

ON RECORD BEHAVIOURS OF RETURNS FOR INTERVAL MAPS

BOYUAN ZHAO

ABSTRACT. Given a dynamical system $(\mathcal{X}, T)$ and an ergodic invariant measure μ , a μ -typical point $x \in \mathcal{X}$ under T visits arbitrarily small neighbourhood of x . We will use techniques based on extreme value theory and Laplace functional for point process to show that, for interval maps with suitable mixing and anti-clustering conditions, the limiting distributions of the record times and values of the return process $Y_j(x) = -\log d(T^j x, x)$, after possible normalisation, has Poisson-type laws.

1. INTRODUCTION

Let $(\mathcal{X}, T, \mu)$ be a dynamical system defined on the metric space $(\mathcal{X}, d)$, and μ is a T -invariant ergodic probability measure. For a μ -typical $x \in \mathcal{X}$, one of the most fundamental theorems in dynamics, namely the *Poincaré recurrence theorem*, asserts that for μ -almost every $x \in \mathcal{X}$,

$$\liminf_{n \rightarrow \infty} d(T^n x, x) = 0.$$

An alternative interpretation is, for a μ -typical point, there exists a subsequence $(n_k)_k$ such that $d(T^{n_k} x, x)$ is strictly decreasing. A rich research area of dynamics is devoted to establishing various quantitative limit laws in relation to recurrences, either regarding the return time or the return distance (see, [BS], [Col], [Sou1, Sou2], [KKP], [HKKP], [GHN], *etc.*) obtaining results in various forms: almost sure convergences, dynamical Borel-Cantelli dichotomy results, fractal dimension-related formulae, extreme value theory etc. In this paper, we are interested in the behaviour of the record value of return distances to the initial position x . More precisely, let

$$Y_j(x) := -\log d(x, T^j x). \quad (1.1)$$

This observable is well-defined for almost every x non-periodic and all $j \in \mathbb{N}$. Set $\tau_1 = 1$ and inductively define the k -th record time $\tau_k := \inf \{m > \tau_{k-1} : Y_m > Y_{\tau_{k-1}}\}$. We will develop a useful approach to determine the behaviour of Y_{τ_k} . In particular, we will use results based on Extreme Value Theory to prove that the distribution of τ_k and Y_{τ_k} follow a particular limit law if the system is fast-mixing and satisfies some other conditions regarding clustering short returns (see Theorem 2.2 below). More precisely, under suitable conditions, we prove that

Theorem 1.1. $\mathbb{P}(Y_{\tau_k} - \log n \leq r) \rightarrow \int_0^r \frac{z^{k-1}}{(k-1)!} e^{-z} dz$. *Equivalently, the k -th record after normalisation has a Gamma distribution.*

To guide us on why we should expect such form of the limiting distribution, let us first look at the *i.i.d.* case, since an observable on a fast mixing dynamical system behaves like a chain of *i.i.d.* random variables after relatively few iterates.

2020 *Mathematics Subject Classification.* 60G70, 37E05, 60G55, 37A25.

Key words and phrases. Return Process, Extreme Value Theory, Record Statistics, Interval Maps, Point Measures.

1.1. Motivation: independent with common distribution case. Let $\{X_i\}_{i=0,1,\dots}$ be a sequence of *i.i.d.* random variables taking values in $[0, 1]$ with common distribution function F and density f . Inheriting the definition above, *i.e.* $Y_j = -\log d(X_0, X_j)$, we observe the following:

Condition on the event $X_0 = x$, then $(Y_j)_{j \geq 1}$ is *i.i.d.* and has distribution

$$\begin{aligned} G_x(y) &:= \mathbb{P}(Y_j \leq y) = \mathbb{P}(|X_j - x| \geq e^{-y}) = 1 - (F(x + e^{-y}) - F(x - e^{-y})) \\ &= 1 - \int_{x-e^{-y}}^{x+e^{-y}} f(t) dt \end{aligned}$$

So Y_j has density $g_x(y) = \frac{d}{dy} G_x(y) = e^{-y} [f(x + e^{-y}) + f(x - e^{-y})]$.

By, for example [Kar, HS, R1], using the hazard function,

$$\mathbb{P}(Y_{\tau_k} \leq y | X_0 = x) = \int_{-\infty}^y \frac{(-\log(1 - G_x(z)))^{k-1}}{(k-1)!} dG_x(z) \quad (1.2)$$

Note that $(F(x + e^{-z}) - F(x - e^{-z})) = \int_{x-e^{-z}}^{x+e^{-z}} f(\xi) d\xi$ is comparable with the probability of visiting an r -neighbourhood around x with $r = e^{-z}$ (in dynamical system context, for a system $(\mathcal{X}, T, \mu)$ this is simply $\mu(B(x, e^{-z}))$). By the Lebesgue differentiation theorem, $F(x + e^{-z}) - F(x - e^{-z}) \approx 2f(x)e^{-z}$, $G_x(y) \approx 1 - 2f(x)e^{-y}$ and $Y_{\tau_k} - \log(2f(x))$ is asymptotically $\Gamma(k, 1)$ (actually it is the special case of Gamma distribution, known as Erlang, since $k \in \mathbb{N}$) distributed, so, integrate over x according to X_0 ,

$$\mathbb{P}(Y_{\tau_k} - \log(2f(x)) \leq y) \approx \int_{-\infty}^{\infty} \int_0^y \frac{t^{k-1}}{(k-1)!} e^{-t} f(x) dt dx.$$

The proof of records for *i.i.d.* random variables follows from proving that the k -th record has the same distribution of the sum of k number of *i.i.d.* random variables with the exact same distribution, plus the fact that the independent increments between record has exponential distribution (see e.g. [R1]). We can establish the result via extreme value theory and point process in a rigorous way.

In classical extreme value theory, given a sequence of *i.i.d.* random variables $\{\xi_i\}_i$ with common distribution F , if for each $\tau \geq 0$, there is $u_n(\tau)$ a sequence of numbers such that

$$n(1 - F(u_n(\tau))) \rightarrow \tau \text{ as } n \rightarrow \infty, \quad (1.3)$$

then (see for example [LLR]) the partial maxima $M_n = \max_{i \leq n} \xi_i$

$$\mathbb{P}(M_n \leq u_n(\tau)) \rightarrow e^{-\tau} \text{ in distribution as } n \rightarrow \infty. \quad (1.4)$$

Thus, coming back to our previous question, again conditioning on $X_0 = x$ again and set $u_n(\tau) = \tau + \log n$,

$$\begin{aligned} n(1 - G_x(u_n)) &= n \int_{x-e^{-\tau/n}}^{x+e^{-\tau/n}} f(t) dt = \frac{n2e^{-\tau}}{|B(x, e^{-\tau/n})|} \int_{B(x, e^{-\tau/n})} f(t) dt \\ &\xrightarrow{n \rightarrow \infty} 2e^{-\tau} f(x), \end{aligned}$$

where the last line holds almost surely by the Lebesgue differentiation theorem. This means, conditioning on $X_0 = x$, the process

$$M_n(t) = \begin{cases} \max_{j \in [1, [tn]]} Y_j - \log n & 1/n < t \leq 1, \\ Y_1 - \log n & 0 < t \leq 1/n. \end{cases}$$

converges to an extremal process $Y(t)$ generated by the distribution $H_x(y) = e^{-2f(x)e^{-y}}$ (see [RR]). So integrating over the conditional event $\{X_0 = x\}$,

$$\mathbb{P} \left(\max_{1 \leq j \leq n} Y_j \leq \tau + \log n \right) \rightarrow \int e^{-2f(x)e^{-\tau}} f(x) dx. \quad (1.5)$$

Remark 1.2. In classical extreme value theory the observable Y_j may be defined more generally by $\varphi(d(x, X_j))$ where φ may take one of the three form, such their normalised partial maximum converges weakly to Gumbel, Fréchet and Weibull distribution, respectively:

- $\varphi(t) = -\log(t)$, then $M_n(t) = M_{[nt]} - \log n$ converges weakly to an extremal process of Gumbel type;
- $\varphi(t) = t^{-\alpha}$ for some $\alpha > 0$, then $M_n(t) = n^{-\alpha} M_{[nt]}$ converges to an extremal process of Fréchet type;
- $\varphi(t) = C - t^\alpha$ for some $C \in \mathbb{R}$ and $\alpha > 0$, then $M_n(t) = n^\alpha (M_{[nt]} - C)$ converges weakly to an extremal process of Weibull type

For the scope of this paper, we will concentrate on the case corresponding to the Gumbel-type distribution.

Lemma 1.3. Let Z be a Poisson random variable with mean λ , then

$$\mathbb{P}(Z \geq k) = \int_0^\lambda \frac{1}{(k-1)!} t^{k-1} e^{-t} dt = \mathbb{P}(\Gamma(k, 1) \leq \lambda)$$

where $\Gamma(k, \alpha)$ is a Gamma distributed random variable with parameters k and α .

Proof. Let $g(\lambda) := \mathbb{P}(Z \geq k) = e^{-\lambda} \sum_{j=k}^\infty \frac{1}{j!} \lambda^j$. It is clear that $g(0) = 0$, and g is differentiable in λ for all $k \geq 1$, hence $g(\lambda) = \int_0^\lambda g'(t) dt$ where

$$g'(\lambda) = \frac{d}{d\lambda} e^{-\lambda} \sum_{j=k}^\infty \frac{1}{j!} \lambda^j = -e^{-\lambda} \sum_{j=k}^\infty \frac{1}{j!} \lambda^j + e^{-\lambda} \sum_{j=k}^\infty \frac{1}{(j-1)!} \lambda^{j-1} = \frac{1}{(k-1)!} e^{-\lambda} \lambda^{k-1}. \quad \square$$

Recall the definitions of τ_k , conditioning on $X_0 = x$ again, we define the counting measures

$$R_n = \sum_{k=1}^\infty \delta_{\tau_k/n}, \quad V_n := \sum_{k=1}^\infty \delta_{\{(Y_{\tau_k} - \log n)/2f(x)\}}$$

By [RR, Theorem 4, Theorem 5], [R2, Corollary 4.22, Corollary 4.23] or [LLR, Theorem 5.7.1, Theorem 5.8.1],

$$R_n \Rightarrow \mathcal{R} \quad \& \quad V_n \Rightarrow \mathcal{V} \quad \text{both in } M_p(0, \infty), \quad (1.6)$$

where $\mathcal{R}$ is a Poisson Random Measure (PRM) with intensity measure, $t^{-1} dt$, and $\mathcal{V}$ a PRM with intensity determined by $S(y) = y$. Recall the Gamma density function, $\gamma_{k,1}(t) = \frac{1}{\Gamma(k)} t^{k-1} e^{-t}$ where $\Gamma(k) = (k-1)!$, then for all $y > 0$:

$$\begin{aligned} \mathbb{P}((Y_{\tau_k} - \log n)/2f(x) \leq y) &= \mathbb{P}(V_n(0, y] \geq k) \rightarrow \mathbb{P}(\text{Poi}(y) \geq k) \\ &= \int_0^y \frac{1}{\Gamma(k)} t^{k-1} e^{-t} dt = \mathbb{P}(\Gamma(k, 1) \leq y), \end{aligned}$$

In the above formula we suppressed the dependence on $X_0 = x$, so after averaging over $X_0 = x$,

$$\mathbb{P} \left(d(X_0, X_{\tau_k}) \geq e^{-2f(x)y/n} \right) \rightarrow \int \mathbb{P}(\Gamma(k, 1) \leq y) f(x) dx \quad (1.7)$$

The above result is comparable with (1.2), with a caveat that τ_k represents the k -th record above the threshold $\log n$. Moreover, from (1.7) one deduces that due to the conditioning on initial condition X_0 , the distribution of the record values will be averaged according to the density f (in fact, X_0 needs not to have the same distribution as the other X_i 's as long as X_0 is defined on the same probability space), which is similar to introducing another level of stochasticity. We can interpret this via the perspective of a *Cox process*.

1.2. Cox process. In this subsection we show that the two-dimensional point process associated to the return value process (1.1) defined for *i.i.d.* random variables form a *Cox process*, which we describe below.

Definition 1.4. Let E be a metric space. A *Cox process* on E is a point process on E such that its distribution is a mixture of Poisson processes. In particular, its **random intensity measure** Λ defined on the state space E (e.g. $\mathbb{R}^d$ or $[0, \infty)$) is given by the following:

- a random measure Λ is drawn according to some distribution.
- Then, conditional on $\Lambda = \lambda$, the point process N is an inhomogeneous Poisson process with intensity measure λ .

In simple terms, a Cox process is a doubly stochastic Poisson point process. It was first properly described in (hence named after the author of) [Cox], and is well studied for a wide range of applied problems including biomathematics and flock modelling. We refer to [LP, §13] for a thorough discussion of its various aspects of statistical properties. The Cox process is also particularly relevant for our return statistics governed by dynamical systems: the point process for a hitting to a small neighbourhood of a fixed initial condition converges weakly to a Poisson process (see §1.3 below), then heuristically, varying the initial position causes the ‘double’ stochasticity.

The **Laplace functional** of a Cox process is

$$\mathbb{E} \left[e^{-\int g dN} \right] = \mathbb{E} \left[\exp \left(- \int_E (1 - e^{-g(z)}) \Lambda_x(dz) \right) \right], \quad (1.8)$$

for all non-negative continuous g on the state space with compact support. By [Kal, Theorem 4.2], to prove convergence of point measures associated to stochastic processes, it is enough to prove the convergence of Laplace functionals for such test functions. Then we can recover the convergence described by (1.5) via the following computation.

Proposition 1.5. *The point process*

$$N_n := \sum_{j=1}^n \delta_{(j/n, Y_j - \log n)}$$

$N_n \Rightarrow \mathcal{N}_{\text{cox}}$ where $\mathcal{N}_{\text{cox}}$ is a Cox point process with random intensity measure $\Lambda_x(dtdy) = 2f(x)e^{-y}dydt$, and Λ_x is drawn with density $f(x)$.

Remark 1.6. Note that we recover the limit in (1.5) by the following:

$$\begin{aligned} \lim_n \mathbb{P}(M_n - \log n \leq \tau) &= \lim_n \mathbb{P}(N_n((0, 1] \times (\tau, \infty)) = 0) \\ &= \mathbb{P}(\mathcal{N}_{\text{cox}}((0, 1] \times (\tau, \infty)) = 0) = \int e^{-\int_0^1 \int_\tau^\infty 2f(x)e^{-y}dydt} f(x)dx. \end{aligned}$$

We need the following lemma to prove such convergence.

