity sets $\tilde{A} \in \tilde{\mathscr{J}}$. Since $\tilde{\mathscr{J}}$ is closed for finite unions and $\partial(\tilde{A} \cup \tilde{B}) \subset \partial\tilde{A} \cup \tilde{B}$ (see [DVJ03, Proposition A1.2.I]), we have that $\cup_{i=1}^m \tilde{A}_i \in \tilde{\mathscr{J}}$ is a $\mathbb{P}$-continuity set and therefore

$$\lim_{n\to\infty} \mathbb{P}_n(\cup_{i=1}^m \tilde{A}_i) = \mathbb{P}(\cup_{i=1}^m \tilde{A}_i).$$

Since $E = \cup_{i\in\mathbb{N}} \tilde{A}_i$, for $\varepsilon > 0$, let $m \in \mathbb{N}$ be such that $\mathbb{P}(\cup_{i=1}^m \tilde{A}_i) > \mathbb{P}(E) - \varepsilon$. Then,

$$\mathbb{P}(E) - \varepsilon < \mathbb{P}(\cup_{i=1}^m \tilde{A}_i) = \lim_{n\to\infty} \mathbb{P}_n(\cup_{i=1}^m \tilde{A}_i) \le \liminf_{n\to\infty} \mathbb{P}_n(E).$$

Since the last inequalities hold for all $\varepsilon > 0$ then

$$\mathbb{P}(E) \le \liminf_{n\to\infty} \mathbb{P}_n(E), \quad \text{for all } E \text{ open}$$

and therefore by Theorem B.1(3), $\mathbb{P}_n$ converges weakly to $\mathbb{P}$. □

In order to study point processes, we need to consider σ-finite measures taking finite measures on bounded sets. Namely, we define:

Definition B.3. A Borel measure μ on a complete, separable, metric space is boundedly finite if $\mu(A) < \infty$ for every bounded Borel set A.

Let $\mathcal{X}$ denote a complete, separable, metric space such as $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$. We denote by $\mathcal{M}_{\mathcal{X}}^{\#}$ the space of boundedly finite Borel measures on $\mathcal{X}$. Following [DVJ03, DVJ08], we consider a notion of convergence in $\mathcal{M}_{\mathcal{X}}^{\#}$ called the *weak hash* or *weak*$^{\#}$ convergence, denoted by $\mu_k \to_{w\#} \mu$ which can be defined by any of the following equivalent conditions (see [DVJ03, Proposition A2.6.II]):

(i) $\lim_{k\to\infty} \int f d\mu_k = \int f d\mu$ for all bounded continuous functions f defined on $\mathcal{X}$ and vanishing outside a bounded set;
(ii) There exists an increasing sequence of bounded open sets B_n converging to $\mathcal{X}$ such that if $\mu_k^{(n)}$ and $\mu^{(n)}$ denote the restrictions of the measures μ_k and μ to B_n, respectively, then $\mu_k^{(n)}$ converges weakly to $\mu^{(n)}$, as $k \to \infty$, for all $n \in \mathbb{N}$. Note that $\mu_k^{(n)}$ and $\mu^{(n)}$ are not necessarily probability measures, so when we say that there is weak convergence we mean that either (1) and (4) from Theorem B.1 apply or (2) and (3) apply with the extra assumption that $\lim_{k\to\infty} \mu_k^{(n)}(B_n) = \mu^{(n)}(B_n)$, for all $n \in \mathbb{N}$.
(iii) $\lim_{k\to\infty} \mu_k(A) = \mu(A)$ for all bounded Borelean A such that $\mu(\partial A) = 0$.

Similarly to Lemma B.2, we show that in order to prove weak$^{\#}$ convergence on $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$, we only need to check $\lim_{k\to\infty} \mu_k(\tilde{A}) = \mu(\tilde{A})$ for all bounded, μ-continuity sets $\tilde{A} \in \tilde{\mathscr{J}}$, as in property (iii).

Lemma B.4. *The class of bounded sets in $\tilde{\mathscr{J}}$ is a convergence determining class for weak$^{\#}$ convergence on the metric space $\tilde{l}_\infty \setminus \{\tilde{\infty}\}$.*

Proof. Suppose $\mu_k \to_{w\#} \mu$. For all $a > 1$, let $\mathscr{F} \ni B_a := \{\mathbf{x} \in \dot{\mathcal{V}}^{\mathbb{Z}}\colon\ 1/a < \|x_0\| < a\}$ and $\tilde{B}_a \in \tilde{\mathscr{J}}$ be associated to B_a as in (3.21). Observe that $\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}) \le a - 1/a$, for all

$\tilde{x}, \tilde{w} \in \tilde{B}_a$. Since there are uncountably many sets of the form B_a and $\partial B_a \subset \{\mathbf{x} \in \dot{\mathcal{V}}^{\mathbb{Z}}\colon\ 1/a = \|x_0\|$ or $\|x_0\| = a\}$, then we can find a strictly increasing, diverging sequence $(a_n)_{n\in\mathbb{N}}$, such that $B_n := B_{a_n}$ is such that $\mu(\partial \tilde{B}_n) = 0$, for all $n \in \mathbb{N}$.

Hence, $(B_n)_{n\in\mathbb{N}}$ is an increasing sequence of bounded, open, μ-continuity sets converging to $\tilde{l}_\infty \backslash \{\tilde{\infty}\}$. Note that by hypothesis we are assuming that $\lim_{k\to\infty} \mu_k(\tilde{A}) = \mu(\tilde{A})$ for all bounded, μ-continuity sets $\tilde{A} \in \tilde{\mathscr{J}}$ and therefore that applies to all the $\tilde{B}_n$, namely, $\lim_{k\to\infty} \mu_k(\tilde{B}_n) = \mu(B_n)$. Therefore, we can apply (3) to prove weak convergence of $\mu_k^{(n)}$ (the restriction of μ_k to B_n) to $\mu^{(n)}$ (the restriction of μ to B_n), which was the main tool used in the proof of Lemma B.2. It follows then, by assumption and Lemma B.2 that $\mu_k^{(n)}$ converges weakly to $\mu^{(n)}$, for all $n \in \mathbb{N}$ and, by (ii), we obtain the weak$^\#$ convergence of μ_k to μ. □

Appendix C. Convergence of point processes on non locally compact spaces

We closely follow [DVJ03, Appendix A2.6] and [DVJ08, Sections 9 and 11]. As before, let $\mathcal{X}$ be a complete separable metric space. The notion of weak$^\#$ convergence is metrizable, in the sense that $\mathcal{M}_{\mathcal{X}}^{\#}$ admits a metric generating what the so-called $w^\#$-topology so that the weak$^\#$ convergence corresponds to convergence in the $w^\#$-topology. Denote by $\mathcal{B}(\mathcal{M}_{\mathcal{X}}^{\#})$ the corresponding Borel σ-algebra. A *random measure* is a random element in $(\mathcal{M}_{\mathcal{X}}^{\#}, \mathcal{B}(\mathcal{M}_{\mathcal{X}}^{\#}))$. A point process N is an integer valued random measure. Let $\mathcal{N}_{\mathcal{X}}^{\#}$ denote the space of boundedly finite integer valued measures. We have that $\mathcal{N}_{\mathcal{X}}^{\#}$ is a closed subset of $\mathcal{M}_{\mathcal{X}}^{\#}$ ([DVJ08, Proposition 9.1.V]) and let $\mathcal{B}(\mathcal{N}_{\mathcal{X}}^{\#})$ denote the corresponding σ-algebra for the $w^\#$-topology. We remark that $\mu \in \mathcal{N}_{\mathcal{X}}^{\#}$ has the following form ([DVJ08, Proposition 9.1.III]):

$$\mu = \sum_{i\in\mathbb{N}} k_i \delta_{x_i}, \quad \text{where } \delta_{x_i} \text{ is the Dirac measure at distinct } x_i \in \mathcal{X}, \text{ and } k_i \in \mathbb{N}. \tag{C.1}$$

For any $\mu \in \mathcal{N}_{\mathcal{X}}^{\#}$ given by (C.1), we define its *support counting measure* μ^* as:

$$\mu^* = \sum_{i\in\mathbb{N}} \delta_{x_i}. \tag{C.2}$$

The boundedly finite measure $\mu \in \mathcal{N}_{\mathcal{X}}^{\#}$ is simple if and only if $\mu = \mu^*$, which is equivalent to say that $k_i = 1$ for all $i \in \mathbb{N}$.

