J. Appl. Probab. **60,** 1031–1068 (2023)
*doi:*10.1017/jpr.2022.114

A CENTRAL LIMIT THEOREM FOR CONSERVATIVE FRAGMENTATION CHAINS

SYLVAIN RUBENTHALER,* *Université Côte d'Azur*

Abstract

We are interested in a fragmentation process. We observe fragments frozen when their sizes are less than ε ($\varepsilon > 0$). It is known (Bertoin and Martínez, 2005) that the empirical measure of these fragments converges in law, under some renormalization. Hoffmann and Krell (2011) showed a bound for the rate of convergence. Here, we show a central limit theorem, under some assumptions. This gives us an exact rate of convergence.

Keywords: Fragmentation; branching process; renewal theory; central-limit theorems; propagation of chaos; *U*-statistics

2020 Mathematics Subject Classification: Primary 60J80
Secondary 60F05

1. Introduction

1.1. Scientific and economic context

One of the main goals in the mining industry is to extract blocks of metallic ore and then separate the metal from the valueless material. To do so, rock is fragmented into smaller and smaller pieces. This is carried out in a series of steps, the first one being blasting, after which the material goes through a sequence of crushers. At each step, the particles are screened, and if they are smaller than the diameter of the mesh of a classifying grid, they go to the next crusher. The process stops when the material has a sufficiently small size (more precisely, small enough to enable physicochemical processing).

This fragmentation process is energetically costly (each crusher consumes a certain quantity of energy to crush the material it is fed). One of the problems that faces the mining industry is that of minimizing the energy used. The optimization parameters are the number of crushers and their technical specifications.

In [4], the authors proposed a mathematical model of what happens in a crusher. In this model, the rock pieces/fragments are fragmented independently of each other, in a random and auto-similar manner. This is consistent with what is observed in the industry, and is supported by [12, 19, 22, 25]. Each fragment has a size s (in $\mathbb{R}^+$) and is then fragmented into smaller fragments of sizes $s_1, s_2, \ldots$ such that the sequence $(s_1/s, s_2/s, \ldots)$ has a law ν which does not depend on s (which is why the fragmentation is said to be auto-similar). This law ν is called the *dislocation measure* (each crusher has its own dislocation measure). The dynamic of the fragmentation process is thus modeled in a stochastic way.

In each crusher, the rock pieces are fragmented repetitively until they are small enough to slide through a mesh whose holes have a fixed diameter. So the fragmentation process stops

Received 2 March 2020; revision received 11 October 2022.
* Postal address: Université Côte d'Azur, CNRS, LJAD, France. Email: rubentha@unice.fr

for each fragment when its size is smaller than the diameter of the mesh, which we denote by ε ($\varepsilon > 0$). We are interested in the *statistical distribution* of the fragments coming out of a crusher. If we renormalize the sizes of these fragments by dividing them by ε, we obtain a measure $\gamma_{-\log(\varepsilon)}$, which we call the *empirical measure* (the reason for the index $-\log(\varepsilon)$ instead of ε will be made clear later). In [4], the authors showed that the energy consumed by the crusher to reduce the rock pieces to fragments whose diameters are smaller than ε can be computed as an integral of a bounded function against the measure $\gamma_{-\log(\varepsilon)}$ [5, 6, 24]. For each crusher, the empirical measure $\gamma_{-\log(\varepsilon)}$ is one of the only two observable variables (the other one being the size of the pieces pushed into the grinder). The specifications of a crusher are summarized in ε and ν.

1.2. State of the art

In [4], the authors showed that the energy consumed by a crusher to reduce rock pieces of a fixed size into fragments whose diameter are smaller than ε behaves asymptotically like a power of ε when ε goes to zero. More precisely, this energy multiplied by a power of ε converges towards a constant of the form $\kappa = \nu(\varphi)$ (the integral of ν, the dislocation measure, against a bounded function φ). They also showed a law of large numbers for the empirical measure $\gamma_{-\log(\varepsilon)}$. More precisely, if f is bounded continuous, $\gamma_{-\log(\varepsilon)}(f)$ converges in law, when ε goes to zero, towards an integral of f against a measure related to ν (this result also appears in [16, p. 399]). We set $\gamma_\infty(f)$ to be this limit (check (5.2), (2.5), and (2.2) to get an exact formula). The empirical measure $\gamma_{-\log(\varepsilon)}$ thus contains information relative to ν and we could extract from it an estimation of κ or of an integral of any function against ν.