Lemma 1.7. *Let $\{a_j\}_{j \in \mathbb{N}}$ be a sequence of non-negative real numbers. Then, if $\sup_{j \in \mathbb{N}} a_j \leq M$ for some $M < \infty$, and $\frac{1}{n} \sum_{j=1}^n a_j \rightarrow a$ for some $a \in \mathbb{R}_{\geq 0}$, then*

$$\lim_{n \rightarrow \infty} \prod_{j=1}^n \left(1 - \frac{1}{n} a_j\right) = \lim_{n \rightarrow \infty} e^{-\frac{1}{n} \sum_{j=1}^n a_j} = e^{-a}.$$

Proof. As $|a_j| \leq M$ for all j , then for big n , there is $|a_j/n| \leq \frac{1}{2}$ for all j . By calculating the MacLaurin series of $\log(1-t)$, there is some $M' < \infty$ such that for all $0 \leq t \leq \frac{1}{2}$,

$$|t + \log(1-t)| = \left| t - t - \frac{t^2}{2} - \frac{t^3}{3} - \dots \right| \leq M' t^2.$$

Therefore,

$$\log \prod_{j=1}^n \left(1 - \frac{1}{n} a_j\right) = \sum_{j=1}^n \log \left(1 - \frac{1}{n} a_j\right) = -\frac{1}{n} \sum_{j=1}^n a_j \pm M' \sum_{j=1}^n \left(\frac{1}{n} a_j\right)^2. \quad \square$$

Proof of Proposition 1.5. We will calculate the Laplace transform of N_n , as in [R2, Proposition 3.5] tells us that the Laplace transform uniquely determines a point measure. Let g be a non-negative continuous function with compact support (without loss of generality suppose the support of g on the z entry is contained in $[a, b) \subset \mathbb{R}_{\geq 0}$). Then, by independence and identical distribution,

$$\mathbb{E} \left[e^{-\int g dN_n} \right] = \mathbb{E} \left[e^{-\sum_{i=1}^n g(i/n, Y_i - \log n)} \right] = \int \left(\prod_{i=1}^n \mathbb{E} \left[e^{-g(i/n, Y_i - \log n)} | X_0 = x \right] \right) f(x) dx.$$

Now we continue to calculate $\mathbb{E} [e^{-g(i/n, Y_i - \log n)}]$, which is always integrable due to continuity and compact support. Let $Z_n = Y_1 - \log n$, then $F_{Z_n} = \mathbb{P}(Z_n < z) = \mathbb{P}(|X_1 - x| > e^{-z}/n) = 1 - \int_{x \pm e^{-z}/n} f(\xi) d\xi$. So

$$f_{Z_n}(z) = e^{-(z+\log n)} \left[f(x + e^{-(z+\log n)}) + f(x - e^{-(z+\log n)}) \right].$$

Therefore,

$$\begin{aligned} \mathbb{E} \left[e^{-g(i/n, Y_1 - \log n)} | X_0 = x \right] &= 1 - \mathbb{E} \left[1 - e^{-g(i/n, Y_1 - \log n)} | X_0 = x \right] \\ &= 1 - \int (1 - e^{-g(i/n, z)}) f_{Z_n}(z) dz \\ &= 1 - \frac{1}{n} \int [f(x + e^{-z}/n) + f(x - e^{-z}/n)] (1 - e^{-g(i/n, z)}) e^{-z} dz \\ &=: 1 - \frac{1}{n} A_{i,n}(x). \end{aligned}$$

By the Lebesgue Differentiation Theorem, for almost every x , $f(x \pm e^{-z}/n) \rightarrow f(x)$ as $n \rightarrow \infty$, hence

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n A_{i,n}(x) = 2f(x) \iint (1 - e^{-g(t, z)}) e^{-z} dz dt$$

By Lemma 1.7,

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[e^{-\int g dN_n} | X_0 = x \right] = \lim_{n \rightarrow \infty} \prod_{i=1}^n \left(1 - \frac{1}{n} A_{i,n} \right) = \exp \left(-2f(x) \iint (1 - e^{-g(t, z)}) e^{-z} dz dt \right)$$

Lastly, applying the Dominated Convergence Theorem,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E} \left[e^{-\int g dN_n} \right] &= \int \lim_{n \rightarrow \infty} \mathbb{E} \left[e^{-\int g dN_n} | X_0 = x \right] f(x) dx \\ &= \int \exp \left(-2f(x) \iint \left(1 - e^{-g(t,z)} \right) e^{-z} dz dt \right) f(x) dx \end{aligned} \quad \square$$

The beauty in establishing the convergence of N_n is that we can then apply the record-extraction mappings to N_n and $\mathcal{N}$ then apply the Continuous Mapping Theorem after checking the necessary almost sure continuity conditions, then the analogous formula as (1.6) can be obtained for the return process for the *i.i.d.* process. Especially, the particular proof for deriving such convergence for record time/value processes does not require conditioning argument. We will formally describe this in dynamical system context in § 5.

1.3. Hitting problem in dynamical setting. From our observations in the *i.i.d.* quantitative recurrence case, one sees that the proof was based on conditioning on the initial position $X_0 = x_0$, after which the remaining calculations behave like a ball-hitting problem, and we apply independence among the random variables. Then, using EVL results for stationary processes followed by integrating over the whole space, one obtains an averaged Gumbel distribution for the partial maxima process and an averaged-Gamma distribution for the records after normalisation. For dynamical systems, to convince ourselves that the record values do converge to the conjectured averaged-Gamma distribution, it is useful and reassuring to verify if the point measures associated with a hitting process with a fixed reference point still converges to a PRM. Indeed, such result holds.

Fix a generic $\tilde{x} \in \mathcal{X}$ and set $Y_j^{\tilde{x}} = -\log d(f^j x, \tilde{x})$, also the partial maximum $M_j^{\tilde{x}} = \max\{Y_1^{\tilde{x}}, \dots, Y_j^{\tilde{x}}\}$. For μ -almost every $\tilde{x}$, $\{Y_j^{\tilde{x}}\}_j$ is a stationary process, and by Lebesgue differentiation theorem, $n\mu \left(Y_j^{\tilde{x}} > \tau + \log n \right) \rightarrow 2\rho(\tilde{x})e^{-\tau}$ if ρ is bounded. By [HT1, Theorem 2.3, Theorem 3.1], under suitable mixing and recurrence conditions (the paper in question called them $\mathcal{D}_r(u_n^{(k)})$ and $\mathcal{D}'_r(u_n^{(k)})$, motivated by Leadbetter's works) satisfied by a wide range of maps enjoying fast decay of correlations, the process

$$M_n^{\tilde{x}}(t) := \left(M_{[nt]}^{\tilde{x}} - \log n \right)$$

converges to an extremal- F process $\mathbb{D}(0, \infty)$ (with respect to the Shorokhod J_1 topology, see [HT1, §1] for more details) generated by the distribution $F(u) = e^{-2\rho(\tilde{x})e^{-u}}$. Moreover, the counting measures

$$R_n = \sum_{j=1}^{\infty} \delta_{j/n} \cdot \mathbf{1}_{\{Y_j^{\tilde{x}} > M_{j-1}^{\tilde{x}}\}}, \quad V_n := \sum_{k=1}^{\infty} \delta_{(M_n^{\tilde{x}}(\tau_k/n))}$$

both converge in $\mathcal{M}_p(0, \infty)$ to PRM with intensity $1/t dt$ and $\lambda[a, b] = b - a$ respectively. In this case,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left((Y_{\tau_k}^{\tilde{x}} - \log n) \leq s \right) &= \lim_{n \rightarrow \infty} \mathbb{P} (V_n[0, s] \geq k) = \sum_{j=k}^{\infty} \frac{1}{j!} s^j e^{-s} \\ &= \int_0^s \frac{1}{(k-1)!} t^{k-1} e^{-t} dt = \mathbb{P}(\Gamma(k, 1) \leq s) \end{aligned}$$

So the normalised record values (that are above 0) do follow a Gamma distribution in limit¹.

¹See Remark 2.4 for further discussion

1.4. Statistics of records for recurrence in dynamics. The stationary nature of the problem can no longer directly facilitate the analysis when we consider the return process $Y_j(x) = -\log d(T^j x, x)$ for a dynamical system. However, various remarkable progress regarding EVL type convergence has been made in this non-stationary setting. For dynamical systems satisfying Young's condition, [Col, Theorem 4.1] established a formula similar to (1.5), and in [HT2, Theorem 2.3], for suitable choices of u_n , the number of returns to $B(x, e^{-u_n})$ converges to a Poisson-like distribution. We give a brief introduction to their results.

Let $(\mathcal{X}, T, \mu)$ be a dynamical system over the unit interval with domain of $T = \mathcal{X} \subset [0, 1]$ that satisfies the following assumptions.

- (A1) μ is a T -invariant probability that has exponential decay of correlations for BV against $L^1(\mu)$ observables, that is, there exists $C_T, \eta > 0$ for all $f \in BV$ and $g \in L^1(\mu)$, for all $k \in \mathbb{N}$,

$$\left| \int f \cdot g \circ T^k d\mu - \int f d\mu \int g d\mu \right| \leq C_T \|f\|_{BV} \|g\|_{L^1(\mu)} e^{-\eta k},$$

where for a function $\varphi : \mathcal{X} \rightarrow \mathbb{R}$, $\|\varphi\|_{BV} = \|\varphi\|_\infty + TV(\varphi)$.

- (A2) suppose $r = r_n \asymp n^{-a}$ for some $a > 0$. Then there exists $C_0, \beta_0 > 0$ such that for $E_j(r) := \{x \in \mathcal{X} : d(x, T^j x) < r\}$ and all $j \leq (\log n)^2$, $\mu(E_j(r)) \leq C_0 r^{\beta_0}$. Consequently, there exists $C_1, \beta_1 > 0$ such that $\mu(\tilde{E}(r)) \leq C_1 r^{\beta_1}$ where $\tilde{E}(r) = \cup_{j=1}^{(\log n)^2} E_j(r)$.

In the language of [HT2], $(\mathcal{X}, T, \mu)$ is said to be *hyperbolic recurrent* if it satisfies (A1) (A2) and μ is an invariant ergodic measure with density $\rho \in BV$. This implies $\|\rho\|_\infty < \infty$.

Theorem 1.8. *For a hyperbolic recurrent dynamical system ([HT2, Theorem 2.3]) $(\mathcal{X}, T, \mu)$, or a C^2 interval map satisfying Young's condition ([Col, Theorem 4.1]),*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\max_{1 \leq j \leq n} Y_j(x) \leq \tau + \log n \right) = \int e^{-2\rho(x)e^{-\tau}} \rho(x) dx$$

$$\lim_{n \rightarrow \infty} \mathbb{P} (\# \{j \leq n : d(T^j x, x) \leq e^{-\tau}/n\} = k) \rightarrow \int \frac{(2\rho(x)e^{-\tau})^k}{k!} e^{-2\rho(x)e^{-\tau}} \rho(x) dx$$

The above formulas say that the among the first n iterates, the largest log value of return distance to initial position converges to an averaged Gumbel law after normalisation, similar to the case in §1.1. In this paper we wish to obtain a more detailed statistics for returns; in particular, our aim is to verify limiting laws similar to (1.6), which encode information for the record behaviours of returns. We will now describe the class of maps for which our main theorem apply.

Notations. We will use the following notations. For a set or an event E , its indicator function is denoted by $\mathbf{1}_E$. For any real numbers A, B, C , write $A = B \pm C$ if $B - C \leq A \leq B + C$; for any sequence a_n , denote $a_n \lesssim n^{-\alpha}$ or $a_n = \mathcal{O}(n^{-\alpha})$ if there exists a uniform constant K such that $a_n \leq K n^{-\alpha}$ for all n large enough, and $a_n \asymp b_n$ if $a_n \lesssim b_n$ and $a_n \gtrsim b_n$.

2. MAIN RESULTS

We now describe the precise setting of the maps our main results apply to. Let $\mathcal{I}$ be a countable index set, and $\mathcal{P} = \{P_a\}_{a \in \mathcal{I}}$ a collection of subintervals in $\mathcal{X}$ with disjoint interiors covering $\mathcal{X}$. We say T is a *uniformly expanding Markov map* if

- (1) for any $a \in \mathcal{I}$, $T_a := T|_{P_a}$ is continuous, injective and $T(P_a)$ is a union of elements in $\mathcal{P}$;
- (2) there is a uniform constant $v > 1$ such that for all $a \in \mathcal{I}$, $|DT_a| \geq v$.

The *repeller* of T , denoted by Ξ , is the collection of points with all their forward iterates contained in $\mathcal{P}$, namely

$$\Xi := \left\{ x \in \mathcal{X} : T^k(x) \in \bigcup_{a \in \mathcal{I}} P_a \text{ for all } k \geq 0 \right\}.$$

The invariant ergodic measure μ in question is supported on Ξ . There is a shift system associated to T : let M be an $\mathcal{I} \times \mathcal{I}$ matrix such that $M_{ab} = 1$ if $T(P_a) \cap P_b \neq \emptyset$ and 0 otherwise. Let Σ denote the space of all *infinite admissible words*, i.e.,

$$\Sigma := \left\{ x = (x_0, x_1, \dots) \in \mathcal{I}^{\mathbb{N}_0} : M_{x_k, x_{k+1}} = 1, \forall k \geq 0 \right\}.$$

The natural projection map $\pi : \Sigma \rightarrow \Xi$ is defined by

$$x = \pi(x_0, x_1, \dots) \text{ if and only if } x \in \bigcap_{i=0}^{\infty} T^{-i} P_{x_i}.$$

The dynamics on Σ is the left shift $\sigma : \Sigma \rightarrow \Sigma$ given by $\sigma(x_0, x_1, \dots) = (x_1, x_2, \dots)$, then π defines a semi-conjugacy $T \circ \pi = \pi \circ \sigma$. The corresponding symbolic measure $\tilde{\mu}$ of μ is given by $\mu = \pi_* \tilde{\mu}$, i.e., for all Borel-measurable set $B \in \mathcal{B}([0, 1])$, $\mu(B) = \tilde{\mu}(\pi^{-1}B)$.

Denote $\mathcal{P}^n := \bigvee_{j=0}^{n-1} f^{-j} \mathcal{P}$, each $P \in \mathcal{P}^n$ corresponds to an n -cylinder in Σ : let $\Sigma_n \subseteq \Sigma$ denote all finite words of length n and for any $\mathbf{i} \in \Sigma$, the n -cylinder defined by $\mathbf{i}$ is

$$[\mathbf{i}] = [i_0, \dots, i_{n-1}] := \{y \in \Sigma : y_j = i_j, j = 0, \dots, n-1\}.$$

Then $\pi[i_0, i_1, \dots, i_{n-1}] = \bigcap_{j=0}^{n-1} T^{-j} P_{i_j} =: P_{\mathbf{i}}$. The *depth* of a cylinder $[\mathbf{i}]$ is the length of $\mathbf{i}$.

2.1. Technical assumptions on $(\mathcal{X}, T, \mu)$. In order to resolve the issue created by non-stationarity of our return process due to the independence of the initial position, logically we will require our system to exhibit fast decay of correlations so that the return process $Y_j(x) = -\log d(x, T^j x)$ resembles an *i.i.d.* sequence after relatively few iterates. In order to make sure the records do not accumulate, we also need to require the set of points that returns to a small neighbourhood of themselves is not very large so that records do not accumulate in a small time interval near 0 after rescaling. Therefore, in this paper, we would focus on systems that are hyperbolic recurrent (see §1.3) as (A1) and (A2) are natural candidates of assumptions that meet our requirements. In particular, (A2) implies the following ‘averaged’ version.