Now, we give formally now the definition of point process.

Definition C.1. A point process N on state space $\mathcal{X}$ is measurable mapping from a probability space $(\Omega, \mathcal{B}, \mathbb{P})$ into $(\mathcal{N}_{\mathcal{X}}^{\#}, \mathcal{B}(\mathcal{N}_{\mathcal{X}}^{\#}))$. A point process N is said to be *simple* if $\mathbb{P}(N(\{x\}) > 1) = 0$, for all $x \in \mathcal{X}$. To each point process N on state space $\mathcal{X}$ we denote by N^* the corresponding *support point process* obtained from N as in (C.2).

Let $(N_n)_{n\in\mathbb{N}}$ be a sequence of point processes and N another point process, all with state space $\mathcal{X}$. We say that N_n converges weakly to N, when the respective distributions P_n, defined by $P_n(A) := \mathbb{P}(N_n \in A)$, for all $A \in \mathcal{B}(\mathcal{N}_{\mathcal{X}}^{\#})$), converge weakly (in the sense of weak convergence of probability measures on the metric space $\mathcal{N}_{\mathcal{X}}^{\#}$) to the distribution P associated to N.

Tightness is a very useful property which gives relative compactness and ultimately weak convergence. We state two conditions that are necessary and sufficient to show that a family

of probability measures $\{P_t,\ t \in \mathcal{T}\}$ is *uniformly tight*, see [DVJ08, Proposition 11.1.VI]. Given any closed sphere $S \subset \mathcal{X}$ and any $\epsilon, \delta >$, there exists a real number M and a compact set $C \subset S$ such that for $t \in \mathcal{T}$,

$$P_t(N_t(S) > M) < \epsilon, \tag{C.3}$$

$$P_t(N_t(S \setminus C) > \delta) < \epsilon. \tag{C.4}$$

We remark that if $\mathcal{X}$ was locally compact, (C.4) would be unnecessary.

Based on these criteria for tightness, in [DVJ08, Proposition 11.1.VII] it is shown that weak convergence of N_n to N follows from the convergence:

$$(N_n(A_1), \dots, N_n(A_k)) \to (N(A_1), \dots, N(A_k)),$$

of joint distributions as random vectors in $\mathbb{R}^k$, for all finite collections $\{A_1, \dots, A_k\}$ of bounded continuity sets $A_i \in \mathcal{B}_{\mathcal{X}}$, for all $i = 1, \dots, k$ and all $k \in \mathbb{N}$. Here, *continuity set* means that $N(\partial A_i) = 0$ a.s.

When the limiting point process is simple, a simpler criterion for convergence can be used.

Proposition C.2. *Let $(N_n)_{n\in\mathbb{N}}$ be a sequence of point processes on the state space $\mathcal{X}$ and let N be simple point process on the same state space. Let $\mathcal{R}$ be a covering dissecting ring*[3] *of continuity sets of $\mathcal{X}$. Then, N_n converges weakly to N if the two following conditions hold:*

(i) $\lim_{n\to\infty} \mathbb{P}(N_n(A) = 0) = \mathbb{P}(N(A) = 0)$ *for all bounded $A \in \mathcal{R}$;*
(ii) for all bounded $A \in \mathcal{R}$ and a nested sequence of partitions $\mathcal{T}_r = \{A_{ri}\colon\ i = 1, \dots, k_r\}$ of A by sets of $\mathcal{R}$ that ultimately separate the points of A,

$$\inf_{\mathcal{T}_r} \limsup_{n\to\infty} \sum_{i=1}^{k_r} \mathbb{P}(N_n(A_{ri}) \geq 2) = 0.$$

Alternatively, we may replace condition (ii) *by the following:*

(II) $\lim_{n\to\infty} \mathbb{E}(N_n(A)) = \mathbb{E}(N(A))$*, for all bounded $A \in \mathcal{R}$.*

Remark C.3. This proposition is very similar to [DVJ08, Proposition 11.1.IX]. However, note that the corresponding statement to our condition (ii), namely [DVJ08, equation (11.1.4)], is incorrectly stated there. Essentially, this condition is supposed to require that there is no accumulation of mass points as the point processes approach the limiting point process. This is related to a property called *ordinary* in [DVJ08, equation (11.1.4)], which is then used in the proof of [DVJ08, Proposition 11.1.IX]. When one observes the condition giving the notion of ordinary point process one realises the need to use the infimum over all partitions, as we did in (ii), which contrasts with the use of the supremum over all partitions used in [DVJ08, equation (11.1.4)].

For completeness we redo the proof of [DVJ08, Proposition 11.1.IX], correcting this typo. We also extend the result by showing that condition (II) can be used to replace condition (ii) when trying to prove weak convergence to a simple point process.

[3] A dissecting ring is a ring generated by finite unions and intersections of the elements of a dissecting system, which consists on a nested sequence of finite partitions of the whole space that eventually separate points. See [DVJ03, Appendices A1.6 and A2.1] for further details.

Proof. We closely follow the proof of [DVJ08, Proposition 11.1.IX]. At the core of the proof this proposition is a result attributed to Rényi and Mönch, see [DVJ08, Theorem 9.2.XII], which states that the distribution of a simple point process on a complete separable metric spaces $\mathcal{X}$ is determined by the values of the avoidance function P_0 on the bounded sets of a dissecting ring $\mathcal{R}$ for $\mathcal{X}$, where P_0 is defined by:

$$P_0(A) = \mathbb{P}(N(A) = 0), \quad \text{for } A \in \mathcal{R}.$$

We are then left to show that the family $\{P_n \colon n \geq n_0\}$ is uniformly tight and that the limit of any convergent subsequent is a simple point process, where P_n is the distribution of the point process N_n.

Assuming that (ii) holds, let A be a closed ball in $\mathcal{R}$. Observing that $\{N_n(A) > k_r\}$ implies $\{N_n(A_{ri}) \geq 2\}\}$ for at least one i, then

$$\sum_{i=1}^{k_r} \mathbb{P}(N_n(A_{ri}) \geq 2) \geq \mathbb{P}(N_n(A) \geq k_r).$$

Note that the sum on the left is non-increasing with r. Hence, given $\epsilon > 0$, by (ii), there exists an r_0 such that for all $r \geq r_0$, we have $\limsup_{n\to\infty} \sum_{i=1}^{k_r} \mathbb{P}(N_n(A_{ri}) \geq 2) < \epsilon$. In particular, there exists $n_0 \in \mathbb{N}$ such that for all $n > n_0$, we have $\mathbb{P}(N_n(A) \geq k_{r_0}) \leq \sum_{i=1}^{k_{r_0}} \mathbb{P}(N_n(A_{r_0 i}) \geq 2) < \epsilon$. Therefore, choosing $M > k_{r_0}$ large enough, we have $\mathbb{P}(N_n(A) \geq M) < \epsilon$, for all $n \in \mathbb{N}$, which means that C.3 holds.