It is worth noting that by studying what happens in various crushers, we could study a family $(\nu_i(f_j))_{i\in I, j\in J}$ (with an index i for the number of the crusher and the index j for the jth test function in a well-chosen basis). Using statistical learning methods, we could from there make a prediction for $\nu(f_j)$ for a new crusher for which we know only the mechanical specifications (shape, power, frequencies of the rotating parts, . . .). It would evidently be interesting to know ν before even building the crusher.

In the same spirit, [14] studied the energy efficiency of two crushers used one after the other. When the final size of the fragments tends to zero, [14] tells us whether it is more efficient energywise to use one crusher or two crushers in a row (another asymptotic is also considered there).

In [15], the authors proved a convergence result for the empirical measure similar to the one in [4], the convergence in law being replaced by an almost sure convergence. In [16], the authors gave a bound on the rate of this convergence, in an L^2 sense, under the assumption that the fragmentation is conservative. This assumption means there is no loss of mass due to the formation of dust during the fragmentation process.

The state of the art as described is shown in Fig. 1. We have convergence results [4, 15] of an empirical quantity towards constants of interest (a different constant for each test function f). Using some transformations, these constants could be used to estimate the constant κ. Thus, it is natural to ask what the exact rate of convergence in this estimation is, if only to be able to build confidence intervals. In [16], we only have a bound on the rate.

When a sequence of empirical measures converges to some measure, it is natural to study the fluctuations, which often turn out to be Gaussian. For such results in the case of empirical measures related to the mollified Boltzmann equation, see [7, 18, 23]. When interested in the limit of an n-tuple as in (1.1), we say we are looking at the convergence of a U-statistic. Textbooks deal with the case where the points defining the empirical measure are independent

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$\begin{array}{ccc} \gamma_{-\log(\varepsilon)} & \underset{\varepsilon\to 0}{\overset{\text{(bound on rate)}}{\longrightarrow}} & \gamma_\infty \\ & & \updownarrow \text{relation} \\ \text{energy} \times (\text{power of } \varepsilon) & \underset{\varepsilon\to 0}{\sim} & \kappa = \nu(\varphi) \end{array}$$

FIGURE 1. State of the art.

or have a known correlation (see [8, 13, 17]). The problem is more complex when the points defining the empirical measure interact with each other as is the case here.

1.3. Goal of the paper

As explained above, we want to obtain the rate of convergence in $\gamma_{-\log(\varepsilon)}$ when ε goes to zero. We want to produce a central limit theorem of the form: for a bounded continuous f, $\varepsilon^\beta(\gamma_{-\log(\varepsilon)}(f) - \gamma_\infty(f))$ converges towards a non-trivial measure when ε goes to zero (the limiting measure will in fact be Gaussian), for some exponent β. The techniques used will allow us to prove the convergence towards a multivariate Gaussian of a vector of the form

$$\varepsilon^\beta(\gamma_{-\log(\varepsilon)}(f_1) - \gamma_\infty(f_1), \ldots, \gamma_{-\log(\varepsilon)}(f_n) - \gamma_\infty(f_n)) \tag{1.1}$$

for functions $f_1, \ldots, f_n$.

More precisely, if by $Z_1, Z_2, \ldots, Z_N$ we denote the fragments sizes that go out from a crusher (with mesh diameter equal to ε), we would like to show that, for a bounded continuous f,

$$\gamma_{-\log(\varepsilon)}(f) := \sum_{i=1}^{N} Z_i f(Z_i/\varepsilon) \longrightarrow \gamma_\infty(f)$$

almost surely (a.s.) when $\varepsilon \to 0$, and that, for all n, and $f_1, \ldots, f_n$ bounded continuous function such that $\gamma_\infty(f_i) = 0$, $\varepsilon^\beta(\gamma_{-\log(\varepsilon)}(f_1), \ldots, \gamma_{-\log(\varepsilon)}(f_n))$ converges in law towards a multivariate Gaussian when ε goes to zero.

The exact results are stated in Proposition 5.1 and Theorem 5.1.

1.4. Outline of the paper

We will state our assumptions along the way (Assumptions 2.1, 2.2, 2.3, and 3.1). Assumption 3.1 can be found at the beginning of Section 3. We define our model in Section 2. The main idea is that we want to follow tags during the fragmentation process. Let us imagine the fragmentation is the process of breaking a stick (modeled by [0, 1]) into smaller sticks. We suppose that the original stick has painted dots, and that during the fragmentation process we take note of the sizes of the sticks supporting the painted dots. When the sizes of these sticks get smaller than ε ($\varepsilon > 0$), the fragmentation is stopped for them and we call them the painted sticks. In Section 3, we make use of classical results on renewal processes and of [21] to show that the size of one painted stick has an asymptotic behavior when ε goes to zero and that we have a bound on the rate with which it reaches this behavior. Section 4 is the most technical. There we study the asymptotics of symmetric functionals of the sizes of the painted sticks (always when ε goes to zero). In Section 5, we precisely define the measure we are interested in (γ_T with $T = -\log(\varepsilon)$). Using the results of Section 4, it is then easy to show a law of large

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

numbers for γ_T (Proposition 5.1) and a central limit theorem (Theorem 5.1). Proposition 5.1 and Theorem 5.1 are our two main results. The proof of Theorem 5.1 is based on a simple computation involving characteristic functions (the same technique was previously used in [9–11, 20]).