Lemma 2.1. *Let $(\mathcal{X}, T, \mu)$ be a dynamical system satisfying (A2) and $\rho \in \mathcal{BV}$. Then for every $r_n \lesssim n^{-1}$,*

$$n \int \sum_{j=1}^{(\log n)^2} \mu(B(\zeta, r_n) \cap T^{-j} B(\zeta, r_n)) d\mu(\zeta) \lesssim \frac{1}{\log n}.$$

Proof. For each $j \leq (\log n)^2$,

$$\begin{aligned} \int \mu(B(\zeta, r_n) \cap T^{-j} B(\zeta, r_n)) d\mu(\zeta) &= \iint \mathbf{1}_{B(\zeta, r_n)}(x) \cdot \mathbf{1}_{B(\zeta, r_n)}(T^j x) d\mu(x) d\mu(\zeta) \\ &= \int \left[\int \mathbf{1}_{B(\zeta, r_n)}(x) \cdot \mathbf{1}_{B(\zeta, r_n)}(T^j x) d\mu(\zeta) \right] d\mu(x), \end{aligned}$$

where the second line follows from Fubini. Notice that

$$\mathbf{1}_{B(\zeta, r_n)}(x) \cdot \mathbf{1}_{B(\zeta, r_n)}(T^j x) = \mathbf{1}_{B(x, r_n)}(\zeta) \mathbf{1}_{B(T^j x, r_n)}(\zeta),$$

so the inner integral equals $\mu(B(x, r_n) \cap B(T^j x, r_n))$. The intersection is non-empty only when $x \in E_j(2r_n)$, and in this case we can trivially bound $\mu(B(x, r_n) \cap B(T^j x, r_n))$ by $\mu(B(x, r_n)) \lesssim r_n$ since $\|\rho\|_\infty < \infty$. Then, by (A2) we can bound the original integral by

$$\int \mu(B(\zeta, r_n) \cap T^{-j}B(\zeta, r_n)) d\mu(\zeta) \leq \int_{E_j(2r_n)} \mu(B(x, r_n)) d\mu(x) \lesssim r_n^{\beta_0+1}.$$

Summing over $j \leq (\log n)^2$, for $r_n \asymp n^{-1}$, as $\beta_0 > 0$, for all n large enough,

$$n(\log n)^2 r_n^{\beta_0+1} \asymp (\log n)^2 n^{-\beta_0} \lesssim \frac{1}{\log n}. \quad \square$$

The next two assumptions will be helpful to prove the rate of convergence of point measures associated with the process Y_j . The first one is on the invariant density ρ and the second is on the corresponding symbolic dynamic of $(\mathcal{X}, T)$.

Condition (LC). The density ρ is said to satisfy (LC) if there exists a closed set $\mathcal{X}_0 \subseteq \mathcal{X}$ with $\mu(\mathcal{X}_0) = 1$ and a uniform constant C_ρ such that for every point $x \in \mathcal{X}_0$, for all $|h|$ small enough,

$$|\rho(x+h) - \rho(x)| \leq \frac{C_\rho}{|\log |h||}.$$

One will see as the proof proceeds, that the logarithmic part of the assumption will only be used once to overcome a technical difficulty regarding convergence rate. If the proof techniques are improved to bypass such difficulty, then the logarithmic continuity can be removed.

Condition (A1+). Say μ is *exponentially ψ -mixing* if there are $\tilde{C}_0, \eta' > 0$ and a monotone decreasing function $\psi(k) \leq \tilde{C}_0 e^{-\eta' k}$ for all $k \in \mathbb{N}$, such that the corresponding symbolic measure $\tilde{\mu}$ satisfies: for all $n, k \in \mathbb{N}$, $\mathbf{i} \in \Sigma_n$ and $\mathbf{j} \in \Sigma^* = \bigcup_{l \geq 1} \Sigma_l$,

$$\left| \frac{\tilde{\mu}(\mathbf{i} \cap \sigma^{-(n+k)} \mathbf{j})}{\tilde{\mu}[\mathbf{i}] \tilde{\mu}[\mathbf{j}]} - 1 \right| \leq \psi(k).$$

The ψ -mixing condition is rather strong, in the sense that it implies ϕ -mixing, β -mixing and other weaker mixing conditions. However, it is satisfied by a wide range of interval maps with a Gibbs-Markov structure.

2.2. Main result. Similar to § 1.3, define the partial maximum $M_j := \max\{Y_1, \dots, Y_j\}$ and the normalised partial maximal process

$$M_n(t) := M_{[tn]} - \log n = \max_{1 \leq j \leq [nt]} -\log d(T^j x, x) - \log n.$$

Recall from §1 the definition of τ_k . It is obvious that $Y_{\tau_k} - \log n = M_n(\tau_k/n)$. Also define the point processes

$$R_n := \sum_{j=1}^{\infty} \delta_{j/n} \cdot \mathbf{1}_{\{Y_j > M_{j-1}\}} \quad V_n := \sum_{\tau_k} \delta_{M_n(\tau_k/n)} \quad (2.1)$$

Theorem 2.2. Assume that $(\mathcal{X}, T, \mu)$ is a piecewise uniformly expanding interval map that is hyperbolic recurrent, and satisfies (A1+) and (LC). Then

- (1) R_n converges weakly to a Poisson random measure $\mathcal{R}$ on $(0, \infty)$ with intensity $1/t dt$, and
- (2) V_n also converges weakly to a Poisson random measure $\mathcal{V}$ on $(0, \infty)$ with intensity measure $\lambda_V[a, b] = (b - a)$.

The above theorem is a consequence of the following:

Theorem 2.3. *Under the assumption of Theorem 2.2, the point process*

$$N_n := \sum_{j=1}^n \delta_{(j/n, Y_j - \log n)}$$

converges weakly to a Cox process $\mathcal{N}$ on $(0, 1] \times \mathbb{R}$ with random intensity measure given by $\Lambda_x(dydt) = 2\rho(x)e^{-y}dydt$, while Λ_x is drawn over $x \in \mathcal{X}$ with density $\rho(x)$.

As far as the author knows, although the rate of convergence (with error $\lesssim 1/\log n$) is not directly derived for the distribution, our result is still one of the few quantitative statement in the literature of dynamical extreme value theory, and is first in establishing a relation between return process and the Cox process. To upgrade to a genuine rate of convergence of the laws may require other techniques such as Stein's method using characterising equations. We also make the following remarks.

Remark 2.4. (1) *For intuition, conditional on the initial point x , the limiting point process is Poisson with intensity $2\rho(x)e^{-z} dt dz$ which coincides with the hitting result in [HT1], then averaging over the initial point therefore produces the Cox process which agrees with the main theorem [HT2].*

(2) *One may be tempted to conclude, as a consequence of Theorem 2.2, that for fixed k , the k -th record also follows a $\Gamma(k, 1)$ distribution. Indeed, by Lemma 1.3, the probability of having more than k records in the interval $[0, r]$ is*

$$\mathbb{P}(V_n[0, r] \geq k) \rightarrow \mathbb{P}(\Gamma(k, 1) \leq r).$$

This is consistent with § 1.1. This appears to be unreasonable: for any fixed k , the ordinary k -th record is almost surely finite, hence the normalisation $Y_{\tau_k} - \log n$ goes to $-\infty$ as $n \rightarrow \infty$ almost surely. The correct way to interpret the main theorem is: the k -th record above the $\log n$ threshold has a $\Gamma(k, 1)$ distribution. In other words, if we set

$$J_n := \#\{j \leq n : Y_j > M_{j-1} \text{ and } Y_j < \log n\},$$

then for any $k \in \mathbb{N}$, $\mathbb{P}(Y_{\tau_{J_n+k}} - \log n \leq r) \rightarrow \mathbb{P}(\Gamma(k, 1) \leq r)$. This interpretation is consistent with motivating results in i.i.d. setting, and reflects the convergence in vague topology: such topology only requires convergence for test functions with compact support, hence the values diverge to $-\infty$ after recentering are simply not seen by the test function.

(3) *By [R2, Proposition 4.12, Proposition 4.13], for the i.i.d. case as in § 1.1, as a consequence of [R2, Proposition 4.11], the Continuous Mapping Theorem [Kal] and embedding into a extremal process, for almost every x ,*

$$\frac{\log \tau_k}{k} \rightarrow 1 \text{ and } \frac{\log(\tau_k - \tau_{k-1})}{k} \rightarrow 1. \quad (2.2)$$

Due to the strongly dependent nature of our dynamical process, the analogous almost sure statement has yet been derived because explicit error bounds of convergence rate are hard to obtain. But it is very possible that these formulas hold. Especially, if there is a way to upgrade the rate of convergence (currently at $1/\log n$) of Laplace functionals to a stronger estimate (so that the error terms are summable in n), then one may be able to apply Borel-Cantelli arguments.

3. EXAMPLES

We give examples of interval maps that satisfy the assumptions for our main results. In particular, these maps have the Gibbs-Markov property hence the corresponding symbolic dynamics is a

subshift of finite type or full-shift on countable alphabet, and the corresponding invariant measure $\tilde{\mu}$ is the unique Gibbs equilibrium for the geometric potential $-\log |T' \circ \pi|$. $\tilde{\mu}$ is ψ -mixing [Bow]. Condition (A2) can be verified by [HNT, Lemma 3.4]²

Example 3.1 (Doubling map). Let $T(x) = 2x \pmod{1}$ be the doubling map, and $\mu = \text{Leb}$. Then this map satisfies all assumptions.

Example 3.2 (Finitely branched Gibbs-Markov maps). Let $T : [0, 1] \rightarrow [0, 1]$ be an interval map with finitely many branches with the Gibbs-Markov property (see the precise definition in [HNT]), and suppose T on each branch is $C^{1+\alpha}$. Let μ be the unique ergodic T -invariant measure absolutely continuous with respect to the Lebesgue measure, then μ also has the Gibbs property. It is well known that maps of this kind are hyperbolic recurrent [HT2], and satisfy (A1+) [Bow]. Within each cylinder set, the invariant density is Lipschitz, hence as the boundary points of cylinder sets are countable, (LC) also holds.

Example 3.3 (Gauss map). Let $G(x) = \frac{1}{x} - \lfloor \frac{1}{x} \rfloor$ be the Gauss map and μ_G the Gauss measure with density $\frac{d\mu_G}{d\text{Leb}}(x) = \frac{1}{(\log 2)(1+x)} dx$. This system is hyperbolic recurrent, and its symbolic representation is a full-shift on natural numbers, the equilibrium with respect to the geometric potential $-\log |G' \circ \pi|$ is a Gibbs measure hence satisfies (A1+). Condition (LC) can be easily verified for every $x \in [0, 1]$.

4. PROOF OF THEOREM 2.3

The main tool to verify the convergence in question is by checking the convergence in the Laplace functional. In accurate terms, for $N_n = \sum_{j=1}^n \delta_{(j/n, Y_j - \log n)}$, we will check the following convergence

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_\mu \left[e^{-\int g dN_n} \right] &= \mathbb{E}_\mu \left[\exp \left(- \iint (1 - e^{-g(t,z)}) 2\rho(x) e^{-z} dz dt \right) \right] \\ &= \int \exp \left(- \int (1 - e^{-g(t,z)}) 2\rho(x) e^{-z} dz dt \right) \rho(x) dx \end{aligned} \quad (4.1)$$

for every non-negative continuous function g with compact support. Then, by [R2, Proposition 3.5], as the Laplace transform uniquely determines a measure, the desired convergence result holds. The major difficulty comes from the non-stationary nature of the random point measures $\{N_n\}_n$, which are defined over return distances, so the precise calculation of $\mathbb{E}_\mu[e^{-\int g dN_n}]$ can be very complicated. Standard extreme value theory setup including Leadbetter's $\mathcal{D}$ and $\mathcal{D}'$ conditions (even the non-stationary version derived by [FFV] as a dynamical adaptation of Husler's condition) are not applicable for this point measure. Instead, we introduce an auxiliary point measure whose Laplace transform is easier to compute, and is useful in approximating that of N_n

4.1. Spatial-Temporal partition. The partition and related methods used in [HT2] will be handy for decorrelating the initial conditions. Let $\mathcal{U}_m = \{\Delta_k\}_k$ be the partition of $\mathcal{X}$ using intervals Δ_k (with disjoint interiors) of length $1/m$ and with centre ζ_k . It is obvious that $\#\mathcal{U}_m = m$, and we choose

$$m = n^{1+\beta'} \text{ and set } \delta = |\Delta|/2.$$

Let $r > 0$. For each $\Delta \in \mathcal{U}_m$, set

$$\Delta_0 := \{y \in \mathcal{X} : d(y, \Delta) \leq r\}, \quad \Delta^\pm := \{y \in \mathcal{X} : d(y, \Delta) \leq r \pm \delta\}.$$

²The statement of [HNT, Lemma 3.4] is a bound for the Lebesgue measure of $E_j(r)$ sets, but since we will take μ which is an acip, the upper bound holds with the same exponent for $\mu(E_j(r))$.

4.2. Approximation of Laplace functional by auxiliary point process. In order to prove the convergence of the Laplace transform, we first need to approximate the behaviour of N_n via some auxiliary point measures. Let $\mathcal{U}_m$ be the partition described above, for each $\Delta \in \mathcal{U}_m$, set

$$Y_j^\Delta(x) := -\log d(T^j x, \zeta), \text{ where } \zeta \text{ is the centre of } \Delta.$$

Let

$$N_n^\Delta := \sum_{j=1}^n \delta_{(j/n, Y_j^\Delta - \log n)}.$$

In comparison to the process N_n , which is of our main interest and related to the returns, N_n^Δ corresponds to the hitting process to the Δ -component. We then define our auxiliary point process: let $\tilde{N}_n$ be a random point measure such that we

- draw $\Delta \in \mathcal{U}_m$ with probability $\mu(\Delta)$;
- then draw $x \in \mathcal{X}$ according to μ .

Note that $\tilde{N}_n$ is still simple. Occasionally, abusing notation, we may write $\tilde{N}_n = \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) N_n^\Delta$ (abusing in the sense that $\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) N_n^\Delta$ in its literal meaning is not a point measure, since its values may be non-integers).

Our first goal is to show that the Laplace transforms of these two point measures are asymptotically equal with an error term of order $1/\log n$. Namely, for each non-negative, continuous function $g : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ with compact support,

$$\begin{aligned} \mathcal{L}_n(g) &:= \mathbb{E}_\mu \left[e^{-\int g dN_n} \right] = \int e^{-\int g dN_n} d\mu, \\ \tilde{\mathcal{L}}_n(g) &:= \mathbb{E}_\mu \left[e^{-\int g d\tilde{N}_n} \right] = \int e^{-\int g d\tilde{N}_n} d\mu = \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \int e^{-\int g dN_n^\Delta} d\mu. \end{aligned}$$

We aim to show $\lim_n |\mathcal{L}_n(g) - \tilde{\mathcal{L}}_n(g)| \rightarrow 0$. By the Mean Value Theorem, $|e^{-u} - e^{-v}| \leq |u - v|$ for all $u, v \geq 0$, since g is non-negative,

$$\mathbb{E}_\mu \left[\left| \int g d\tilde{N}_n - \int g dN_n \right| \right] \leq \mathcal{L}_n(g) - \tilde{\mathcal{L}}_n(g) \leq \mathbb{E}_\mu \left[\left| \int g dN_n - \int g d\tilde{N}_n \right| \right].$$

Thus it suffices to prove the following proposition.

Proposition 4.1. *Let $g : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ be a non-negative continuous function with compact support and $N_n, \tilde{N}_n$ defined as above. Then if $(\mathcal{X}, T, \mu)$ is hyperbolic recurrent,*

$$\mathbb{E}_\mu \left[\left| \int g dN - \int g d\tilde{N}_n \right| \right] \lesssim \frac{1}{\log n}. \quad (4.2)$$

Remark 4.2. *When verifying the convergence of the Laplace functional for N_n^Δ for any $\Delta \in \mathcal{U}_m$, it is enough to restrict to C^1 non-negative test functions. This is because C^1 non-negative test functions are dense: for any $\varepsilon > 0$ and any g a non-negative continuous functions with compact support $K = [a, b] \subset [-D, D] \subset (0, 1] \times \mathbb{R}$, there exists $g_\varepsilon \geq 0$ that is C^1 and supported on K , such that $\|g - g_\varepsilon\|_\infty < \varepsilon$, hence*

$$\left| \int g dN_n^\Delta - \int g_\varepsilon dN_n^\Delta \right| \leq \varepsilon N_n^\Delta(K') \leq \varepsilon N_n^\Delta((0, 1] \times [-D, D])$$

for some $0 \leq D < \infty$. Taking expectations on both sides,

$$\mathbb{E}_\mu[N_n^\Delta((0, 1] \times [-D, D])] \leq n\mu(B(\zeta, e^D/n)) \leq 2e^D \|\rho\|_\infty < \infty,$$

where ζ is again the centre of Δ . This inequality does not depend on Δ and ε was arbitrarily small, hence for any g non-negative continuous function with compact support, for any $\Delta \in \mathcal{U}_m$, $\mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right]$ can be replaced by a Laplace functional of a C^1 function with arbitrarily small cost.