Now, we verify that the same happens when we assume (II) instead of (ii). Using Chebyshev's inequality, we obtain for all $n \in \mathbb{N}$

$$\mathbb{P}(N_n(S) > M) \leq \frac{\mathbb{E}(N_n(S))}{M}.$$

By (II), there exists $K \in \mathbb{N}$ such that $\mathbb{E}(N_n(S)) < K$, for all $n \in \mathbb{N}$. Therefore, taking M sufficiently large so that $K/M < \epsilon$, we obtain that $\mathbb{P}(N_n(S) > M) < \epsilon$, for all $n \in \mathbb{N}$ and, therefore, C.3 holds.

We next show that (i) implies C.4, which can be restated here in the following form: given $\epsilon > 0$. there exists a compact set C such that $\mathbb{P}(N_n(S \setminus C) = 0) > 1 - \epsilon$, for all $n \in \mathbb{N}$. We choose C so that for the limit distribution (which corresponds to that of a simple point process), we have

$$\mathbb{P}(N(S \setminus C) = 0) > 1 - \epsilon/2.$$

From (i), we have that $\lim_{n\to\infty} \mathbb{P}(N_n(S \setminus C) = 0) = \mathbb{P}(N(S \setminus C) = 0)$, from which we obtain that $\mathbb{P}(N_n(S \setminus C) = 0) > 1 - \epsilon$, for all n sufficiently large and, by taking a larger C if necessary, for all $n \in \mathbb{N}$, which means C.4 holds.

Since, both conditions C.3 and C.4 hold then N_n admits a weakly convergent subsequence, say $(N_{n_\ell})_{\ell\in\mathbb{N}}$, which converges weakly to $\bar{N}$. By (i), we have

$$\mathbb{P}(N(A) = 0) = \mathbb{P}(\bar{N} = 0),$$

but we still do not know if $\bar{N}$ is simple, hence this only gives us that N and $\bar{N}^*$ have the same distribution, where $\bar{N}^*$ is the support point process associated to $\bar{N}$.

Hence, we are left to prove that both (ii) and (II) imply that $\bar{N}$ is simple, *i.e.*, $\bar{N} = \bar{N}^*$, a.s.

Note that for all r, we have

$$\sum_{i=1}^{k_r} \mathbb{P}(\bar{N}(A_{ri}) \geq 2) = \lim_{\ell\to\infty} \sum_{i=1}^{k_r} \mathbb{P}(N_{n_\ell}(A_{ri}) \geq 2).$$

Assuming (ii), we obtain

$$\inf_{\mathcal{T}_r} \sum_{i=1}^{k_r} \mathbb{P}(\bar{N}(A_{ri}) \geq 2) = \inf_{\mathcal{T}_r} \lim_{\ell\to\infty} \sum_{i=1}^{k_r} \mathbb{P}(N_{n_\ell}(A_{ri}) \geq 2) \leq \inf_{\mathcal{T}_r} \limsup_{n\to\infty} \sum_{i=1}^{k_r} \mathbb{P}(N_n(A_{ri}) \geq 2) = 0.$$

This means that the point process $\bar{N}$ is ordinary, as defined in [DVJ08, Definition 9.3.XI], and, therefore, by [DVJ08, Proposition 9.3.XII], $\bar{N}$ is simple.

Now, assuming (II) instead, we have for all bounded $A \in \mathcal{R}$,

$$\mathbb{E}(N(A)) = \mathbb{E}(\bar{N}^*(A)) \leq \mathbb{E}(\bar{N}(A)) = \lim_{\ell\to\infty} \mathbb{E}(N_{n_\ell}(A)) = \lim_{n\to\infty} \mathbb{E}(N_n(A)) = \mathbb{E}(N(A)),$$

which means that $\mathbb{E}(\bar{N}^*(A)) = \mathbb{E}(\bar{N}(A))$ for all bounded $A \in \mathcal{R}$, which implies that $\bar{N}$ must be simple. □

Remark C.4. In fact, since given any sequence of positive random variables $(X_n)_{n\in\mathbb{N}}$ converging in distribution to X, we have $\liminf_{n\to\infty} \mathbb{E}(X_n) \geq \mathbb{E}(X)$, then we can strengthen Proposition C.2 by replacing (II) by $\limsup_{n\to\infty} \mathbb{E}(N_n(A)) \leq \mathbb{E}(N(A))$, which actually only need verifying for all bounded sets in semi-ring generating $\mathcal{R}$.

We consider now the particular case where $\mathcal{X} = \mathbb{R}_0^+ \times \tilde{l}_\infty \setminus \{\tilde{\infty}\}$. In this space, we consider the ring $\mathcal{R}$ and its subclass of sets $\mathcal{I}$ defined by:

$$\mathcal{R} := \left\{ \bigcup_{\ell=1}^{m} J_\ell \times \tilde{A}_\ell \colon\ m \in \mathbb{N},\ J_\ell = [a_\ell, b_\ell),\ \text{and } \tilde{A}_\ell \in \tilde{\mathscr{R}} \right\}; \tag{C.5}$$

$$\mathcal{I} := \left\{ \bigcup_{\ell=1}^{m} J_\ell \times \tilde{A}_\ell \colon\ m \in \mathbb{N},\ J_\ell = [a_\ell, b_\ell),\ \text{and } \tilde{A}_\ell \in \tilde{\mathscr{J}} \right\}. \tag{C.6}$$

Lemma C.5. *The conclusion of Proposition* C.2 *holds if the condition are verified only for all bounded sets of* $\mathcal{I}$.

Proof. We focus on the condition (i), which is the strongest. The others follow easily as one can already guess from Remark C.4. The statement follows from the fact that $\tilde{\mathscr{J}}$ is closed for unions and then using the inclusion exclusion formula, one can write the probability involving any set $\tilde{A}_\ell \in \tilde{\mathscr{R}}$ using the probability of events involving only elements of $\tilde{\mathscr{J}}$.

For example, assume that $\tilde{A}_1 = \tilde{B}_1 \cap \tilde{D}_1$, where $\tilde{B}_1, \tilde{D}_1 \in \tilde{\mathscr{J}}$. Observe that

$$\begin{aligned}
\mathbb{P}(N(\cup_{\ell=1}^m J_\ell \times \tilde{A}_\ell) = 0) =& \mathbb{P}\left(\{N(J_1 \times (\tilde{B}_1 \cap \tilde{D}_1)) = 0\} \cap \{N(\cup_{\ell=1}^m J_\ell \times \tilde{A}_\ell) = 0\}\right) \\
=& \mathbb{P}\left(\{N(J_1 \times \tilde{B}_1) = 0\} \cap \{N(\cup_{\ell=1}^m J_\ell \times \tilde{A}_\ell) = 0\}\right) \\
&+ \mathbb{P}\left(\{N(J_1 \times \tilde{D}_1) = 0\} \cap \{N(\cup_{\ell=1}^m J_\ell \times \tilde{A}_\ell) = 0\}\right) \\
&- \mathbb{P}\left(\{N(J_1 \times (\tilde{B}_1 \cup \tilde{D}_1)) = 0\} \cap \{N(\cup_{\ell=1}^m J_\ell \times \tilde{A}_\ell) = 0\}\right),
\end{aligned}$$

which means that all events on the 3 last terms correspond to valued of the avoidance function over sets of $\mathcal{I}$, as we wanted. In the same way, if $\tilde{A}_1 = \tilde{B}_1 \setminus \tilde{D}_1$, for $\tilde{B}_1, \tilde{D}_1 \in \tilde{\mathscr{J}}$, for example, we could use the fact that $\mathbb{P}(N(\tilde{A}_1) = 0) = \mathbb{P}(N(\tilde{B}_1 \cup \tilde{D}_1) = 0) - \mathbb{P}(N(\tilde{D}_1) = 0)$ and obtain a similar formula for the avoidance function using only sets in $\mathcal{I}$ as we did above.