1.5. Notation

For x in $\mathbb{R}$, we set $\lceil x \rceil = \inf\{n \in \mathbb{Z} : n \geq x\}$, $\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n \leq x\}$. The symbol $\sqcup$ means 'disjoint union'. For n in $\mathbb{N}^*$, we set $[n] = \{1, 2, \ldots, n\}$. For f an application from a set E to a set F, we write $f : E \hookrightarrow F$ if f is injective and, for k in $\mathbb{N}^*$, if $F = E$, we set

$$f^{\circ k} = \underbrace{f \circ f \circ \cdots \circ f}_{k \text{ times}}.$$

For any set E, we set $\mathcal{P}(E)$ to be the set of subsets of E.

2. Statistical model

2.1. Fragmentation chains

Let $\varepsilon > 0$. As in [16], we start with the space

$$\mathcal{S}^{\downarrow} = \left\{ \mathbf{s} = (s_1, s_2, \ldots), \ s_1 \geq s_2 \geq \cdots \geq 0, \ \sum_{i=1}^{+\infty} s_i \leq 1 \right\}.$$

A fragmentation chain is a process in $\mathcal{S}^{\downarrow}$ characterized by

- a dislocation measure ν, which is a finite measure on $\mathcal{S}^{\downarrow}$;
- a description of the law of the times between fragmentations.

A fragmentation chain with dislocation measure ν is a Markov process $X = (X(t), t \geq 0)$ with values in $\mathcal{S}^{\downarrow}$. Its evolution can be described as follows: a fragment with size x lives for some time (which may or may not be random) then splits and gives rise to a family of smaller fragments distributed as $x\xi$, where ξ is distributed according to $\nu(\cdot)/\nu(\mathcal{S}^{\downarrow})$. We suppose the lifetime of a fragment of size x is an exponential time of parameter $x^{\alpha}\nu(\mathcal{S}^{\downarrow})$, for some α. We could make different assumptions here on the lifetime of fragments, but this would not change our results. Indeed, as we are interested in the sizes of the fragments frozen as soon as they are smaller than ε, the time they need to become this small is not important.

We denote by $\mathbb{P}_m$ the law of X started from the initial configuration $(m, 0, 0, \ldots)$ with m in $(0, 1]$. The law of X is entirely determined by α and $\nu(\cdot)$ [2, Theorem 3].

We make the same assumption as in [16].

Assumption 2.1. *$\nu(\mathcal{S}^{\downarrow}) = 1$ and $\nu(s_1 \in]0; 1[) = 1$.*

Let $\mathcal{U} := \{0\} \cup \bigcup_{n=1}^{+\infty} (\mathbb{N}^*)^n$ denote the infinite genealogical tree. For u in $\mathcal{U}$, we use the conventional notation $u = ()$ if $u = \{0\}$ and $u = (u_1, \ldots, u_n)$ if $u \in (\mathbb{N}^*)^n$ with $n \in \mathbb{N}^*$. This way, any u in $\mathcal{U}$ can be denoted by $u = (u_1, \ldots, u_n)$ for some $u_1, \ldots, u_n$ and with n in $\mathbb{N}$. Now, for $u = (u_1, \ldots, u_n) \in \mathcal{U}$ and $i \in \mathbb{N}^*$, we say that u is in the nth generation and we write $|u| = n$; we write $ui = (u_1, \ldots, u_n, i)$, $u(k) = (u_1, \ldots, u_k)$ for all $k \in [n]$. For any $u = (u_1, \ldots, u_n)$ and $v = ui$ ($i \in \mathbb{N}^*$), we say that u is the mother of v. For any u in $\mathcal{U}\backslash\{0\}$ ($\mathcal{U}$ deprived of its root), u has exactly one mother and we denote it by $\boldsymbol{m}(u)$. The set $\mathcal{U}$ is ordered alphanumerically:

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

- If u and v are in $\mathcal{U}$ and $|u| < |v|$ then $u < v$.
- If u and v are in $\mathcal{U}$ and $|u| = |v| = n$, $u = (u_1, \ldots, u_n)$, and $v = (v_1, \ldots, v_n)$ with $u_1 = v_1, \ldots, u_k = v_k$, $u_{k+1} < v_{k+1}$ then $u < v$.