The same argument applies to N_n as well: by Lemma 4.3 below and assumption (A2),

$$\mathbb{E}_\mu[N_n((0, 1] \times [-D, D])] \leq \sum_{j=1}^n E_j(e^D/n) \lesssim n^{-\beta_1} + ne^D/n < \infty.$$

Lemma 4.3. For $(\mathcal{X}, T, \mu)$ a hyperbolic recurrent system, for $j > (\log n)^2$ and $r = r_n \asymp n^{-1}$,

$$\mu(E_j(r_n)) \leq 3\delta^{-1}e^{-(\log n)^2} + 2\|\rho\|_\infty^2(r + \delta) \lesssim r \asymp n^{-1}. \quad (4.3)$$

Proof. For each $\Delta \in \mathcal{U}_m$, if $x \in E_j(r) \cap \Delta$, then $d(T^j x, \zeta) \leq d(T^j x, x) + d(x, \zeta) \leq r + \delta$. In other words, $x \in E_j(r) \cap \Delta$ implies that $\mathbf{1}_\Delta(x)\mathbf{1}_{\Delta^+}(T^j x) = 1$. Hence, using decay of correlations (A1) and boundedness of the density ρ , as $\|\mathbf{1}_\Delta\|_{BV} \leq 3$,

$$\begin{aligned} \mu(E_j(r)) &\leq \sum_{\Delta \in \mathcal{U}_m} \int \mathbf{1}_\Delta \cdot \mathbf{1}_{\Delta^+} \circ T^j d\mu \leq \sum_{\Delta \in \mathcal{U}_m} \left(\int \mathbf{1}_\Delta d\mu \int \mathbf{1}_{\Delta^+} d\mu + \|\mathbf{1}_\Delta\|_{BV} \|\mathbf{1}_{\Delta^+}\|_{L^1(\mu)} e^{-\eta j} \right) \\ &\leq 3\delta^{-1}e^{-\eta j} + \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta)\mu(\Delta^+) \leq \delta^{-1}e^{-(\log n)^2} + 2\|\rho\|_\infty^2(r + \delta). \quad \square \end{aligned}$$

Proof of Proposition 4.1. Let g be a C^1 non-negative function with compact support. Given $\mathcal{U}_m$, we have

$$\begin{aligned} &\int g dN_n(x) - \int g d\tilde{N}_n(x) \\ &= \sum_{\Delta \in \mathcal{U}_m} \mathbf{1}_\Delta(x) \sum_{j=1}^n g(j/n, Y_j(x) - \log n) - \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^n \mu(\Delta) g(j/n, Y_j^\Delta(x) - \log n) \\ &= \sum_{\Delta \in \mathcal{U}_m} \mathbf{1}_\Delta(x) \sum_{j=1}^n g(j/n, Y_j(x) - \log n) - g(j/n, Y_j^\Delta(x) - \log n) \\ &\quad + \sum_{\Delta \in \mathcal{U}_m} (\mathbf{1}_\Delta(x) - \mu(\Delta)) \sum_{j=1}^n g(j/n, Y_j^\Delta(x) - \log n) \\ &=: \text{Err}_{n,1} + \text{Err}_{n,2}. \end{aligned}$$

Note that $\mathbb{E}_\mu \left[\left| \int g dN - \int g d\tilde{N}_n \right| \right] \leq \mathbb{E}_\mu [|\text{Err}_{n,1}|] + \mathbb{E}_\mu [|\text{Err}_{n,2}|]$. Because g has compact support, many of the $g(j/n, \cdot)$ terms are 0 unless $T^j x$ is either really close to the initial condition x , or the centre ζ of Δ . This naturally allows us to bound these error terms by restricting x on $E_j(r_n)$ sets with $r_n \asymp n^{-1}$.

Step 1. We first estimate $E_\mu[\text{Err}_{n,2}]$. Let

$$\phi_{n,j}^\Delta(\cdot) := g(j/n, -\log(d(\cdot, \zeta)) - \log n).$$

Without loss of generality, suppose the support of g in its z -coordinate is contained in $[a, b]$, and recall that $r_n(z) := \frac{e^{-z}}{n}$,

$$\phi_{n,j}^\Delta(T^j x) \neq 0 \text{ when } d(T^j x, \zeta) \in \left[e^{-b}/n, e^{-a}/n \right] =: [r_n(b), r_n(a)].$$

This gives

$$\|\phi_{n,j}^\Delta\|_{L^1(\mu)} \leq \|g\|_\infty \cdot \mu(\text{supp}(\phi_{n,j}^\Delta)) \leq \|g\|_\infty \cdot \mu(B(\zeta, r_n(a))).$$

Since the density $\rho = \frac{d\mu}{dLeb} \in \mathcal{BV}$ is bounded, we can conclude that $\|\phi_{n,j}^\Delta\|_{L^1(\mu)} \lesssim n^{-1}$. Also, in order that $\mathbf{1}_\Delta(x)\phi_{n,j}^\Delta(x) \not\equiv 0$, we must have $d(T^j x, x) \leq e^{-a}/n + \delta$, that is, $x \in E_j(r_n(a) + \delta)$.

By assumptions (A1) and (A2),

$$\begin{aligned} & \mathbb{E}_\mu \left[\sum_{\Delta \in \mathcal{U}_m} \int \mathbf{1}_\Delta(x) \sum_{j=1}^n \phi_{n,j}^\Delta(T^j x) d\mu(x) \right] \\ &= \sum_{\Delta \in \mathcal{U}_m} \int \mathbf{1}_\Delta(x) \sum_{j=1}^{(\log n)^2} \phi_{n,j}^\Delta(T^j x) d\mu(x) + \sum_{\Delta \in \mathcal{U}_m} \int \mathbf{1}_\Delta(x) \sum_{j > (\log n)^2}^n \phi_{n,j}^\Delta(T^j x) d\mu(x) \\ &\leq C_0(\log n)^2 \|g\|_\infty (r_n(a) + \delta)^{\beta_0} + \sum_{\Delta \in \mathcal{U}_m} \sum_{j > (\log n)^2}^n \int \mathbf{1}_\Delta d\mu \int \phi_{n,j}^\Delta d\mu + C_T \|\mathbf{1}_\Delta\|_{BV} \|\phi_{n,j}^\Delta\|_{L^1(\mu)} e^{-\eta(\log n)^2} \\ &\leq C_0(\log n)^2 \|g\|_\infty (r_n(a) + \delta)^{\beta_0} + n^{2+\beta'} C_T 3 \|g\|_\infty n^{-\eta \log n} + \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{j > (\log n)^2}^n \|\phi_{n,j}^\Delta\|_{L^1(\mu)} \end{aligned}$$

Then,

$$\begin{aligned} \mathbb{E}_\mu[\text{Err}_{n,2}] &\leq C_0(\log n)^2 \|g\|_\infty (r_n(a) + \delta)^{\beta_0} + n^{2+\beta'} C_T 3 \|g\|_\infty n^{-\eta \log n} - \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^{(\log n)^2} \mu(\Delta) \|\phi_{n,j}^\Delta\|_{L^1(\mu)} \\ &\lesssim (\log n)^2 n^{-\beta_0} + n^{2+\beta'} n^{-\eta \log n} - (\log n)^2 n^{-1} \lesssim \frac{1}{\log n}. \end{aligned}$$

Similarly, by completely ignoring the first $(\log n)^2$ iterates in $\sum \mathbf{1}_\Delta \sum_{j=1}^n \phi_{n,j}$, we bound the expectation from below by

$$\begin{aligned} \mathbb{E}_\mu[\text{Err}_{n,2}] &\geq - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{j=1}^{(\log n)^2} \|\phi_{n,j}^\Delta\|_{L^1(\mu)} + n^{2+\beta'} C_T 3 \|g\|_\infty n^{-\eta \log n} \\ &\gtrsim - \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) (\log n)^2 n^{-1} - n^{2+\beta'} n^{-\eta \log n} \gtrsim -(\log n)^{-1}. \end{aligned}$$

We therefore conclude that $\lim_{n \rightarrow \infty} |\mathbb{E}_\mu[\text{Err}_{n,2}]| = 0$ with an error term uniformly asymptotically bounded by $\frac{1}{\log n}$.

Step 2. Next we treat $\text{Err}_{n,1}$.

Fix $\Delta \in \mathcal{U}_m$, recall by definition, for all $x \in \Delta$, $d(x, \zeta) \leq \delta$, hence by triangular inequality, for all j ,

$$|d(x, T^j x) - d(T^j x, \zeta)| \leq d(x, \zeta) \leq \delta = n^{-(1+\beta')}. \quad (4.4)$$

If $d(T^j x, \zeta) > r_n(a) + \delta$ then $d(T^j x, x) > r_n(a)$ hence both $g(j/n, Y_j(x) - \log n)$ and $g(j/n, Y_j^\Delta - \log n)$ are automatically zero. Similarly if $d(T^j x, \zeta) < r_n(b) - \delta$ then both g terms are 0. So neither of $g(j/n, Y_j - \log n)$ and $g(j/n, Y_j^\Delta - \log n)$ is automatically zero, for all $j = 1, \dots, n$, only when $d(T^j x, x) \in [r_n(b) + \delta, r_n(a) - \delta]$. In this case, $\frac{|d(T^j x, x) - d(T^j x, \zeta)|}{d(T^j x, \zeta)} \leq \frac{\delta}{e^{-b} n^{-1/2}} \leq 2e^b n^{-\beta'}$. As for all z

small, by Taylor expansion at 0 we have the rough upper bound $|\log(1 - z)| \leq 2|z|$,

$$|Y_j(x) - \log n - (Y_j^\Delta(x) - \log n)| = \left| \log \left(1 - \frac{d(T^j x, x) - d(T^j x, \zeta)}{d(T^j x, x)} \right) \right| \leq 4e^b n^{-\beta'}. \quad (4.5)$$

Define

$$\omega_\ell(g) := \sup\{|g(t, u) - g(t, v)| : t, u, v \in \text{supp}(g), |u - v| \leq \ell\}, \quad (4.6)$$

as g is C^1 and has compact support in z coordinate, $\omega(g)_\ell \lesssim \ell$. Also define the sets

$$\begin{aligned} \Delta_j^{\text{bulk}} &:= \left\{ x \in \Delta : d(T^j x, \zeta) \in \left[\frac{e^{-b}}{n} + \delta, \frac{e^{-a}}{n} - \delta \right] \right\} \\ \Delta_j^{\text{shell}} &:= \left\{ x \in \Delta : d(T^j x, \zeta) \in \left[\frac{e^{-b}}{n} - 2\delta, \frac{e^{-b}}{n} + \delta \right] \cup \left[\frac{e^{-a}}{n} - \delta, \frac{e^{-a}}{n} + 2\delta \right] \right\} \end{aligned}$$

Note that

- $\Delta_j^{\text{bulk}} \subset E_j(r_n(a)) \cap \Delta$, $\bigcup_{\Delta \in \mathcal{U}_m} \Delta_j^{\text{shell}} \subset E_j(r_n(a) + 3\delta)$;
- $\Delta_j^{\text{bulk}} \cap \Delta_j^{\text{shell}} = \emptyset$;
- $g|_{\Delta \setminus (\Delta_j^{\text{bulk}} \cup \Delta_j^{\text{shell}})} \equiv 0$;
- For any $r_1 < r_2$, let $A(\zeta, r_1, r_2) := B(\zeta, r_2) \setminus B(\zeta, r_1)$. Then

$$\mathbf{1}_{\Delta_j^{\text{shell}}} = \mathbf{1}_\Delta \cdot (\mathbf{1}_{A(\zeta, r_n(a) - \delta, r_n(a) + 2\delta)} \cup \mathbf{1}_{A(\zeta, r_n(b) - \delta, r_n(b) + \delta)}) \circ T^j \quad (4.7)$$

Let us reduce the cumbersome notation above by defining

$$A'(\zeta, a, b, \delta) := A(\zeta, r_n(a) - \delta, r_n(a) + 2\delta) \cup A(\zeta, r_n(b) - \delta, r_n(b) + \delta).$$

By boundedness of the density ρ

$$\mu(A'(\zeta, a, b, \delta)) = \mu(B(\zeta, r_n(a) + 2\delta) \setminus B(\zeta, r_n(a) - \delta)) + \mu(B(\zeta, r_n(b) + \delta) \setminus B(\zeta, r_n(b) - \delta)) \lesssim \delta. \quad (4.8)$$

Using both (A1) (decay of correlations) and (A2) (estimate of short returns), and the invariance of μ together with (4.8),

$$\begin{aligned} \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^n \int \mathbf{1}_{\Delta_j^{\text{shell}}} d\mu &= \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^{(\log n)^2} \int \mathbf{1}_{\Delta_j^{\text{shell}}} d\mu + \sum_{\Delta \in \mathcal{U}_m} \sum_{j > (\log n)^2}^n \int \mathbf{1}_{\Delta_j^{\text{shell}}} d\mu \\ &\leq \tilde{E}_{(\log n)^2}(r_n(a) + 3\delta) + \sum_{\Delta \in \mathcal{U}_m} \sum_{j > (\log n)^2}^n \left(\mu(\Delta) \mu(A'(\zeta, a, b, \delta)) + 3e^{-\eta(\log n)^2} \right) \\ &\lesssim n^{-\beta_1} + n\delta^{-1}(\delta^2 + n^{-\eta(\log n)^2}) \lesssim n^{-\beta_1} + n^{-\beta'}. \end{aligned}$$

Therefore,

$$\begin{aligned} &\sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^n \int_{\Delta_j^{\text{shell}}} |g(j/n, Y_j(x) - \log n) - g(j/n, Y_j^\Delta(x) - \log n)| d\mu(x) \\ &\leq 2\|g\|_\infty \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^n \int \mathbf{1}_{\Delta_j^{\text{shell}}} d\mu \lesssim n^{-\beta_1} + n^{-\beta'} \end{aligned}$$

For any fixed j and different Δ and $\Delta' \in \mathcal{U}_m$, $\Delta_j^{\text{bulk}} \cap \Delta_j'^{\text{bulk}} = \emptyset$, and

$$\begin{aligned} & \int \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^n \mathbf{1}_{\Delta_j^{\text{bulk}}}(x) |g(j/n, Y_j(x) - \log n) - g(j/n, Y_j^\Delta(x) - \log n)| d\mu(x) \\ & \leq \sum_{j=1}^n \int_{\bigcup_{\Delta \in \mathcal{U}_m} \Delta_j^{\text{bulk}}} |g(j/n, Y_j(x) - \log n) - g(j/n, Y_j^\Delta(x) - \log n)| \\ & \leq \omega_{4e^{b_n - \beta'}}(g) \sum_{j=1}^n \mu \left(\bigcup_{\Delta \in \mathcal{U}_m} \Delta_j^{\text{bulk}} \right) \leq \omega_{4e^{b_n - \beta'}}(g) \sum_{j=1}^n \mu(E_j(r_n(a))) \end{aligned}$$

By (A2) and (4.3),

$$\sum_{j=1}^n \mu(E_j(r_n(a))) \leq C_1(r_n(a))^{\beta_1} + n \left(3\delta^{-1} e^{-(\log n)^2} + 2\|\rho\|_\infty^2 (r_n(a) + \delta) \right) < \infty.$$

Lastly, combining all the error estimates and use the fact that $g \in C^1$ so that $\omega_{4e^{b_n - \beta'}}(g) \lesssim n^{-\beta'}$,

$$\begin{aligned} \mathbb{E}_\mu[|\text{Err}_{n,1}|] & \leq \int \sum_{\Delta \in \mathcal{U}_m} \sum_{j=1}^n \mathbf{1}_\Delta(x) |g(j/n, Y_j(x) - \log n) - g(j/n, Y_j^\Delta(x) - \log n)| d\mu(x) \\ & \lesssim n^{-\beta_1} + n^{-\beta'} + \omega_{4e^{b_n - \beta'}}(g) \lesssim n^{-\beta_1} + n^{-\beta'} + n^{-\beta'} \lesssim \frac{1}{\log n}. \end{aligned} \quad \square$$

4.3. Limit of Laplace functional for the auxiliary process. It remains to show that the Laplace function of $\tilde{N}_n$ does converge to the desired limit with a similar rate, *i.e.* in the form of (4.1). Especially, we will show that the Laplace functional of the sum over N_n^Δ for converges to that of a Cox process. Almost all quantities in the following section are defined with respect to some point $\zeta \in \mathcal{X}$. When an object has Δ as a superscript, it means the relevant point $\zeta = \zeta_\Delta$ is the centre of a partition set $\Delta \in \mathcal{U}_m$. Otherwise, it is an arbitrary $\zeta \in \mathcal{X}$.