Hence, we have just shown that we can handle unions, intersections and exclusions of sets of $\tilde{\mathscr{J}}$. Noting that $\tilde{A}_\ell \in \tilde{\mathscr{R}}$ can always be written by using a finite number of these set operations involving elements of $\tilde{\mathscr{J}}$, the conclusion follows. □

Appendix D. Asymptotic independence in disjoint time intervals

We now state a series of auxiliary lemmas which will enable us to establish the second equality in Proposition 3.17. The first lemma, essentially says that in order to create a gap between the blocks we create an error which is the size of gap times the probability of having a rare event. We refer to [FFMa20, Lemma 3.5] for a proof.

Lemma D.1. *For any fixed $A \in \mathcal{B}_{\mathcal{V}}$ and $s, t', m \in \mathbb{N}$, we have:*

$$\left|\mathbb{P}(\mathscr{W}_{0,s+t'+m}(A)) - \mathbb{P}(\mathscr{W}_{0,s}(A) \cap \mathscr{W}_{s+t',s+t'+m}(A))\right| \leq t'\mathbb{P}(A).$$

The second lemma gives the error for estimating the probability of having at least one rare event inside a block by the sum of the probabilities of having one individual rare event.

Lemma D.2. *For any fixed $A \in \mathcal{B}_{\mathcal{V}}$ and integers s, t, m, we have:*

$$|\mathbb{P}(\mathscr{W}_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) - (1 - s\mathbb{P}(A))\mathbb{P}(\mathscr{W}_{0,m}(A))| \leq \left|s\mathbb{P}(A)\mathbb{P}(\mathscr{W}_{0,m}(A)) - \sum_{j=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{s+t-j,s+t-j+m}(A))\right| + s\mathbb{P}(A \cap \mathscr{W}^c_{1,s}(A)).$$

Proof. Note that the first term in the bound is measuring the mixing across a time gap of size t and the second term is measuring the probability that two events A cluster in the first block. Adding and subtracting and using the triangle inequality we obtain that

$$\begin{aligned}\left|\mathbb{P}(\mathscr{W}_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) - (1 - s\mathbb{P}(A))\mathbb{P}(\mathscr{W}_{0,m}(A))\right| \leq \\ \left|s\mathbb{P}(A)\mathbb{P}(\mathscr{W}_{0,m}(A)) - \sum_{j=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{s+t-j,s+t-j+m}(A))\right| + \\ + \left|\mathbb{P}(\mathscr{W}_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) - \mathbb{P}(\mathscr{W}_{0,m}(A)) + \sum_{j=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{s+t-j,s+t-j+m}(A))\right|. \quad \text{(D.1)}\end{aligned}$$

For the second term on the right, by stationarity, we have

$$\begin{aligned}\mathbb{P}(\mathscr{W}_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) &= \mathbb{P}(\mathscr{W}_{s+t,s+t+m}(A)) - \mathbb{P}(\mathscr{W}^c_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) \\ &= \mathbb{P}(\mathscr{W}_{0,m}(A)) - \mathbb{P}(\mathscr{W}^c_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)).\end{aligned}$$

Now, decomposing the event $\mathscr{W}_{0,s}^c(A) = \bigcup_{i=0}^{s-1} \sigma^{-i}(A)$ according to the last occurrence of the event A, we have

$$\begin{aligned}\mathbb{P}(\mathscr{W}_{0,s}^c(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) &= \sum_{i=0}^{s-1} \mathbb{P}(\sigma^{-i}(A) \cap \mathscr{W}_{i+1,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) \\ &= \sum_{i=0}^{s-1} \mathbb{P}(\sigma^{-i}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) - \sum_{i=0}^{s-1} \mathbb{P}(\sigma^{-i}(A) \cap \mathscr{W}_{i+1,s}^c(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) \\ &= \sum_{i=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{s+t-i,s+t-i+m}(A)) - \sum_{i=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{1,s-i}^c(A) \cap \mathscr{W}_{s+t-i,s+t-i+m}(A)).\end{aligned}$$

Therefore, we may write

$$0 \le \sum_{i=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{s+t-i,s+t-i+m}(A)) - \mathbb{P}(\mathscr{W}_{0,s}^c(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) \le \sum_{i=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{1,s-i}^c(A))$$

Hence, using these last three computations we get:

$$\begin{aligned}&\left|\mathbb{P}(\mathscr{W}_{0,s}(A) \cap \mathscr{W}_{s+t,s+t+m}(A)) - \mathbb{P}(\mathscr{W}_{0,m}(A)) + \sum_{j=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{s+t-j,s+t-j+m}(A))\right| \\ &\qquad \le \sum_{i=0}^{s-1} \mathbb{P}(A \cap \mathscr{W}_{1,s-i}^c(A)) \le s\mathbb{P}(A \cap \mathscr{W}_{1,s}^c(A))\end{aligned}$$

The result now follows from plugging the last estimate into (D.1). □

We are now ready to prove Proposition 3.17.

Proof of Proposition 3.17. We split time into blocks of of length $r_n := \lfloor n/k_n \rfloor$, $I_1 = [0, r_n)$, $I_2 = [r_n, 2r_n)$,..., $I_{k_n} = [(k_n - 1)r_n, k_n r_n)$, $I_{k_n+1} = [k_n r_n, n)$. Let $S_\ell = S_\ell(k_n)$ be the number of blocks I_j contained in $J_{n,\ell}$, *i.e.*, ,

$$S_\ell := \#\{i \in \{1, \dots, k_n\} : i \in \tilde{J}_{k_n,\ell}\} = \lceil (b_\ell - a_\ell)k_n \rceil.$$

For each $\ell = 1, \dots, m$, we define

$$\mathscr{A}_\ell := \bigcap_{i=\ell}^{m} \mathscr{W}_{J_{n,i}} \left(A_{n,i}^{(q_n)}\right).$$

Set $i_\ell := \min\{j \in \{1, \dots, k_n\} : I_j \subset J_{n,\ell}\} = \lceil k_n a_\ell \rceil$. Then $I_{i_\ell}, I_{i_\ell+1}, \dots, I_{i_\ell+S_\ell-1} \subset J_{n,\ell}$. Now, fix ℓ and for each $i \in \{i_\ell, \dots, i_\ell + S_\ell - 1\}$ let

$$\mathscr{B}_{i,\ell} := \bigcap_{j=i}^{i_\ell+S_\ell-1} \mathscr{W}_{I_j}(A_{n,\ell}^{(q_n)}),\ I_i^* := [(i-1)r_n, ir_n - t_n) \text{ and } I_i' := I_i \setminus I_i^*.$$

Note that $|I_i^*| = r_n - t_n =: r_n^*$ and $|I_i'| = t_n$.