Suppose we have a process X which has the law $\mathbb{P}_m$. For all ω, we can index the fragments that are formed by the process X with elements of $\mathcal{U}$ in a recursive way.

- We start with a fragment of size m indexed by $u = ()$.
- If a fragment x, with a birth time t_1 and a split time t_2, is indexed by u in $\mathcal{U}$, at time t_2 this fragment splits into smaller fragments of sizes $(xs_1, xs_2, \ldots)$ with $(s_1, s_2, \ldots)$ of law $\nu(\cdot)/\nu(\mathcal{S}^{\downarrow})$. We index the fragment of size xs_1 by $u1$, the fragment of size xs_2 by $u2$, and so on.

A mark is an application from $\mathcal{U}$ to some other set. We associate a mark $\xi_{\ldots}$ on the tree $\mathcal{U}$ to each path of the process X. The mark at node u is ξ_u, where ξ_u is the size of the fragment indexed by u. The distribution of this random mark can be described recursively as follows.

Proposition 2.1. (Consequence of Proposition 1.3 (p. 25) of [3].) *There exists a family of independent and identically distributed (i.i.d.) variables indexed by the nodes of the genealogical tree* $\left((\widetilde{\xi}_{ui})_{i\in\mathbb{N}^*}, u \in \mathcal{U}\right)$, *where each* $(\widetilde{\xi}_{ui})_{i\in\mathbb{N}^*}$ *is distributed according to the law* $\nu(\cdot)/\nu(\mathcal{S}^{\downarrow})$, *and such that, given the marks* $(\xi_v, |v| \leq n)$ *of the first n generations, the marks at generation* $n+1$ *are given by* $\xi_{ui} = \widetilde{\xi}_{ui}\xi_u$, *where* $u = (u_1, \ldots, u_n)$ *and* $ui = (u_1, \ldots, u_n, i)$ *is the ith child of u.*

2.2. Tagged fragments

From now on, we suppose that we start with a block of size $m = 1$. We assume that the total mass of the fragments remains constant through time, as follows.

Assumption 2.2. (Conservation property.) $\nu\left(\sum_{i=1}^{+\infty} s_i = 1\right) = 1$.

This assumption was already present in [16]. We observe that the Malthusian exponent of [3, p. 27] is equal to 1 under our assumptions. Without this assumption, the link between the empirical measure $\gamma_{-\log(\varepsilon)}$ and the tagged fragments, (5.1), vanishes and our proofs of Proposition 5.1 and Theorem 5.1 fail.

We can now define tagged fragments. We use the representation of fragmentation chains as random infinite marked trees to define a fragmentation chain with q tags. Suppose we have a fragmentation process X of law $\mathbb{P}_1$. We take $(Y_1, Y_2, \ldots, Y_q)$ to be q i.i.d. variables of law $\mathcal{U}([0, 1])$. We set, for all u in $\mathcal{U}$, (ξ_u, A_u, I_u) with ξ_u defined as above. The random variables A_u take values in the subsets of [q]. The random variables I_u are intervals. These variables are defined as follows.

- We set $A_{\{0\}} = [q]$, $I_{\{0\}} = (0, 1]$ ($I_{\{0\}}$ is of length $\xi_{\{0\}} = 1$).
- For all $n \in \mathbb{N}$, suppose we are given the marks of the first n generations. Suppose that, for u in the nth generation, $I_u = (a_u, a_u + \xi_u]$ for some $a_u \in \mathbb{R}$ (it is of length ξ_u). Then the marks at generation $n+1$ are given by Proposition 2.1 (concerning $\xi_{\cdot}$) and, for all u such that $|u| = n$ and for all i in $\mathbb{N}^*$,

$$I_{ui} = \left(a_u + \xi_u\left(\widetilde{\xi}_{u1} + \cdots + \widetilde{\xi}_{u(i-1)}\right), a_u + \xi_u\left(\widetilde{\xi}_{u1} + \cdots + \widetilde{\xi}_{ui}\right)\right],$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$k \in A_{ui}$ if and only if $Y_k \in I_{ui}$ (I_{ui} is then of length ξ_{ui}). We observe that for all $j \in [q]$, $u \in \mathcal{U}$, $i \in \mathbb{N}^*$,

$$\mathbb{P}\big(j \in A_{ui} \,|\, j \in A_u, \widetilde{\xi}_{ui}\big) = \widetilde{\xi}_{ui}. \tag{2.1}$$

In this definition, we imagine having q dots on the interval [0, 1], and we impose that dot j has the position Y_j (for all j in $[q]$). During the fragmentation process, if we know that dot j is in the interval I_u of length ξ_u, then the probability that this dot is on I_{ui} (for some i in $\mathbb{N}^*$, I_{ui} of length ξ_{ui}) is equal to $\xi_{ui}/\xi_u = \widetilde{\xi}_{ui}$.