Proposition 4.4. *Let g be an arbitrary non-negative C^1 function with compact support. Define*

$$\Psi_\infty^g(\zeta) := \exp \left(- \iint \left(1 - e^{-g(t,z)} \right) 2\rho(\zeta) e^{-z} dz dt \right).$$

Under the conditions of Theorem 2.2,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \left| \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] - \Psi_\infty^g(\zeta_\Delta) \right| \lesssim \frac{1}{\log n}, \quad (4.9)$$

where ζ_Δ is the centre of $\Delta \in \mathcal{U}_m$.

We will again prove this proposition in several steps. Fix g a C^1 non-negative test function such that $\text{supp}(g) \subset (0, 1] \times [a, b]$. First, approximate g by simple functions. Fix $K \in \mathbb{N}$ and let

$$0 = t_0 < t_1 < \cdots < t_K = 1$$

be a partition of equal length over the support of g in its t coordinate. Similarly set

$$a = z_0 < z_1 < \cdots < z_L = b,$$

which is a partition of the z -support and without loss of generality is assumed to be contained in $[a, b]$. Set

$$I_k = \left\{ j \in \{1, \dots, n\} : \frac{j}{n} \in [t_{k-1}, t_k) \right\}. \quad (4.10)$$

With the above partition, we prove the following error bound.

Lemma 4.5. *For $\tilde{g} = \tilde{g}_L$ defined as above, for every $\Delta \in \mathcal{U}_m$,*

$$\left| \mathbb{E}_\mu \left[e^{-\int \tilde{g} dN_n^\Delta} \right] - \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] \right| \lesssim \max \left\{ \frac{1}{L}, \frac{1}{K} \right\}. \quad (4.11)$$

Proof. By C^1 and compact support, $\|D(g)\|_\infty < \infty$, so using the Mean Value Theorem,

$$\sup_{t,z} |g(t,z) - \tilde{g}(t,z)| \leq \|D(g)\|_\infty \max \left\{ \frac{1}{L}, \frac{1}{K} \right\}.$$

Then up to a uniform constant depending on g ,

$$\begin{aligned} \left| \mathbb{E}_\mu \left[e^{-\int \tilde{g} dN_n^\Delta} \right] - \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] \right| &\leq \left| \mathbb{E}_\mu \left[\int g dN_n^\Delta \right] - \mathbb{E}_\mu \left[\int \tilde{g} dN_n^\Delta \right] \right| \lesssim \max \left\{ \frac{1}{L}, \frac{1}{K} \right\} E_\mu [N_n^\Delta([a, b])] \\ &\lesssim \max \left\{ \frac{1}{L}, \frac{1}{K} \right\} n\mu(B(\zeta, e^{-a}/n)) \lesssim \max \left\{ \frac{1}{L}, \frac{1}{K} \right\}. \quad \square \end{aligned}$$

We choose $K = L = \log n$, so the error between their Laplace functionals is of order $1/\log n$, then compute $\mathbb{E}_\mu \left[e^{-\int \tilde{g} dN_n^\Delta} \right]$. Define the annuli

$$A_l(\zeta) = B(\zeta, r_n(z_{l-1})) \setminus B(\zeta, r_n(z_l)), \quad r_n(z) = \frac{e^{-z}}{n}.$$

hence the counting variable

$$M_{n,k,l} = \sum_{j \in I_k} \mathbf{1}_{A_l(\zeta)}(T^j x), \text{ for each } 1 \leq k \leq K, \text{ and } 1 \leq l \leq L,$$

where $I_k = \{j : j/n \in [t_{k-1}, t_k]\}$. Then,

$$\begin{aligned} \int \tilde{g} dN_n^\Delta(x) &= \sum_{j=1}^n \tilde{g}(j/n, Y_j^\Delta(x) - \log n) \\ &= \sum_{l=1}^L \sum_{k=1}^K \sum_{j=1}^n g(t_{k-1}, z_{l-1}) \mathbf{1}_{[t_{k-1}, t_k]}(j/n) \mathbf{1}_{[z_{l-1}, z_l]}(Y_j^\Delta(x) - \log n) \\ &= \sum_{l=1}^L \sum_{k=1}^K g(t_{k-1}, z_{l-1}) M_{n,k,l}^\Delta(x) \end{aligned}$$

Hence $\mathbb{E}_\mu \left[e^{-\int \tilde{g} dN_n^\Delta} \right]$ is a function of the random vectors $(M_{n,k,1}^\Delta, \dots, M_{n,k,L}^\Delta)$.

Define the total variation metric d_{TV} : for two non-negative integer valued random variables, X, Y ,

$$d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) := \frac{1}{2} \sum_{k=0}^{\infty} |\mathbb{P}(X = k) - \mathbb{P}(Y = k)|$$

where $\mathcal{L}$ denotes the law (distribution) of a random variable. In particular, if the two random variables are defined over the same probability space, then $d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) \leq \mathbb{P}(X \neq Y)$.

Below we also recall a multivariate Poisson approximation theorem from [Bar, Theorem 1], developed from [LCam].

Theorem 4.6 (Multivariate Le Cam's Poisson Approximation theorem). *Let $X_1, \dots, X_n$ be independent Bernoulli random d -vectors such that $\mathbb{P}(X_j = \epsilon_i) = p_j^{(i)}$, where ϵ_i represents the i 'th coordinate unit vector. Let $p_j = \sum_{i=1}^d p_j^{(i)}$, then*

$$d_{\text{TV}} \left(\mathcal{L} \left(\sum_{i=1}^n X_i \right), \mathcal{L} (Z_1, \dots, Z_d) \right) \leq \sum_{j=1}^d p_j^2.$$

where $Z_1, \dots, Z_d$ are independent Poisson vectors, such that

$$Z_i \sim \text{Poi} \left(\sum_{j=1}^d p_j^{(i)} \right).$$

The main idea towards proving Proposition 4.4 is via coupling the locally defined point measures, N_n^Δ , with Poisson random variables via the total variation metric. More precisely, we will slice the time window into blocks of size $(\log n)^3$, separated by some gaps, then apply the decorrelation argument. The stronger type of mixing (A1+) is needed as we need to deal with decorrelate the events of hitting to annuli sets between each blocks, and (A1) is not sufficient as some of the indicator functions after composing with the dynamic T can have large BV norms.

Lemma 4.7 (Space-time Total-Variation approximation). *Under the conditions of Theorem 2.2,*

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) d_{\text{TV}} \left(\mathcal{L} \left((M_{n,k,l}^\Delta)_{1 \leq k \leq K, 1 \leq l \leq L} \right), \mathcal{L} \left((Z_{k,l}^\Delta)_{1 \leq k \leq K, 1 \leq l \leq L} \right) \right) \lesssim \frac{1}{\log n},$$

where ζ_Δ is the centre of $\Delta \in \mathcal{U}_m$, and the random variables $Z_{k,l}$ are independent Poisson random variables with parameters

$$\lambda_{k,l} = 2\rho(\zeta_\Delta)(t_k - t_{k-1}) \int_{z_{l-1}}^{z_l} e^{-z} dz.$$

Proof. Proof of this lemma splits into four steps, with similar ideas used in [CC], namely coupling the relevant random variables with independent Bernoulli-like random variables.

Step 1: Breaking into gaps.

We write $b_n = (\log n)^3$, $\ell_n = (\log n)^2$ and $c_n = b_n + \ell_n$, and set $k_n = \left\lfloor \frac{n}{c_n} \right\rfloor$. We then partition $1, \dots, n$ into blocks

$$B_i = \{(i-1)c_n + 1, \dots, (i-1)c_n + b_n\}, \quad 1 \leq i \leq k_n,$$

with gaps of length ℓ_n between consecutive blocks. For $1 \leq k \leq K$ and $1 \leq l \leq L$. Let $\zeta \in \mathcal{X}$, define the block counts

$$S_{i,k,l} = \sum_{j \in B_i \cap I_k} \mathbf{1}_{A_l(\zeta)}(T^j x).$$

Thus we can write

$$M_{n,k,l} = \sum_{i=1}^{k_n} S_{i,k,l} + R_{n,k,l},$$

where $R_{n,k,l} = \sum_{j \notin \cup_{i=1}^{k_n} B_i} \mathbf{1}_{I_k}(j) \mathbf{1}_{A_l(\zeta)}(T^j x)$ and I_k as in (4.10).

Let $R_n^{\text{gap}} := \sum_{k,l} R_{n,k,l}$, and summing over k, l , we have, uniformly in ζ ,

$$\sum_{k,l} \mathbb{E}_\mu[R_{n,k,l}] \leq \sum_{j \notin \cup_{i=1}^{k_n} B_i} \mathbb{E}_\mu[\mathbf{1}_{B(\zeta, r_n(z_0))} \circ T^j] \lesssim k_n \ell_n n^{-1} \lesssim \frac{1}{\log n}.$$

Hence, summing over all centres ζ_Δ ,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) d_{\text{TV}} \left(\mathcal{L} \left((M_{n,k,l}^\Delta)_{k,l} \right), \mathcal{L} \left((S_{i,k,l}^\Delta)_{i,k,l} \right) \right) \lesssim \frac{1}{\log n}. \quad (4.12)$$

Step 2: At most one hit within blocks.

Fix $\zeta \in \mathcal{X}$, suppressing the notation dependence on ζ , we set $N_i(\zeta) = \sum_{k=1}^K \sum_{l=1}^L S_{i,k,l}$. Since the annuli sets $A_l(\zeta)$ are disjoint and $\bigcup_l A_l(\zeta) \subset B(\zeta, r_n(z_0))$,

$$\{N_i(\zeta) \geq 2\} \subset \{\exists j, j' : T^j x, T^{j'} x \in B(\zeta, r_n(z_0))\}.$$

By Markov's inequality

$$\mathbb{P}(N_i(\zeta) \geq 2) \leq \frac{1}{2} \sum_{j, j' \in B_i : j \neq j'} \mu \left(B(\zeta, r_n(z_0)) \cap T^{-(j'-j)} B(\zeta, r_n(z_0)) \right).$$

We will bound the averaged probability of having more than one hits to any of these annuli sets within each block B_i , *i.e.*

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{i=1}^{k_n} \mathbb{P}(N_i^\Delta \geq 2) \leq \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{i=1}^{k_n} \sum_{j, j' \in B_i : j < j'} \mathbb{E}_\mu \left[\mathbf{1}_{B(\zeta_\Delta, r_n(z_0))} \circ T^j \cdot \mathbf{1}_{B(\zeta_\Delta, r_n(z_0))} \circ T^{j'} \right].$$

We split the sum according to whether $j' - j \leq \ell_n$ or $j' - j > \ell_n$.

Short returns: $j' - j \leq \ell_n$.

For fixed ζ ,

$$\sum_{\substack{j, j' \in B_i \\ 0 < j' - j \leq \ell_n}} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-(j'-j)} B(\zeta, r_n(z_0)) \right) \leq b_n \sum_{k=1}^{\ell_n} \mu \left(B(\zeta, r_n(z_0)) \cap T^{-k} B(\zeta, r_n(z_0)) \right).$$

Summing over all k_n blocks of digits and averaging over $\Delta \in \mathcal{U}_m$,

$$\begin{aligned} & \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{i=1}^{k_n} \sum_{\substack{j, j' \in B_i \\ 0 < j' - j \leq \ell_n}} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-(j'-j)} B(\zeta_\Delta, r_n(z_0)) \right) \\ & \leq k_n b_n \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{j=1}^{\ell_n} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-j} B(\zeta_\Delta, r_n(z_0)) \right). \end{aligned}$$

The function $\zeta \mapsto \sum_{j=1}^{\ell_n} \mu \left(B(\zeta, r_n(z_0)) \cap T^{-j} B(\zeta, r_n(z_0)) \right)$ is Lipschitz with constant $\lesssim \ell_n \|\rho\|_\infty$. Hence by mean value theorem, the Riemann sum over the partition $\mathcal{U}_m$ of diameter $\delta = n^{-(1+\beta')}$ satisfies

$$\begin{aligned} & \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{k=1}^{\ell_n} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-k} B(\zeta_\Delta, r_n(z_0)) \right) \\ & \leq \int \sum_{k=1}^{\ell_n} \mu \left(B(\zeta, r_n(z_0)) \cap T^{-k} B(\zeta, r_n(z_0)) \right) d\mu(\zeta) + \mathcal{O} \left((\log n)^2 n^{-(1+\beta')} \right). \end{aligned}$$

By Lemma 2.1,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{k=1}^{\ell_n} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-k} B(\zeta_\Delta, r_n(z_0)) \right) \lesssim \frac{1}{n \log n}.$$

Multiplying by $k_n b_n \lesssim n$,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{i=1}^{k_n} \sum_{\substack{j, j' \in B_i \\ 0 < j' - j \leq \ell_n}} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-(j' - j)} B(\zeta_\Delta, r_n(z_0)) \right) \lesssim \frac{1}{\log n}. \quad (4.13)$$

Long returns: $j' - j > \ell_n$.