Using Lemmas D.1 and D.2, we obtain

$$\Big|\mathbb{P}(\mathscr{B}_{i,\ell}\cap\mathscr{A}_{\ell+1})-(1-r_n\mathbb{P}(A_{n,\ell}^{(q_n)}))\mathbb{P}(\mathscr{B}_{i+1,\ell}\cap\mathscr{A}_{\ell+1})\Big|\le r_n^*\mathbb{P}(A_{n,\ell}^{(q_n)}\cap\mathscr{W}_{1,r_n^*}(A_{n,\ell}^{(q_n)}))+2t_n\mathbb{P}(A_{n,\ell}^{(q_n)})$$
$$+\sum_{j=0}^{r_n^*-1}\Big|\mathbb{P}(A_{n,\ell}^{(q_n)})\mathbb{P}(\mathscr{B}_{i+1,\ell}\cap\mathscr{A}_{\ell+1})-\mathbb{P}\left(\sigma^{-j-(i-1)r_n}A_{n,\ell}^{(q_n)}\cap\mathscr{B}_{i+1,\ell}\cap\mathscr{A}_{\ell+1}\right)\Big|,$$

Now using condition $Д_q^*(u_n)$, we obtain

$$\Big|\mathbb{P}(\mathscr{B}_{i,\ell}\cap\mathscr{A}_{\ell+1})-(1-r_n\mathbb{P}(A_{n,\ell}^{(q_n)}))\mathbb{P}(\mathscr{B}_{i+1,\ell}\cap\mathscr{A}_{\ell+1})\Big|\le\Upsilon_{k_n,n,\ell}$$

where

$$\Upsilon_{k_n,n,\ell}:=r_n^*\mathbb{P}(A_{n,\ell}^{(q_n)}\cap\mathscr{W}_{1,r_n^*}(A_{n,\ell}^{(q_n)}))+2t_n\mathbb{P}(A_{n,\ell}^{(q_n)})+r_n^*\gamma(q,n,t_n).$$

By (3.34), we may assume that n is sufficiently large so that by induction, we obtain:

$$\Big|\mathbb{P}(\mathscr{B}_{i_\ell,\ell}\cap\mathscr{A}_{\ell+1})-\left(1-r_n\mathbb{P}(A_{n,\ell}^{(q_n)})\right)^{S_\ell}\mathbb{P}(\mathscr{A}_{\ell+1})\Big|\le S_\ell\Upsilon_{k_n,n,\ell}.$$

and consequently, again by an induction argument,

$$\Big|\mathbb{P}(\mathscr{A}_1)-\prod_{\ell=1}^m\left(1-r_n\mathbb{P}(A_{n,\ell}^{(q_n)})\right)^{S_\ell}\Big|\le\sum_{\ell=1}^m S_\ell\Upsilon_{k_n,n,\ell}.$$

Since, for each $\ell=1,\dots,m$, we have $S_\ell=\lceil(b_\ell-a_\ell)k_n\rceil\sim k_n|J_\ell|$, it follows that, by (3.34), we have

$$\lim_{n\to+\infty}\prod_{\ell=1}^m\left(1-r_n\mathbb{P}(A_{n,\ell}^{(q_n)})\right)^{S_\ell}=\lim_{n\to+\infty}\prod_{\ell=1}^m\left(1-\left\lfloor\frac{n}{k_n}\right\rfloor\mathbb{P}(A_{n,\ell}^{(q_n)})\right)^{k_n|J_\ell|}=\prod_{\ell=1}^m\mathrm{e}^{-\nu(A_\ell)|J_\ell|}$$

To conclude the proof it suffices to show that $\lim_{n\to+\infty}\sum_{\ell=1}^m S_\ell\Upsilon_{k_n,n,\ell}=0$, which means that the following terms need to vanish

$$r_n^*\sum_{\ell=1}^m k_n|J_\ell|\mathbb{P}\left(A_{n,\ell}^{(q_n)}\cap\mathscr{W}_{1,r_n^*}(A_{n,\ell}^{(q_n)})\right)+2t_n\sum_{\ell=1}^m k_n|J_\ell|\mathbb{P}(A_{n,\ell}^{(q_n)})+r_n^*\sum_{\ell=1}^m k_n|J_\ell|\gamma(q,n,t_n).$$

The first term vanishes by $Д_q'(u_n)$. The second term also vanishes because, by (3.2), we have $k_nt_n=o(n)$ and, by (3.34), we have $n\mathbb{P}(A_{n,\ell}^{(q_n)})\to\nu(A_\ell)\ge 0$. The third term vanishes on account of condition $Д_q^*(u_n)$. □

Appendix E. Remaining proofs for the dynamical examples

Section 2.2 gave concrete dynamical examples where the theory in this paper holds. As noted in the discussion after Theorem 4.1, previous work shows that many of the required conditions hold. In this section we fill in the missing proofs, starting by showing that the piling process is well-defined for observables whose norm is maximised at periodic points. Note that the existence of the limit in (3.15) follows from a simpler version of the ideas here.

Recall that we are dealing with an acip μ and that for our periodic point ζ we are always assuming that the density exists and is bounded, say $\frac{d\mu}{\mathrm{Leb}}=D\in(0,\infty)$. For definiteness assume that we are dealing with a non-invertible map, the observable is as in (2.6), $\mathcal{M}=\{\zeta\}$ and ζ is a repelling fixed point (therefore $p=1$). In this case the process $(Y_j)_{j\in\mathbb{Z}}$ is such that,

for some $s \in \mathbb{N}_0$, we have $Y_j = \infty$, for all $j \leq -s$ and $Y_j = U \cdot \|(DT_\zeta)^j(\Theta)\|^d \frac{(DT_\zeta)^j(\Theta)}{\|(DT_\zeta)^j(\Theta)\|}$, for all $j > -s$, where U is a uniformly distributed random variable on the interval $[0,1]$ independent of Θ, which has a uniform distribution on $\mathbb{S}^{d-1}$ and $(DT)^i_\zeta$ denotes the i-fold product of the derivative of T at ζ (which is invertible since we are assuming that ζ is repelling).

This clearly satisfies (2)–(4) for the piling process, so we need to show (1), i.e.,

$$\mathcal{L}\left(\frac{1}{\tau}\mathbb{X}_n^{r_n+s,r_n+t} \,\middle|\, \|\mathbf{X}_{r_n}\| > u_n(\tau)\right) \xrightarrow[n\to\infty]{} \mathcal{L}\left((Y_j)_{j=s,\dots,t}\right), \text{ for all } s < t \in \mathbb{Z} \text{ and all } \tau > 0.$$