For any fixed $\zeta \in \mathcal{X}$, by (A1) decay of correlations,

$$\begin{aligned} & \mu \left(B(\zeta, r_n(z_0)) \cap T^{-(j' - j)} B(\zeta, r_n(z_0)) \right) \\ & \leq \mu(B(\zeta, r_n(z_0)))^2 + C_T \|\mathbf{1}_{B(\zeta, r_n(z_0))}\|_{BV} \|\mathbf{1}_{B(\zeta, r_n(z_0))}\|_{L^1(\mu)} e^{-\eta(j' - j)}. \end{aligned}$$

Since $\mu(B(\zeta, r_n(z_0))) \lesssim n^{-1}$, and $\|\mathbf{1}_{B(\zeta, r_n(z_0))}\|_{BV} \lesssim 1$, this is uniformly (in ζ) bounded by $n^{-2} + n^{-1}e^{-\eta(j' - j)}$. Summing over at most b_n^2 pairs within each B_i block and averaging over $\Delta \in \mathcal{U}_m$,

$$\begin{aligned} & \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{i=1}^{k_n} \sum_{j < j' \in B_i: j' - j > \ell_n} \mu \left(B(\zeta_\Delta, r_n(z_0)) \cap T^{-(j' - j)} B(\zeta_\Delta, r_n(z_0)) \right) \\ & \lesssim k_n b_n^2 \left(n^{-2} + n^{-1} e^{-\eta(\log n)^2} \right) \lesssim (\log n)^3 n^{-1}. \end{aligned} \quad (4.14)$$

Combining (4.13) and (4.14),

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \sum_{i=1}^{k_n} \mathbb{P}(N_i^\Delta \geq 2) \lesssim \frac{1}{\log n} + \frac{(\log n)^3}{n} \lesssim \frac{1}{\log n}. \quad (4.15)$$

Now define the random $k \times l$ vector $S_i^\Delta := (S_{i,k,l}^\Delta)_{k,l}$, and set $\beta_i^\Delta = S_i^\Delta \mathbf{1}_{\{N_i(\zeta_\Delta) \leq 1\}}$. Then β_i^Δ takes values in

$$\{0\} \cup \{e_{k,l} : 1 \leq k \leq K, 1 \leq l \leq L\},$$

where $e_{k,l}$ denotes the corresponding coordinate matrix. By the coupling characterisation of total variation,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) d_{TV} \left(\mathcal{L}((S_i^\Delta)_i), \mathcal{L}((\beta_i^\Delta)_i) \right) \lesssim \frac{1}{\log n}. \quad (4.16)$$

Step 3: Decorrelation of the block variables. Let β_i denote the random variable with the same construction as above, only substituting an arbitrary $\zeta \in \mathcal{X}$ for a centre of a partition set $\zeta_\Delta \in \Delta$. We will show that, uniformly in ζ ,

$$d_{TV} \left(\mathcal{L}(\beta_1, \dots, \beta_{k_n}), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i) \right) \lesssim (\log n) n^{-2}. \quad (4.17)$$

The decay of correlation condition (A1) is not quite enough to prove (4.17) because the indicator function of $\mathbf{1}\{\beta_i = s\}$ may have very large BV norm. The main tool to bypass such difficulty is symbolic coding and using cylinders to approximate the annulus sets $A_l(\zeta)$ by finite-depth cylinder unions. Then, we can apply the mixing property which takes account of all events in the relative sigma-algebras.

Step 3(a): Cylinder approximation of the annuli.

The following technique is inspired by [BJK]. Let $\mathcal{I}$ be the Markov partition of T as described at the beginning of § 2, and let $\mathcal{I}^{(m)}$ denote the partition into cylinders of depth m . By the uniform expansion assumption there exist constants $C_2 > 0$ and $v > 1$ such that

$$\sup_{I \in \mathcal{I}^{(m)}} \text{diam}(I) \leq C_2 v^{-m}, \quad m \geq 1,$$

and the invariant measure μ has a bounded density. For $\gamma > 0$, consider the stopping-cylinder partition

$$\mathcal{W}(\gamma) := \left\{ I \in \mathcal{I}^{(m)} : \text{diam}(I) \leq \gamma < \text{diam}(I_-) \right\},$$

where I_- denotes the corresponding parent cylinder obtained by deleting the last symbol. It is easy to see that $\mathcal{W}(\gamma)$ is a partition of the repeller. Moreover, there exists $m(\gamma) = \mathcal{O}(\log(1/\gamma))$ such that the symbolic representation of every cylinder set in $\mathcal{W}(\gamma)$ has length at most $m(\gamma)$.

For each annulus set $A_l(\zeta)$, define its inner approximation by

$$A_{l,n}(\zeta) := \bigcup_{\substack{I \in \mathcal{W}(\gamma_n) \\ I \subset A_l(\zeta)}} I, \quad \gamma_n := n^{-3}.$$

Although this union may be countably infinite, any cylinder set appearing in $A_{l,n}(\zeta)$ is determined by the first m_n symbols in Σ , where

$$m_n := m(\gamma_n) \lesssim \log n.$$

Since each $A_l(\zeta)$ has a uniformly bounded number of boundary points, every cylinder contributing to $A_l(\zeta) \setminus A_{l,n}(\zeta)$ must meet one of those boundary points. As each stopping cylinder has diameter at most γ_n , boundedness of the density ρ yields

$$\mu(A_l(\zeta) \triangle A_{l,n}(\zeta)) \lesssim \gamma_n,$$

uniformly in ζ and l . Write $\beta_i^{(n)}$ for the block vector obtained from β_i by replacing every indicator $\mathbf{1}_{A_l(\zeta)} \circ T^j$ by $\mathbf{1}_{A_{l,n}(\zeta)} \circ T^j$. Since $L = \log n$, invariance of μ and the union bound give

$$\begin{aligned} & \mu \left(\exists j \leq n, l \leq L : \mathbf{1}_{A_l(\zeta)}(T^j x) \neq \mathbf{1}_{A_{l,n}(\zeta)}(T^j x) \right) \\ & \leq \sum_{j=1}^n \sum_{l=1}^L \mu(A_l(\zeta) \triangle A_{l,n}(\zeta)) \leq C n \gamma_n L \lesssim (\log n) n^{-2}. \end{aligned} \tag{4.18}$$

Step 3(b): Decorrelation of the approximating block variables. The elements in each approximating target set are determined by at most m_n symbols. Write $\mathcal{G}_p^q$ as the σ -algebra generated by the set $\sigma^p(\mathcal{C}_q)$, where $\mathcal{C}_q$ is the set of q -cylinders. The approximating block vector $\beta_i^{(n)}$ takes values in the finite set $\{0\} \cup \{e_{k,l} : 1 \leq k \leq K, 1 \leq l \leq L\}$, and is measurable with respect to

$$\mathcal{F}_i := \mathcal{G}_{s_i+1}^{s_i+b_n+m_n}, \quad s_i = (i-1)(b_n + \ell_n).$$

Consecutive σ -algebras are separated by

$$w_n := \ell_n - m_n - 1 \asymp (\log n)^2.$$

Let

$$\mathcal{L}_i := \mathcal{L}(\beta_1^{(n)}, \dots, \beta_i^{(n)}) \otimes \bigotimes_{j=i+1}^{k_n} \mathcal{L}(\beta_j^{(n)}), \quad 0 \leq i \leq k_n.$$

Then $\mathcal{L}_0 = \bigotimes_{j=1}^{k_n} \mathcal{L}(\beta_j^{(n)})$ is the product of the marginals and $\mathcal{L}_{k_n} = \mathcal{L}(\beta_1^{(n)}, \dots, \beta_{k_n}^{(n)})$ is the joint law.

Fix $1 \leq i \leq k_n$ and a bounded test function h with $\|h\|_\infty \leq 1$. Define

$$\Phi(\beta_1^{(n)}, \dots, \beta_i^{(n)}) := \mathbb{E}_{\bigotimes_{j>i} \mathcal{L}(\beta_j^{(n)})} \left[h(\beta_1^{(n)}, \dots, \beta_i^{(n)}; \beta_{i+1}^{(n)}, \dots, \beta_{k_n}^{(n)}) \right].$$

Then

$$\mathbb{E}_{\mathcal{L}_i}[h] = \int \Phi d\mathcal{L}(\beta_1^{(n)}, \dots, \beta_i^{(n)}), \quad \mathbb{E}_{\mathcal{L}_{i-1}}[h] = \int \Phi d \left[\mathcal{L}(\beta_1^{(n)}, \dots, \beta_{i-1}^{(n)}) \otimes \mathcal{L}(\beta_i^{(n)}) \right].$$

The two laws differ precisely by the correlation between the σ -algebras $\mathcal{A}_1^{i-1} := \sigma(\beta_1^{(n)}, \dots, \beta_{i-1}^{(n)})$ and $\mathcal{A}_i^i := \sigma(\beta_i^{(n)})$. These are generated by cylinder sets separated by w_n symbols, so the ψ -mixing condition (A1+) gives

$$d_{\text{TV}} \left(\mathcal{L}(\beta_1^{(n)}, \dots, \beta_i^{(n)}), \mathcal{L}(\beta_1^{(n)}, \dots, \beta_{i-1}^{(n)}) \otimes \mathcal{L}(\beta_i^{(n)}) \right) \leq \psi(w_n) \lesssim e^{-\eta'(\log n)^2}.$$

Since $\|\Phi\|_\infty \leq 1$,

$$|\mathbb{E}_{\mathcal{L}_i}[h] - \mathbb{E}_{\mathcal{L}_{i-1}}[h]| \leq \psi(w_n),$$

and taking the supremum over h yields

$$d_{\text{TV}}(\mathcal{L}_i, \mathcal{L}_{i-1}) \leq \psi(w_n) \lesssim n^{-3}.$$

Summing over i and applying the triangle inequality,

$$d_{\text{TV}} \left(\mathcal{L}(\beta_1^{(n)}, \dots, \beta_{k_n}^{(n)}), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i^{(n)}) \right) \leq \sum_{i=1}^{k_n} d_{\text{TV}}(\mathcal{L}_i, \mathcal{L}_{i-1}) \leq k_n \psi(w_n) \lesssim n^{-2}. \quad (4.19)$$

This estimate is uniform in ζ .

Step 3(c): Transferring back to the original block variables.

By the triangle inequality,

$$d_{\text{TV}} \left(\mathcal{L}(\beta_1, \dots, \beta_{k_n}), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i) \right) \leq T_1 + T_2 + T_3,$$

where

$$\begin{aligned} T_1 &:= d_{\text{TV}} \left(\mathcal{L}(\beta_1, \dots, \beta_{k_n}), \mathcal{L}(\beta_1^{(n)}, \dots, \beta_{k_n}^{(n)}) \right), \\ T_2 &:= d_{\text{TV}} \left(\mathcal{L}(\beta_1^{(n)}, \dots, \beta_{k_n}^{(n)}), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i^{(n)}) \right), \\ T_3 &:= d_{\text{TV}} \left(\bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i^{(n)}), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i) \right). \end{aligned}$$

Bound on T_1 . The two vectors differ only if, for some $j \leq n$ and some $l \leq L$, the orbit hits the symmetric difference $A_l(\zeta) \triangle A_{l,n}(\zeta)$. Hence, by (4.18),

$$T_1 \leq \mu \left(\exists j \leq n, l \leq L : \mathbf{1}_{A_l(\zeta)}(T^j x) \neq \mathbf{1}_{A_{l,n}(\zeta)}(T^j x) \right) \lesssim (\log n) n^{-2}.$$

Bound on T_2 . This is the mixing term, bounded in (4.19): $T_2 \lesssim n^{-2}$.

Bound on T_3 . For products of marginals, the total-variation distance is bounded by the sum of the marginal distances:

$$T_3 \leq \sum_{i=1}^{k_n} d_{\text{TV}} \left(\mathcal{L}(\beta_i^{(n)}), \mathcal{L}(\beta_i) \right) \leq \sum_{i=1}^{k_n} \mu(\beta_i^{(n)} \neq \beta_i).$$

Each β_i depends on the orbit over the block B_i of length b_n . The two differ only if the orbit hits a symmetric difference at some time in B_i . Hence

$$\mu(\beta_i^{(n)} \neq \beta_i) \leq \sum_{j \in B_i} \sum_{l=1}^L \mu(A_l(\zeta) \Delta A_{l,n}(\zeta)) \leq b_n L C \gamma_n.$$

Summing over $i = 1, \dots, k_n$, and as $k_n b_n \asymp n$ and $\gamma_n = n^{-3}$, as we chose $L = \log n$,

$$T_3 \leq k_n \cdot b_n \cdot L \cdot C \gamma_n \lesssim (\log n) n^{-2}.$$

Combining the three bounds, for any $\zeta \in \mathcal{X}$,

$$d_{\text{TV}} \left(\mathcal{L}(\beta_1, \dots, \beta_{k_n}), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i) \right) \leq T_1 + T_2 + T_3 \lesssim (\log n) n^{-2}.$$

Averaging over the partition $\mathcal{U}_m$,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) d_{\text{TV}} \left(\mathcal{L}(\beta_1^\Delta, \dots, \beta_{k_n}^\Delta), \bigotimes_{i=1}^{k_n} \mathcal{L}(\beta_i^\Delta) \right) \lesssim (\log n) n^{-2}. \quad (4.20)$$

Step 3(d): Passage to the sums. Lastly, let $\tilde{\beta}_1, \dots, \tilde{\beta}_{k_n}$ be independent random vectors with $\tilde{\beta}_i \stackrel{d}{=} \beta_i$. As the map $(x_1, \dots, x_{k_n}) \mapsto \sum_{i=1}^{k_n} x_i$ is measurable and contracting with respect to the metric d_{TV} , by (4.20)

$$\sum_{\Delta} \mu(\Delta) d_{\text{TV}} \left(\mathcal{L} \left(\sum_{i=1}^{k_n} \beta_i^\Delta \right), \mathcal{L} \left(\sum_{i=1}^{k_n} \tilde{\beta}_i^\Delta \right) \right) \lesssim (\log n) n^{-2}.$$

Step 4: Multivariate Poisson approximation. Now, suppress dependence on $\Delta \in \mathcal{U}_m$ again, define β_i for an arbitrary $\zeta \in \mathcal{X}$ analogously as in step 3. For each block i , set $p_i^{(k,l)} = \mathbb{P}(\beta_i = e_{k,l})$ and $p_i = \sum_{k=1}^K \sum_{l=1}^L p_i^{(k,l)}$. Since a block can contain a relevant point only when $T^j x \in B(\zeta, r_n(z_0))$ for some $j \in B_i$, so again by Markov's inequality, for all i , up to a uniform constant that depends only on ρ and z_0 ,

$$p_i \leq \mathbb{P} \left[\sum_{j \in B_i} \mathbf{1}_{B(\zeta, r_n(z_0))} \geq 1 \right] \leq \sum_{j \in B_i} \mu(B(\zeta, r_n(z_0))) \lesssim \frac{b_n}{n}.$$

Therefore

$$\sum_{i=1}^{k_n} p_i^2 \leq k_n \max_i p_i^2 \lesssim k_n \frac{b_n^2}{n^2} \lesssim \frac{(\log n)^3}{n}.$$

By the multivariate Le Cam theorem,

$$d_{\text{TV}} \left(\mathcal{L} \left(\sum_i \beta_i \right), \mathcal{L}((Z_{k,l}^{(n)})_{k,l}) \right) \lesssim \sum_i p_i^2 \lesssim \frac{(\log n)^3}{n},$$

where the $Z_{k,l}$ are independent Poisson random variables with

$$\mathbb{E}Z_{k,l} = \lambda_{k,l}^{(n)} = \sum_{i=1}^{k_n} p_i^{(k,l)}.$$

We now calculate these means. First, notice that

$$p_i^{(k,l)} = \mathbb{P}(S_{i,k,l} = 1, N_i \leq 1) = \mathbb{P}(S_{i,k,l} = 1) - \mathbb{P}(S_{i,k,l} = 1, N_i \geq 2),$$

and by the formula $\mathbb{E}[X] = \sum_{q=0}^{\infty} \mathbb{P}(X \geq q) = \sum_{q=1}^{\infty} q \mathbb{P}(X = q)$ for all integer-valued non-negative random variables,

$$\mathbb{P}(S_{i,k,l} = 1) = \mathbb{E}_{\mu}[S_{i,k,l}] - \mathbb{E}_{\mu}[S_{i,k,l} \mathbf{1}_{\{S_{i,k,l} \geq 2\}}],$$

we have

$$\left| p_i^{(k,l)} - \mathbb{E}_{\mu}[S_{i,k,l}] \right| \leq \mathbb{E}_{\mu}[S_{i,k,l} \mathbf{1}_{\{S_{i,k,l} \geq 2\}}] + \mathbb{P}(S_{i,k,l} = 1, N_i \geq 2).$$

The latter term is trivially bounded by $\mathbb{P}(N_i \geq 2)$ and for the first term, notice that

$$S_{i,k,l} \mathbf{1}_{\{S_{i,k,l} \geq 2\}} \leq S_{i,k,l}(S_{i,k,l} - 1) = \sum_{j \neq j' \in B_i \cap I_k} \mathbf{1}_{A_l(\zeta)} \circ T^j \cdot \mathbf{1}_{A_l(\zeta)} \circ T^{j'},$$

hence we can bound the expectation of such product by

$$\mathbb{E}_{\mu}[S_{i,k,l}(S_{i,k,l} - 1)] \leq \sum_{j \neq j' \in B_i} \mu \left(T^{-j} B(\zeta, r_n(z_0)) \cap T^{-j'} B(\zeta, r_n(z_0)) \right)$$

which was calculated in Step 2 already. Therefore, summing over all k_n blocks we get

$$\left| \lambda_{k,l}^{(n)} - \sum_{i=1}^{k_n} \mathbb{E}_{\mu}[S_{i,k,l}] \right| \lesssim \frac{(\log n)^3}{n}.$$

Obviously, for all k and all l , $\sum_{i=1}^{k_n} S_{i,k,l}$ and $M_{n,k,l}^{\zeta}$ differ in at most $\min\{\#I_k, k_n \ell_n\}$ terms, so

$$\left| \mathbb{E}_{\mu} \left[\sum_{i=1}^{k_n} S_{i,k,l} \right] - \mathbb{E}_{\mu}[M_{n,k,l}] \right| \lesssim \frac{\#I_k}{n}.$$

It remains to compare $n\mu(A_l(\zeta))$ with $2\rho(\zeta) \int_{z_{l-1}}^{z_l} e^{-z} dz$. Define

$$F_n(z, \zeta) := n\mu \left(B \left(\zeta, \frac{e^{-z}}{n} \right) \right), \quad z \in [a, b]$$

where $[a, b]$ is contained in the support of the g in its z -coordinate. Then $n\mu(A_l(\zeta)) = F_n(z_{l-1}, \zeta) - F_n(z_l, \zeta)$. By the logarithmic continuity assumption (LC) on ρ , uniformly for $z \in [a, b]$ and for μ -a.e. ζ ,

$$F_n(z, \zeta) = 2\rho(\zeta)e^{-z} + \mathcal{E}_n(z, \zeta),$$

where $|\mathcal{E}_n(z, \zeta)| \lesssim 1/\log n$. Moreover, $\mathcal{E}_n(\cdot, \zeta)$ has bounded variation:

$$\text{Var}_{[a,b]} \mathcal{E}_n(\cdot, \zeta) \lesssim \frac{1}{\log n}.$$

Indeed, by (LC)

$$\left| \frac{d}{dz} \mathcal{E}_n(z, \zeta) \right| = \left| -e^{-z} \left[\rho \left(\zeta + \frac{e^{-z}}{n} \right) + \rho \left(\zeta - \frac{e^{-z}}{n} \right) - 2\rho(\zeta) \right] \right| \lesssim \frac{1}{\log n},$$

Hence

$$\sum_{l=1}^L \left| n\mu(A_l(\zeta)) - 2\rho(\zeta) \int_{z_{l-1}}^{z_l} e^{-z} dz \right| = \sum_{l=1}^L |\mathcal{E}_n(z_{l-1}, \zeta) - \mathcal{E}_n(z_l, \zeta)| \lesssim \frac{1}{\log n}.$$

Multiplying by $t_k - t_{k-1}$ and summing over k gives

$$\sum_{k=1}^K \sum_{l=1}^L \left| \lambda_{k,l}^{(n)} - 2\rho(\zeta)(t_k - t_{k-1}) \int_{z_{l-1}}^{z_l} e^{-z} dz \right| \lesssim \frac{1}{\log n}.$$

Thus, if $Z_{k,l}$ are independent Poisson random variables with means $\lambda_{k,l} = 2\rho(\zeta)(t_k - t_{k-1}) \int_{z_{l-1}}^{z_l} e^{-z} dz$, then

$$d_{\text{TV}} \left(\mathcal{L} \left((Z_{k,l}^{(n)})_{k,l} \right), \mathcal{L} \left((Z_{k,l})_{k,l} \right) \right) \leq \sum_{k,l} |\lambda_{k,l}^{(n)} - \lambda_{k,l}| \lesssim \frac{1}{\log n}.$$

This bound is uniform in ζ .

Combining the four steps for fixed ζ , and averaging over $\Delta \in \mathcal{U}_m$, for $Z_{k,l}^\Delta$ each with mean $\lambda_{k,l}^\Delta$

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) d_{\text{TV}} \left(\mathcal{L} \left((M_{n,k,l}^\Delta)_{k,l} \right), \mathcal{L} \left((Z_{k,l}^\Delta)_{k,l} \right) \right) \lesssim \frac{1}{\log n} + \frac{(\log n)^3}{n} + \frac{1}{\log n} \lesssim \frac{1}{\log n}. \quad \square$$

Proof of Proposition 4.4. Recall the notation convention: a quantity carrying a Δ superscript is evaluated at $\zeta = \zeta_\Delta$ which is the centre of a partition set $\Delta \in \mathcal{U}_m$. One without Δ is evaluated at an arbitrary $\zeta \in \mathcal{X}$.

For each $\Delta \in \mathcal{U}_m$, let $Z_{k,\ell}^\Delta$ be the independent Poisson random that appeared in Lemma 4.7 with expectation

$$\lambda_{k,\ell}^{(n),\Delta} = \sum_{i=1}^{k_n} p_i^{(k,\ell)}(\zeta_\Delta) = \mathbb{P}(S_{i,k,\ell}^\Delta = 1) - \mathbb{P}(S_{i,k,\ell}^\Delta = 1, N_i^\Delta \geq 2).$$

For any two point processes ξ_1, ξ_2 and any measurable function h on the space of point measures,

$$|\mathbb{E}[h(\xi_1)] - \mathbb{E}[h(\xi_2)]| \leq \|h\|_\infty \cdot d_{\text{TV}}(\mathcal{L}(X), \mathcal{L}(Y)). \quad (4.21)$$

Take $\xi_1 = N_n^\Delta$ and ξ_2 a point process whose law coincides with $\mathcal{L}((Z_{k,\ell}^\Delta)_{k,\ell})$, and set

$$F(\xi) := \exp \left(- \int g d\xi \right),$$

for some non-negative C^1 function with compact support (hence $\|F\|_\infty \leq 1$). By Lemma 4.7, as N_n^Δ is a combination of $M_{n,k,l}^\Delta$,

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \left| \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] - \mathbb{E} \left[e^{-\sum_{k,\ell} g(t_{k-1}, z_{l-1}) Z_{k,\ell}^\Delta} \right] \right| \lesssim \frac{1}{\log n}. \quad (4.22)$$

For independent Poisson random variables $Z_{k,\ell}^\Delta$ with means $\lambda_{k,\ell}^{(n),\Delta}$, a standard computation (for example, see [R2]) gives

$$\mathbb{E} \left[\exp \left(- \sum_{k,\ell} g(t_{k-1}, z_{l-1}) Z_{k,\ell}^\Delta \right) \right] = \prod_{k,\ell} \exp \left(- \lambda_{k,\ell}^{(n),\Delta} \left(1 - e^{-g(t_{k-1}, z_{l-1})} \right) \right).$$

Taking logarithms,

$$\log \mathbb{E} \left[\exp \left(- \sum_{k,\ell} g(t_{k-1}, z_{l-1}) Z_{k,\ell}^\Delta \right) \right] = - \sum_{k,\ell} \lambda_{k,\ell}^{(n),\Delta} \left(1 - e^{-g(t_{k-1}, z_{l-1})} \right).$$

By the proof of Lemma 4.7, uniformly in $\zeta \in \mathcal{X}$ up to a measure zero set,

$$\lambda_{k,\ell}^{(n)} = 2\rho(\zeta)(t_k - t_{k-1}) \int_{z_{\ell-1}}^{z_\ell} e^{-z} dz + \mathcal{O}\left(\frac{1}{\log n}\right).$$

Applying this at $\zeta = \zeta_\Delta$, substituting into the logarithm of the expectation above, as $n \rightarrow \infty$, up to an error term $\lesssim 1/\log n$ that is uniform in Δ ,

$$-\sum_{k,\ell} \lambda_{k,\ell}^{(n),\Delta} \left(1 - e^{-g(t_{k-1}, z_{l-1})}\right) \rightarrow -2\rho(\zeta_\Delta) \sum_{k,\ell} (t_k - t_{k-1}) \int_{z_{\ell-1}}^{z_\ell} \left(1 - e^{-g(t_{k-1}, z_{l-1})}\right) e^{-z} dz.$$

The double sum is a Riemann sum for the integral

$$-2\rho(\zeta_\Delta) \int_0^1 \int_{-\infty}^{\infty} \left(1 - e^{-\tilde{g}(t,z)}\right) e^{-z} dz dt.$$

Hence, uniformly in Δ ,

$$\left| \mathbb{E} \left[e^{-\sum_{k,\ell} g(t_{k-1}, z_{l-1}) Z_{k,\ell}^\Delta} \right] - \Psi_\infty^g(\zeta_\Delta) \right| \lesssim \frac{1}{\log n}. \quad (4.23)$$

For the final step, by the triangle inequality,

$$\begin{aligned} & \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \left| \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] - \Psi_\infty^g(\zeta_\Delta) \right| \\ & \leq \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \left| \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] - \mathbb{E} \left[e^{-\sum_{k,\ell} g(t_{k-1}, z_{l-1}) Z_{k,\ell}^\Delta} \right] \right| \\ & \quad + \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \left| \mathbb{E} \left[e^{-\sum_{k,\ell} g(t_{k-1}, z_{l-1}) Z_{k,\ell}^\Delta} \right] - \Psi_\infty^g(\zeta_\Delta) \right|. \end{aligned}$$

By (4.22), the first sum is $\lesssim 1/\log n$. By (4.23), the second term $\lesssim \frac{1}{\log n} \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) = \frac{1}{\log n}$. Therefore

$$\sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \left| \mathbb{E}_\mu \left[e^{-\int g dN_n^\Delta} \right] - \Psi_\infty^g(\zeta_\Delta) \right| \lesssim \frac{1}{\log n} \quad \square$$

Finally, by Proposition 4.4, the formula $\mathbb{E}_\mu[e^{-\int g d\tilde{N}_n}] = \sum_{\Delta \in \mathcal{U}_m} \mu(\Delta) \mathbb{E}_\mu[e^{-\int g dN_n^\Delta}]$, and Proposition 4.1, we may conclude with the following.

Proposition 4.8. *For all C^1 non-negative test functions with compact support $g = g(t, z)$, for μ -almost every ζ that occurs as the centre of $\Delta \in \mathcal{U}_m$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{E}_\mu \left[e^{-\int g dN_n} \right] = \int \exp \left(- \iint \left(1 - e^{-g(t,z)} \right) 2\rho(x) e^{-z} dz dt \right) \rho(x) dx.$$

This validates Theorem 2.3.

5. PROOF OF THEOREM 2.2

We will prove Theorem 2.2 from Theorem 2.3 via applying the Continuous Mapping Theorem [Bi], following the machinery comprehensively established in [R2, §4] and [HT1, §5]. We give an extremely brief introduction here. Let (M, d) and (M', d') be two metric spaces, and a mapping $h : M \rightarrow M'$ that is almost surely continuous, which we define below. We call an element $\xi \in M$ a discontinuity with respect to h if there exists a sequence $\{\xi_n\}_n$ such that $d(\xi_n, \xi) \rightarrow 0$ as $n \rightarrow \infty$ while $d'(h(\xi_n), h(\xi)) \not\rightarrow 0$. The map is a.s. continuous if the set of discontinuities has zero μ mass.

For completeness, we can take $d, d' = d_v$ which is compatible with the vague topology and can be constructed via the following: take a countable dense base set $\{f_n\}_n \subset C_c^1((0, 1] \times \mathbb{R})$, and define the metric by

$$d_v(\mu_1, \mu_2) = \sum_{j=1}^{\infty} \frac{1}{2^j} \min\{1, |\mu_1(f_j) - \mu_2(f_j)|\}.$$

It is known that $\mu_n \rightarrow \mu$ vaguely if and only if $d_v(\mu_n, \mu) \rightarrow 0$ as $n \rightarrow \infty$. Applying the CMT will then imply that $h(\xi_n) \xrightarrow{d'} h(\xi)$ for all continuity points. Let us first recall the precise statement of CMT.

Theorem 5.1 (Continuous Mapping Theorem, [Bil]). *Let (M_1, d_1) and (M_2, d_2) be metric spaces and ξ_n, ξ are random elements on $(M_1, \mathcal{B}(M_1))$. Suppose that $\xi_n \rightarrow \xi$ weakly. If a function $h : M_1 \rightarrow M_2$ satisfies*

$$\mathbb{P}(\xi \in \{x_0 \in M_1 : x_0 \text{ is a discontinuity point for } h\}) = 0$$

then $h(\xi_n) \rightarrow h(\xi)$ weakly as well.

In the proceeding proof of Theorem 2.2, we will take $M_1 = \mathcal{M}_p((0, 1] \times \mathbb{R})$ and $M_2 = \mathcal{M}_p(\mathbb{R})$, where $\mathcal{M}_p(X)$ denotes the set of locally finite measures on a space X and is equipped with the vague topology. A notable difference between the results presented in this paper and those of [HT1] is that we are not seeking to establish a convergence of the partial maxima process $M_n(t)$ to a continuous classical extremal process $Y_G(t)$ generated by one of the three canonical extremal distribution G introduced in §1.1. In fact, it should be clear from [Col] that the distribution of partial maximum of the return process $Y_j(x) := -\log d(T^j x, x)$ converges in limit to a averaged-Gumbel law with density $g_x(z) = \rho(x)e^{-2\rho(x)z}$.

The following lemma verifies the conditions for applying the CMT. Define the functions which retrieve the behaviour of records: let $E = (0, 1] \times \mathbb{R}$, and $\xi = \sum_i \delta_{(t_i, z_i)}$, let

$$\begin{aligned} H_1 : \mathcal{M}_p(E) &\rightarrow \mathbb{D}(\mathbb{R}), & H_1(\xi)(t) &:= \sup\{z_i : t_i \leq t\}; \\ H_2 : \mathcal{M}_p(E) &\rightarrow \mathbb{D}(\mathbb{R}), & H_2(\xi)(z) &:= \inf\{t_i : z_i > z\}; \\ H_3 : \mathbb{D}(\mathbb{R}) &\rightarrow \mathcal{M}_p(\mathbb{R}), & H_3(y) &:= \sum_{s: y(s^-) \neq y(s^+)} \delta_s. \end{aligned}$$

Both $h_1 := H_3 \circ H_1$ and $h_2 := H_3 \circ H_2$ are functions from $\mathcal{M}_p(E)$ to $\mathcal{M}_p(\mathbb{R})$. Our first step is to verify that h_1, h_2 are almost surely continuous with respect to $\mathcal{N}$.

Lemma 5.2. *Let $E = (0, 1] \times \mathbb{R}$ and let $\mathcal{N}$ the limiting Cox point process as in Theorem 2.3. Let D_R, D_V be the set of discontinuities for h_1 and h_2 respectively. Then*

$$\mathbb{P}(\mathcal{N} \in D_R) = 0 \quad \text{and} \quad \mathbb{P}(\mathcal{N} \in D_V) = 0.$$

Proof. For convenience, we recall the Laplace functional of $\mathcal{N}$

$$\mathbb{E}[e^{-\int g d\mathcal{N}}] = \int_X \exp\left(-\int_E (1 - e^{-g(t, z)}) 2\rho(x) e^{-z} dt dz\right) \rho(x) dx,$$

For Borel sets $A, B \subset E$, the *second factorial moment measure* of $\mathcal{N}$ is

$$M^{(2)}(A \times B) = \mathbb{E}\left[\sum_{i \neq j} \mathbf{1}_{\{X_i \in A\}} \mathbf{1}_{\{X_j \in B\}}\right] = \int_X \Lambda_x(A) \Lambda_x(B) \rho(x) dx. \quad (5.1)$$

where the sum in the middle is over ordered pairs of distinct points $\delta_{X_i} = \delta_{(t_i, z_i)} \in \mathcal{N}$. The second identity follows from the Laplace functional with $g = \lambda \mathbf{1}_A + \lambda' \mathbf{1}_B$, then differentiating against λ, λ'

at $\lambda, \lambda' = 0$. In our case, as $\Lambda_x(dt dz) = 2\rho(x)e^{-z}dt dz$, the product measure $\Lambda_x \otimes \Lambda_x$ has density $4\rho(x)^2e^{-z-z'}dtdz dt'dz$ which is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^4$.