Without loss of generality we assume that g is positive in a neighbourhood of ζ, so $\|\mathbf{X}_{r_n}(x)\| > u_n(\tau)$ can be written $g\left(\operatorname{dist}(T^{r_n}(x),\zeta)\right) > u_n(\tau)$. From (2.14) we see that asymptotically $u_n(\tau) \sim g\left(\left(\frac{\tau}{Cn}\right)^{1/d}\right)$, where $C = DL_\zeta$ for $L_\zeta = \lim_{r\to 0}\frac{\operatorname{Leb}(B_r(\zeta))}{r}$ (this exists in the cases considered here). In this section we will assume that this is an equality since all our estimates are asymptotic in n, so similarly $u_n^{-1}(z) = Cn\left(g^{-1}(z)\right)^d$, where we will also assume this is well-defined. Since we are assuming that $\|\mathbf{X}_{r_n}(x)\| = u_n(\tau)$, let $v < 1$ be such that $\mathbf{X}_{r_n}(x) = u_n(v\tau)$, which translates as $\operatorname{dist}(T^{r_n}(x),\zeta) = \left(\frac{v\tau}{Cn}\right)^{1/d}$. Using the linearisation domain around the fixed point ζ from Hartman-Grobman theory, $\operatorname{dist}(T^{r_n+k}(x),\zeta) \sim \left(\frac{v\tau}{Cn}\right)^{1/d}\|(DT_\zeta)^k(w)\|$, where $w = \frac{\Phi_\zeta(T^{r_n}(x))}{\|\Phi_\zeta(T^{r_n}(x))\|}$ and Φ_ζ is as in (2.6). It follows that $\mathbf{X}_{r_n+k}(x) \sim g\left(\left(\frac{v\tau}{Cn}\right)^{1/d}\|(DT_\zeta)^k(w)\|\right)\frac{(DT_\zeta)^k(w)}{\|(DT_\zeta)^k(w)\|}$, from which we find

$$\frac{u_n^{-1}\left(\mathbf{X}_{r_n+k}(x)\right)}{\tau}\frac{\mathbf{X}_{r_n+k}(x)}{\|\mathbf{X}_{r_n+k}(x)\|} \sim v\|(DT_\zeta)^k(w)\|^d\frac{(DT_\zeta)^k(w)}{\|(DT_\zeta)^k(w)\|}.$$

All of these asymptotics become equalities as $n \to \infty$. The fact that $T^{r_n}(x)$ is chosen according to the invariant measure μ on the small neighbourhood $B_{g^{-1}(u_n(\tau))}(\zeta) = \{\|X_{r_n}(x)\| < u_n(\tau)\}$ of ζ and the fact that μ behaves asymptotically like Lebesgue measure on small neighbourhoods of ζ give that the finite Y_js are indeed of the form required. The infinite terms appear because, due to the repelling nature of ζ, the entrance of the orbit of x at time r_n in $B_{g^{-1}(u_n(\tau))}(\zeta)$ can only be preceded by a finite number, $s \in \mathbb{N}_0$, of consecutive hits to $B_{g^{-1}(u_n(\tau))}(\zeta)$. But then $T^{r_n-s-1}(x)$, for example, must belong to neighbourhood of $T^{-1}(\zeta)\setminus\{\zeta\}$, which are at a fixed distance from ζ: let $c = \operatorname{dist}(\zeta, T^{-1}(\zeta)\setminus\{\zeta\}) > 0$. Then for n sufficiently large we have $\operatorname{dist}(T^{r_n-s-1}(x),\zeta) > c/2$ and hence $u_n^{-1}(\|X_{r_n-s-1}\|) \sim Cn(c/2)^d$, which clearly diverges to ∞ as $n \to \infty$.

In the invertible case, for the toral diffeomorphisms considered, we have $T^{-1}(\zeta) = \zeta$ and therefore for all $j \in \mathbb{Z}$, we have $Y_j = U \cdot \|(DT_\zeta)^j(\Theta)\|^d\frac{(DT_\zeta)^j(\Theta)}{\|(DT_\zeta)^j(\Theta)\|}$, where U and Θ are as above.

We observe that if instead ζ were a periodic point of period $q > 1$ then elements of Y would be the same as for the fixed point case, but with $q-1$ ∞ terms between the entries qk and $q(k+1)$, for all $k \in \mathbb{Z}$.

Now that we understand $(Y_j)_j$ it is easy to see that the condition

$$\mathbb{P}(\text{all finite } Q_j\text{'s are mutually different}) = 1$$

in Theorem 2.11 is also trivially satisfied in these examples as it can be interpreted as ζ being hyperbolic.

In the one dimensional case, concerning most of the examples given in the noninvertible scenario in Section 2.2, if we consider ζ as a repelling fixed point such that $|DT_\zeta| = \chi > 1$ then, in this case, $(Y_j)_j$ assumes the simpler form:

$$(\dots, \infty, \infty, U\chi^{-s}, \dots, U\chi^{-1}, U, U\chi, U\chi^2, \dots)$$

Finally, to obtain the piling process we need to condition the Y_j on the event $\inf_{j\le -1} \|Y_j\| \ge 1$, which is to say that in the computation of the Y_j, the exceedance occurring at time r_n is indeed the first exceedance of the cluster, which means $s = 0$ and therefore $(Z_j)_j$ is of the form

$$(\dots, \infty, \infty, U, U\chi, U\chi^2, \dots)$$

It remains to show that the Manneville-Pomeau case satisfies our conditions. We first show $Д'_{q_n}$. We use the same technique as in [FFTV16], which is to employ [HSV99, Lemma 3.5] which estimates for a set U,

$$\sum_{j=1}^{N} \frac{1}{\mu(U)} \mu(U \cap T_\gamma^{-j}(U)) \le N \sup_{j=\tau_U,\dots,N} \sup_U \frac{P^j(\mathbb{1}_U h)}{h}$$

where h is the density and P is the transfer operator and τ_U is the first return time of U to itself. For the next estimates, the argument in [HSV99] required that the set U is a cylinder set, but in fact the distortion properties follow from the iterates of interest remaining a bounded distance from 0. Hence the same methods will work for a small Borel set in the linearisation domain of z. Using the facts that $Ph = h$ and P is a positive operator along with some distortion estimates,

$$\frac{P^j(\mathbb{1}_U h)}{h} \lesssim \frac{1}{\inf h} \frac{\mu(U)}{\mathrm{Leb}(T_\gamma^{\tau_U} U)}$$

where Leb is Lebesgue. Since $\inf h > 0$ we can ignore the first term.

In our setting this translates to an estimate of $Д'_{q_n}$ becoming:

$$n \sum_{j=\max\{q_n,\tau'_n\}}^{r_n} \mu(U_n^{(q_n)} \cap T_\gamma^{-j} U_n) \lesssim r_n \frac{\mu(U_n^{(q_n)})}{\mathrm{Leb}(T_\gamma^{\tau'_n} U_n^{(q_n)})},$$

where τ'_n is the smallest $j \ge 1$ with $U_n^{(q_n)} \cap T_\gamma^{-j} U_n \ne \emptyset$.

We will suppose that ζ is a periodic point of period q and $|DT_\gamma^q(\zeta)| > 1$. We set $q_n = q$ for all $n \in \mathbb{N}$. Hartman-Grobman theory around a repelling periodic point says that before the set U_n can return to hit $U_n^{(q_n)}$ it must have (after $\asymp \log n$ iterates) reached the scale of the linearisation domain, which we can treat as order 1. Hence $\mathrm{Leb}(T_\gamma^{\tau'_n} U_n^{(q_n)}) \gtrsim 1$, and we find, since also $\mu(U_n^{(q_n)}) \sim \tau/n$,

$$n \sum_{j=\tau'_n}^{r_n} \mu(U_n^{(q_n)} \cap T_\gamma^{-j} U_n) \lesssim \frac{r_n}{n}.$$

Therefore we can obtain our required estimates with $r_n = o(n)$. Adding this fact to [FFTV16, (19)] we find that the requirement on γ so that both $Д_{q_n}$ and $Д'_{q_n}$ hold is $3 + \gamma + 2\gamma^2 - \frac{1}{\gamma} < 0$, eg $\gamma \in (0, 0.289)$.

As final remark, we recall that the theory developed here applies to higher dimensional cases, where $d > 1$. We include here a finite time sample path generated by a uniformly expanding 2-dimensional system covered by Theorem 2.5.