Define the set of pairs with the same time coordinate:

$$D_{\text{time}} = \{(t, z, t', z') \in E \times E : t = t'\} \subset \mathbb{R}^4.$$

This set has Lebesgue measure zero in $\mathbb{R}^4$, because it is contained in a three dimensional hyperplane of points satisfying $t = t'$. By absolute continuity of $\Lambda_x \otimes \Lambda_x$ with respect to the Lebesgue measure, we conclude that

$$(\Lambda_x \otimes \Lambda_x)(D_{\text{time}}) = 0 \quad \text{for every } x.$$

Substituting into (5.1),

$$M^{(2)}(D_{\text{time}}) = \int_{\mathcal{X}} (\Lambda_x \otimes \Lambda_x)(D_{\text{time}}) \rho(x) dx = 0.$$

Define random variable

$$S_t := \sum_{i \neq j} \mathbf{1}_{\{t_i = t_j\}}$$

that counts the number of ordered pairs of distinct points of $\mathcal{N}$ with the same t -coordinate. Its expectation is

$$\mathbb{E}[S_t] = M^{(2)}(D_{\text{time}}) = 0.$$

Since $S_t \geq 0$ is integer-valued, $\mathbb{E}[S_t] = 0$ implies $S_t = 0$ almost surely by Markov's inequality. Therefore, almost surely, no two distinct atoms of $\mathcal{N}$ share the same t -coordinate. The proof for almost surely no two atoms have the same z -coordinate follows verbatim.

Next, we show that $\mathcal{N}$ is locally finite. Let $K \subset E$ be a compact set, then by definition, as $\|\rho\|_\infty < \infty$

$$\mathbb{E}[\mathcal{N}(K)] = \int \Lambda_x(K) \rho(x) dx < \infty.$$

As $\mathcal{N}(K)$ is non-negative, this entails $\mathbb{P}(\mathcal{N}(K) < \infty) = 1$. To prove local finiteness, take a countable family of compact sets $\{K_n\}_n$ such that their union covers E and

$$\mathbb{P} \left(\left(\bigcap_n \{\mathcal{N}(K_n) < \infty\} \right)^c \right) \leq \sum_n \mathbb{P}(\mathcal{N}(K_n) = \infty) = 0.$$

This is sufficient to let us conclude with

$$\mathbb{P}(\mathcal{N}(K) < \infty \text{ for all compact } K \subset E) = 1.$$

In particular, for every fixed $t^* \in (0, 1]$ and $z^* \in \mathbb{R}$, we can show that $\mathbb{P}(\mathcal{N}(\{t^*\} \times \mathbb{R}) \geq 1) = 0$ and $\mathbb{P}(\mathcal{N}(0, 1] \times \{z^*\}) \geq 1) = 0$ by direct computation of the expectation then apply Markov's inequality. This means there is no accumulation of atoms in either the t or z coordinate, *i.e.* almost surely

$$\mathcal{N}(\{t^*\} \times \mathbb{R}) = 0 \quad \text{and} \quad \mathcal{N}((0, 1] \times \{z^*\}) = 0.$$

By [R2, §4] and [HT1, §5] what we have proved so far is enough to conclude that $\mathbb{P}(\mathcal{N} \in D_R) = 0$. The proof for $\mathbb{P}(\mathcal{N} \in D_V) = 0$ follows verbatim. $\square$

By Lemma 5.2 we are allowed to apply CMT to conclude that $R_n \rightarrow h_1(\mathcal{N})$ and $V_n \rightarrow h_2(\mathcal{N})$ both in distribution. It only remains to identify the limiting distribution, *i.e.* the distribution of $h_1(\mathcal{N})$ and $h_2(\mathcal{N})$.

Proof. Proof of Theorem 2.2 By definition of $\mathcal{N}$, the law of $\mathcal{N}$ has the form $\mathbb{P}(\mathcal{N}(\cdot) \leq k) = \int \mathbb{P}(\Pi_x(\cdot) \leq k) \rho(x) dx$ for all $k \geq 0$, where Π_x is a Poisson random measure on E with intensity $2\rho(x)e^{-z}dtdz$ (in other words, its intensity measure is $\text{Leb} \times \lambda_{G_x}$ where $G_x(z) = 2\rho(x)e^{-z}$). By measurability of the map h_1 , for all Borel subset $A \subset (0, 1]$ (without loss of generality take $A = (a, b]$)

$$\mathbb{P}((h_1(\mathcal{N}))(A) \leq k) = \int_X \mathbb{P}((h_1(\Pi_x))(A) \leq k) \rho(x) dx \quad (5.2)$$

where by classical results on record statistics via Poisson point processes [RR]³, the number of points in $h_1(\Pi_x)$ on $(a, b]$ is Poisson distributed with mean $\log(b/a)$. As this is independent of x , the right-hand side of (5.2) after integrating over x equals $\mathbb{P}(\mathcal{R}(A) \leq k)$, where $\mathcal{R}$ is a Poisson point process with intensity $1/tdt$. By Lemma 5.2 and the CMT, $h_1(N_n)$ converges in distribution to $h_1(\mathcal{N})$, hence we conclude that the limiting distribution of $\sum_j \delta_{j/n} \mathbf{1}_{\{Y_j > M_{j-1}\}}$ converges to that of a Poisson process with intensity $1/tdt$.

Similarly, using an analogous argument, for any Borel set $B \subset \mathbb{R}$,

$$\mathbb{P}(h_2(\mathcal{N})(B) \leq k) = \int \mathbb{P}(h_2(\Pi_x)(B) \leq k) \rho(x) dx.$$

By [HT1, §5], the distribution of $h_2(\Pi_x)$ has intensity measure

$$\begin{aligned} \lambda_V[a, b] &= -\log(-\log G_x(b)) + \log(-\log G_x(a)) = -\log\left(-\log(e^{-2\rho(x)e^{-b}})\right) + \log\left(-\log(e^{-2\rho(x)e^{-a}})\right) \\ &= b - \log(2\rho(x)) - a + \log(2\rho(x)) = b - a \end{aligned}$$

which is again independent of x . Then, the limiting distribution of $h_2(N_n)$ equals that of a Poisson point process $\mathcal{V}$ on $\mathbb{R}$ with intensity $\lambda[a, b] = b - a$. $\square$

6. REMARKS ON DIFFERENT OBSERVABLES

The main results presented in §2 deal with the observable defined for the value of return distances that results in the averaged-Gumbel type for the limiting distribution of the maximum process. The analogous statement for Fréchet-type holds. Take $Y_j(x) = d(x, T^j x)^{-\alpha}$ and define the associated point process $N_n = \sum_{j=1}^n \delta_{(j/n, n^{-\alpha} Y_j)}$. Then

$$N_n \Rightarrow \mathcal{N}^F, \mathcal{N}^F \text{ a Cox process with random intensity } \Lambda_x(dtdz) = 2\rho(x)\alpha z^{-(1+\alpha)}dtdz.$$

The record-value process V_n converges in distribution to a PRM with intensity λ_F on $(0, \infty)$ given by $\lambda_F[a, b] = \alpha \log(b/a)$. For the Weibull-type, we would define $Y_j(x) = C - d(x, T^j x)^\alpha$ and $N_n = \sum_{j=1}^n \delta_{(j/n, n^\alpha(Y_j - C))}$. Then

$$N_n \Rightarrow \mathcal{N}^W, \mathcal{N}^W \text{ a Cox process with random intensity } \Lambda_x(dtdz) = 2\rho(x)\alpha(-z)^{\alpha-1}dtdz,$$

and V_n converges in distribution to a PRM with intensity Λ_W on satisfying $\lambda_W[a, b] = \alpha \log(b/a)$ as well. The limiting distribution of the record-time process R_n is universal for all three cases, *i.e.* PRM with intensity measure $1/tdt$.

The proofs of these Fréchet and Weibull observable cases follow almost verbatim from those of the Gumbel case without needing to change any assumptions or the proof structure. For the Fréchet case: the proof of Proposition 4.1 relies on using compactness of the support of the test function g in its z coordinate, to argue that the difference of

$$g(j/n, Y_j/n^\alpha) - g(j/n, Y_j^\Delta/n^\alpha)$$

³The original statement in [RR] was stated for homogeneous Poisson point process, where Π_x is inhomogeneous. However, this can be resolved without much difficulty by introducing a change of variable.

is not automatically zero only when x is in some $E_j(r_n)$ set corresponding to return distance r_n . Note that $Y_j(x)/n^\alpha \in [a, b]$ implies $d(T^j x, x) \in [1/nb^{1/\alpha}, 1/na^{1/\alpha}]$, and $Y_j^\Delta(x)/n^\alpha \in [a, b]$ implies $d(T^j x, x) \in [1/nb^{1/\alpha} - \delta, 1/na^{1/\alpha} + \delta]$ where 2δ is the diameter of the partition sets $\Delta \in \mathcal{U}_m$. Similarly for the Weibull case, $n^\alpha(Y_j(x) - C) \in [a, b] \subset (-\infty, 0]$ implies $d(T^j x, x) \in [(-b)^{1/\alpha}/n, (-a)^{1/\alpha}/n]$, and similarly for $n^\alpha(Y_j^\Delta - C)$. Consequently,

$$\|g(j/n, d(\zeta, T^j \cdot)^{-\alpha})\|_{L^1(\mu)} \lesssim n^{-1} \text{ and } \|g(j/n, n^\alpha(d(\zeta, T^j \cdot)^\alpha - C))\|_{L^1(\mu)} \lesssim n^{-1}$$

remain unchanged. The proofs for convergence of the Laplace functionals for the auxiliary process are analogous to that of Proposition 4.4, requiring only changing the parameters of the Poisson variables and the diameters of the annuli sets.

DISCLAIMER OF AI USAGE

This paper used generative AI tools marginally. In particular, no paid plans features were ever employed. The main usage of a large language model (Deepseek) was for

- (1) facilitating literature research (including the literature on Cox process, the precise reference of multivariate Le Cam theorem and the checking the precise requirements to apply the Continuous Mapping Theorem),
- (2) spotting typographical errors, and
- (3) keeping consistency of my notations throughout (especially for the proof of Prop 4.1 and Prop 4.4)

There were 2 places where LLM did have an impactful contribution to my proofs:

- (1) during Step 2 of Lemma 4.7, where AI pointed out the error terms of the Riemann sum can be dealt with via Lipschitz continuity;
- (2) helping me finish the mixing argument in Step 3(b) of the proof of Lemma 4.7.

REFERENCES

- [BJK] B. Bárány, N. Jurga, and I. Kolossváry, *On the convergence rate of the chaos game*, Int. Math. Res. Not. IMRN **(5)** (2023), 4456–4500. 4.3
- [Bar] A. Barbour, *Stein's method and Poisson process convergence*, J. Appl. Probab. **25A** (1988) 175–184. 4.3
- [BS] L. Barreira and B. Saussol, *Hausdorff dimension of measures via Poincaré recurrence*, Comm. Math. Phys., **219** (2001) 443–463. 1
- [Bil] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York (1968). 5, 5.1
- [Bow] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Springer, Berlin Heidelberg (1975). 3, 3.2
- [CC] J. Chazottes and P. Collet, *Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **33** (2013) 49–80. 4.3
- [Col] P. Collet, *Statistics of closest return for some non-uniformly hyperbolic systems* Ergodic Theory and Dynamical Systems, **21**(2) (2001) 401–420. 1, 1.4, 1.8, 5
- [Cox] D. Cox, *Some statistical methods connected with series of events*, J. R. Stat. Soc. **17**(2) (1955) 129–164. 1.2
- [FFV] A.C.M. Freits, J. M. Freitas and S. Vaienti, *Extreme Value Laws for non stationary processes generated by sequential and random dynamical systems*, Annales de l'Institut Henri Poincaré - Probabilités et Statistiques **53**(3) (2017) 1341–1370. 4
- [GHN] C. Gupta, M. Holland and M. Nicol, *Extreme value theory and return time statistics for dispersing billiard maps and flows, Lozi maps and Lorenz-like maps*, Ergodic Theory and Dynam. Systems. **31**(5) (2011) 1363–1390. 1
- [HNT] M. Holland, M. Nicol and A. Török, *Extreme value distributions for non-uniformly expanding dynamical systems*, Trans. Am. Math. Soc. **364** (2012) 661–688. 3, 3.2, 2

- [HKKP] M. Holland, M. Kirsebom, P. Kunde, and T. Persson, *Dichotomy results for eventually always hitting time statistics and almost sure growth of extremes*, Trans. Amer. Math. Soc. **377** (2024) 3927-3982. 1
- [HT1] M. Holland and M. Todd, *Weak convergence to extremal processes and record events for non-uniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems **39** (2019) 980-1001 1.3, 2.4, 5, 5, 5, 5
- [HT2] M. Holland and M. Todd, *On distributional limit laws for recurrence*, Nonlinearity **38**, (2025) no. 075028. 1.4, 1.8, 2.4, 3.2, 4.1
- [HS] P. Holmes and W. Strawderman, *A note on the waiting times between record observations*, J. Appl. Probab. **6**(3) (1969) 711-714. 1.1
- [Kal] O. Kallenberg. Random Measures. Akademie-Verlag, Berlin, 3. 1983. 1.2, 2.4
- [Kar] S. Karlin, *A first course in stochastic processes*, Academic Press, New York, 1966. 1.1
- [KKP] P. Kirsebom, M. Kunde and T. Persson, *On Shrinking Targets and Self-Returning Points*, Ann. Sc. Norm. Super. Pisa Cl. Sci., **24**, (2023) 1499-1535. 1
- [LP] G. Last and M. Penrose, *Lectures on the Poisson Process*. Cambridge University Press (2017). 1.2
- [LLR] M. Leadbetter, G. Lindgren and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes* (Springer Series in Statistics). Springer, New York (1980). 1.1, 1.1
- [LCam] L. Le Cam, *An approximation theorem for the Poisson binomial distribution*, Pacific J. Math. **10** (1960) 1181-1197. 4.3
- [R1] S. Resnick, *Limit laws for record values*, J. Stoch. Process and Their Appl. **1** (1973) 67-82. 1.1, 1.1
- [R2] S. Resnick, *Extreme Values, Regular Variation and Point Processes*. Springer, New York **1** (1987). 1.1, 1.2, 2.4, 4, 4.3, 5, 5
- [RR] S. Resnick and M. Rubinovitch, *The structure of extremal processes*, Advances Appl. Probab. **5** (1973) 287-307. 1.1, 1.1, 5, 3
- [Sou1] B. Saussol, *Recurrence rate in rapidly mixing dynamical systems*, Discrete and Continuous Dynam. Systems, **15** (2004) 259-267. 1
- [Sou2] B. Saussol, *Recurrence rate in rapidly mixing dynamical systems*, Discrete and Continuous Dynamical Systems, **15**(1) (2006) 259-267. 1

BEIJING INTERNATIONAL CENTER FOR MATHEMATICAL RESEARCH (BICMR), PEKING UNIVERSITY,, NO.5 YIHEYUAN ROAD,, HAIDIAN DISTRICT,, BEIJING, CHINA, 100081

Email address: boyuan.zhao@pku.edu.cn