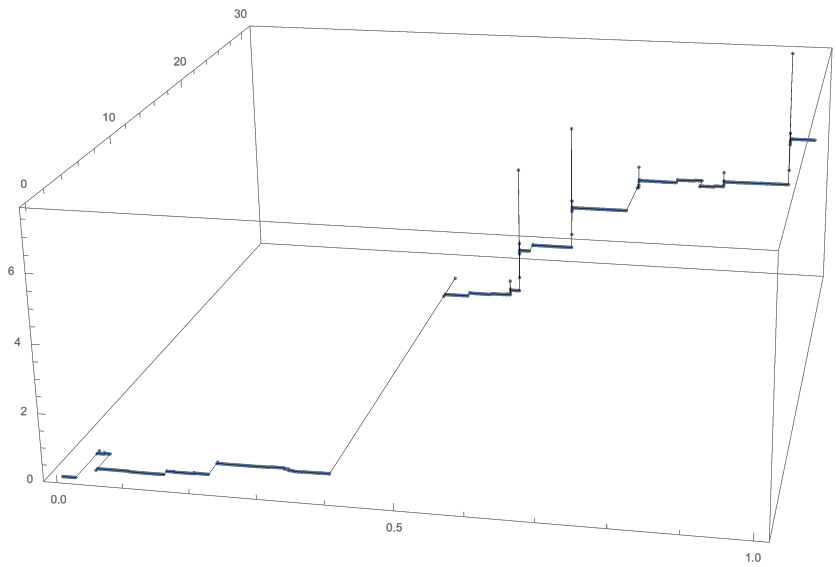

FIGURE E.1. Plot of finite sample simulation of $S_n(t)$ with $n = 5000$, where $X_j = \psi \circ T^j(x, y)$, where $T(x, y) = (3x \mod 1, 3y/2 \mod 1)$, $\psi(x, y) = (|x - 1/8|^{-2} - |x - 3/8|^{-2}, |y - 8/19|^{-2} - 2|y - 12/19|^{-2} + |y - 12/13|^{-2})$. Here we have completed the graph in order to make the jumps more visible. Note that $\zeta = 8/19$ is a periodic point of period 3 of the map $3y/2 \mod 1$ and $12/19, 12/13$ are its orbital points.

REFERENCES

[AFF20] M. Abadi, A.C.M. Freitas, and J.M. Freitas, *Dynamical counterexamples regarding the extremal index and the mean of the limiting cluster size distribution*, J. Lond. Math. Soc. (2) **102** (2020), no. 2, 670–694.

[AFFR16] Davide Azevedo, Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Fagner B. Rodrigues, *Clustering of extreme events created by multiple correlated maxima*, Phys. D **315** (2016), 33–48. MR 3426918

[AFFR17] ———, *Extreme Value Laws for Dynamical Systems with Countable Extremal Sets*, J. Stat. Phys. **167** (2017), no. 5, 1244–1261. MR 3647060

[AFLV11] José F. Alves, Jorge M. Freitas, Stefano Luzzatto, and Sandro Vaienti, *From rates of mixing to recurrence times via large deviations*, Adv. Math. **228** (2011), no. 2, 1203–1236. MR 2822221 (2012g:37049)

[Bil99] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication. MR MR1700749 (2000e:60008)

[BKS12] Bojan Basrak, Danijel Krizmanić, and Johan Segers, *A functional limit theorem for dependent sequences with infinite variance stable limits*, Ann. Probab. **40** (2012), no. 5, 2008–2033. MR 3025708

[BM08] Péter Bálint and Ian Melbourne, *Decay of correlations and invariance principles for dispersing billiards with cusps, and related planar billiard flows*, J. Stat. Phys. **133** (2008), no. 3, 435–447. MR 2448631

[BPS18] Bojan Basrak, Hrvoje Planinić, and Philippe Soulier, *An invariance principle for sums and record times of regularly varying stationary sequences*, Probab. Theory Related Fields **172** (2018), no. 3-4, 869–914. MR 3877549

[BS09] Bojan Basrak and Johan Segers, *Regularly varying multivariate time series*, Stochastic Process. Appl. **119** (2009), no. 4, 1055–1080. MR 2508565

[CFF+15] Maria Carvalho, Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, Mark Holland, and Matthew Nicol, *Extremal dichotomy for uniformly hyperbolic systems*, Dyn. Syst. **30** (2015), no. 4, 383–403. MR 3430306

[CHN19] Meagan Carney, Mark Holland, and Matthew Nicol, *Extremes and extremal indices for level set observables on hyperbolic systems*, Preprint arXiv:1909.04748, 2019.

[Dav83] Richard A. Davis, *Stable limits for partial sums of dependent random variables*, Ann. Probab. **11** (1983), no. 2, 262–269. MR 690127 (84e:60028)

[DH95] Richard A. Davis and Tailen Hsing, *Point process and partial sum convergence for weakly dependent random variables with infinite variance*, Ann. Probab. **23** (1995), no. 2, 879–917. MR 1334176 (96g:60031)

[DP84] Manfred Denker and Walter Philipp, *Approximation by Brownian motion for Gibbs measures and flows under a function*, Ergodic Theory Dynam. Systems **4** (1984), no. 4, 541–552. MR 779712

[DVJ03] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes. Vol. I*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2003, Elementary theory and methods. MR 1950431 (2004c:60001)

[DVJ08] ———, *An introduction to the theory of point processes. Vol. II*, second ed., Probability and its Applications (New York), Springer, New York, 2008, General theory and structure. MR 2371524 (2009b:60150)

[Dwa64] Meyer Dwass, *Extremal processes*, Ann. Math. Statist **35** (1964), 1718–1725. MR 0177440

[FF08] Ana Cristina Moreira Freitas and Jorge Milhazes Freitas, *Extreme values for Benedicks-Carleson quadratic maps*, Ergodic Theory Dynam. Systems **28** (2008), no. 4, 1117–1133. MR 2437222 (2010a:37013)

[FFMa18] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Mário Magalhães, *Convergence of marked point processes of excesses for dynamical systems*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 9, 2131–2179. MR 3836843

[FFMa20] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Mário Magalhães, *Complete convergence and records for dynamically generated stochastic processes*, Trans. Amer. Math. Soc. **373** (2020), no. 1, 435–478. MR 4042881

[FFRS20] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, Fagner B. Rodrigues, and Jorge Valentim Soares, *Rare events for Cantor target sets*, Comm. Math. Phys. **378** (2020), no. 1, 75–115. MR 4124982

[FFT12] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Mike Todd, *The extremal index, hitting time statistics and periodicity*, Adv. Math. **231** (2012), no. 5, 2626–2665. MR 2970462

[FFT13] ———, *The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics*, Comm. Math. Phys. **321** (2013), 483–527.

[FFT15] ———, *Speed of convergence for laws of rare events and escape rates*, Stochastic Process. Appl. **125** (2015), no. 4, 1653–1687. MR 3310360

[FFTV16] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, Mike Todd, and Sandro Vaienti, *Rare events for the Manneville-Pomeau map*, Stochastic Processes and their Applications **126** (2016), no. 11, 3463–3479.

[FGGV18] D. Faranda, H. Ghoudi, P. Guiraud, and S. Vaienti, *Extreme value theory for synchronization of coupled map lattices*, Nonlinearity **31** (2018), no. 7, 3326–3358. MR 3816758

[FMT03] Michael Field, Ian Melbourne, and Andrew Török, *Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions*, Ergodic Theory Dynam. Systems **23** (2003), no. 1, 87–110. MR 1971198

[Gou04] Sébastien Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probab. Theory Related Fields **128** (2004), no. 1, 82–122. MR 2027296

[Gou08] ———, *Stable laws for the doubling map*, Unpublished notes available at https://www.math.sciences.univ-nantes.fr/ gouezel/articles/DoublingStable.pdf, 2008.

[Gou10] ———, *Almost sure invariance principle for dynamical systems by spectral methods*, Ann. Probab. **38** (2010), no. 4, 1639–1671. MR 2663640

[HK82] Franz Hofbauer and Gerhard Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. **180** (1982), no. 1, 119–140. MR 656227

[HM07] Mark Holland and Ian Melbourne, *Central limit theorems and invariance principles for Lorenz attractors*, J. Lond. Math. Soc. (2) **76** (2007), no. 2, 345–364. MR 2363420

[Hsi87] Tailen Hsing, *On the characterization of certain point processes*, Stochastic Process. Appl. **26** (1987), no. 2, 297–316. MR 923111

[HSV99] Masaki Hirata, Benoît Saussol, and Sandro Vaienti, *Statistics of return times: a general framework and new applications*, Comm. Math. Phys. **206** (1999), no. 1, 33–55. MR 1736991 (2001c:37007)

[HT19] Mark Holland and Mike Todd, *Weak convergence to extremal processes and record events for nonuniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **39** (2019), no. 4, 980–1001. MR 3916843

[JMP+19] Paul Jung, Ian Melbourne, Françoise Pène, Paulo Varandas, and Hong-Kun Zhang, *Necessary and sufficient condition for $\mathcal{M}_2$-convergence to a lévy process for billiards with cusps at flat points*, Preprint arXiv:1902.08958, 2019.

[JPZ20] Paul Jung, Françoise Pène, and Hong-Kun Zhang, *Convergence to α-stable lévy motion for chaotic billiards with several cusps at flat points*, Nonlinearity **33** (2020), no. 2, 807–839.

[Lam64] John Lamperti, *On extreme order statistics*, Ann. Math. Statist **35** (1964), 1726–1737. MR 0170371

[Lea83] M. R. Leadbetter, *Extremes and local dependence in stationary sequences*, Z. Wahrsch. Verw. Gebiete **65** (1983), no. 2, 291–306. MR MR722133 (85b:60033)

[LFF+16] Valerio Lucarini, Davide Faranda, Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, Mark Holland, Tobias Kuna, Matthew Nicol, Mike Todd, and Sandro Vaienti, *Extremes and recurrence in dynamical systems*, Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts, Wiley, Hoboken, NJ, 2016.

[LLR83] M. R. Leadbetter, Georg Lindgren, and Holger Rootzén, *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York, 1983. MR MR691492 (84h:60050)

[LSV99] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems **19** (1999), no. 3, 671–685. MR MR1695915 (2000d:37029)

[LWZ81] Raoul LePage, Michael Woodroofe, and Joel Zinn, *Convergence to a stable distribution via order statistics*, Ann. Probab. **9** (1981), no. 4, 624–632. MR 624688 (82k:60049)

[MN05] Ian Melbourne and Matthew Nicol, *Almost sure invariance principle for nonuniformly hyperbolic systems*, Comm. Math. Phys. **260** (2005), no. 1, 131–146. MR 2175992

[MN09] ———, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, Ann. Probab. **37** (2009), no. 2, 478–505. MR 2510014

[Mor77] Toshio Mori, *Limit distributions of two-dimensional point processes generated by strong-mixing sequences*, Yokohama Math. J. **25** (1977), no. 2, 155–168. MR 0467887

[MP16] Giorgio Mantica and Luca Perotti, *Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis*, J. Phys. A **49** (2016), no. 37, 374001, 21. MR 3544737

[MT02] Ian Melbourne and Andrei Török, *Central limit theorems and invariance principles for time-one maps of hyperbolic flows*, Comm. Math. Phys. **229** (2002), no. 1, 57–71. MR 1917674

[MV20] Ian Melbourne and Paulo Varandas, *Convergence to a Lévy process in the Skorohod $\mathcal{M}_1$ and $\mathcal{M}_2$ topologies for nonuniformly hyperbolic systems, including billiards with cusps*, Comm. Math. Phys. **375** (2020), no. 1, 653–678. MR 4082176

[MZ15] Ian Melbourne and Roland Zweimüller, *Weak convergence to stable Lévy processes for nonuniformly hyperbolic dynamical systems*, Ann. Inst. Henri Poincaré Probab. Stat. **51** (2015), no. 2, 545–556. MR 3335015

[Nov02] S. Y. Novak, *Multilevel clustering of extremes*, Stochastic Process. Appl. **97** (2002), no. 1, 59–75. MR 1870720

[O'B87] George L. O'Brien, *Extreme values for stationary and Markov sequences*, Ann. Probab. **15** (1987), no. 1, 281–291. MR MR877604 (88e:60042)

[Pic71] James Pickands, III, *The two-dimensional Poisson process and extremal processes*, J. Appl. Probability **8** (1971), 745–756. MR 0295453

[PM80] Yves Pomeau and Paul Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Comm. Math. Phys. **74** (1980), no. 2, 189–197. MR 576270 (81g:58024)

[PS20] Françoise Pène and Benoît Saussol, *Spatio-temporal poisson processes for visits to small sets*, to appear in Israel J. Math. (arXiv:1803.06865), 2020.

[Res75] Sidney I. Resnick, *Weak convergence to extremal processes*, Ann. Probability **3** (1975), no. 6, 951–960. MR 0428396

[Res87] ———, *Extreme values, regular variation, and point processes*, Applied Probability. A Series of the Applied Probability Trust, vol. 4, Springer-Verlag, New York, 1987. MR 900810 (89b:60241)

[Sat13] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 2013, Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR 3185174

[Seg05] Johan Segers, *Approximate distributions of clusters of extremes*, Statist. Probab. Lett. **74** (2005), no. 4, 330–336. MR 2186477

[Smi92] Richard L. Smith, *The extremal index for a Markov chain*, J. Appl. Probab. **29** (1992), no. 1, 37–45. MR 1147765

[TK10a] Marta Tyran-Kamińska, *Convergence to Lévy stable processes under some weak dependence conditions*, Stochastic Process. Appl. **120** (2010), no. 9, 1629–1650. MR 2673968

[TK10b] ———, *Weak convergence to Lévy stable processes in dynamical systems*, Stoch. Dyn. **10** (2010), no. 2, 263–289. MR 2652889

[Whi02] Ward Whitt, *Stochastic-process limits*, Springer Series in Operations Research, Springer-Verlag, New York, 2002, An introduction to stochastic-process limits and their application to queues. MR 1876437

[You98] Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650. MR MR1637655 (99h:58140)

[You99] ———, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188. MR MR1750438 (2001j:37062)

Ana Cristina Moreira Freitas, Centro de Matemática & Faculdade de Economia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal

Email address: amoreira@fep.up.pt

URL: http://www.fep.up.pt/docentes/amoreira/

Jorge Milhazes Freitas, Centro de Matemática & Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal

Email address: jmfreita@fc.up.pt

URL: http://www.fc.up.pt/pessoas/jmfreita/

Mike Todd, Mathematical Institute, University of St Andrews, North Haugh, St Andrews, KY16 9SS, Scotland,

Email address: m.todd@st-andrews.ac.uk

URL: http://www.mcs.st-and.ac.uk/~miket/