In the case $q = 1$, the branch $\{u \in \mathcal{U} : A_u \neq \emptyset\}$ has the same law as the randomly tagged branch of [3, Section 1.2.3]. The presentation is simpler in our case because the Malthusian exponent is 1 under Assumption 2.2.

2.3. Observation scheme

We freeze the process when the fragments become smaller than a given threshold $\varepsilon > 0$. That is, we have the data $(\xi_u)_{u \in \mathcal{U}_\varepsilon}$, where $\mathcal{U}_\varepsilon = \{u \in \mathcal{U},\ \xi_{\boldsymbol{m}(u)} \geq \varepsilon,\ \xi_u < \varepsilon\}$.

We now look at q tagged fragments ($q \in \mathbb{N}^*$). For each i in $[q]$, we call $L_0^{(i)} = 1, L_1^{(i)}, L_2^{(i)}, \ldots$ the successive sizes of the fragment having the tag i. More precisely, for each $n \in \mathbb{N}^*$, there is almost surely exactly one $u \in \mathcal{U}$ such that $|u| = n$ and $i \in A_u$; and so, $L_n^{(i)} = \xi_u$. For each i, the process $S_0^{(i)} = -\log\big(L_0^{(i)}\big) = 0 \leq S_1^{(i)} = -\log\big(L_1^{(i)}\big) \leq \cdots$ is a renewal process without delay, with waiting time following a law π (see [1, Chapter V] for an introduction to renewal processes). The waiting times (for i in $[q]$) are $S_0^{(i)}, S_1^{(i)} - S_0^{(i)}, S_2^{(i)} - S_1^{(i)}, \ldots$ The renewal times (for i in $[q]$) are $S_0^{(i)}, S_1^{(i)}, S_2^{(i)}, \ldots$ The law π is defined by the following:

For all bounded measurable f: $[0, 1] \to [0, +\infty)$,

$$\int_{\mathcal{S}^{\downarrow}} \sum_{i=1}^{+\infty} s_i f(s_i) \nu(\mathrm{d}\boldsymbol{s}) = \int_0^{+\infty} f(\mathrm{e}^{-x}) \pi(\mathrm{d}x) \tag{2.2}$$

(see [3, Proposition 1.6, p. 34], or [16, (3) and (4), p. 398]). Under Assumptions 2.1 and 2.2, [3, Proposition 1.6] is true, even without the Malthusian hypothesis of [3].

We make the following assumption on π.

Assumption 2.3. *There exist a and b,* $0 < a < b < +\infty$, *such that the support of* π *is* $[a,b]$. *We set* $\delta = \mathrm{e}^{-b}$.

We have added a comment about Assumption 2.3 in Remark 4.1. We believe that we could replace it by the following.

Assumption 2.4. *The support of* π *is* $(0, +\infty)$.

However, this would lead to difficult computations.

We set

$$T = -\log(\varepsilon). \tag{2.3}$$

We set, for all $i \in [q]$, $t \geq 0$,

$$\begin{aligned} N_t^{(i)} &= \inf\big\{j : S_j^{(i)} > t\big\}, \\ B_t^{(i)} &= S_{N_t^{(i)}}^{(i)} - t, \\ C_t^{(i)} &= t - S_{N_t^{(i)}-1}^{(i)} \end{aligned} \tag{2.4}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

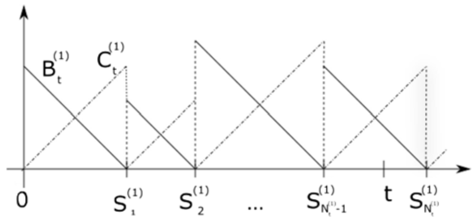

FIGURE 2. Process $B^{(1)}$ and $C^{(1)}$.

(see Fig. 2 for an illustration). The processes $B^{(i)}$, $C^{(i)}$, and $N^{(i)}$ are time-homogeneous Markov processes [1, Proposition 1.5, p. 141]. All of them are càdlàg (i.e. right-continuous with a left-hand-side limit). We call $B^{(i)}$ the residual lifetime of the fragment tagged by i. We call $C^{(i)}$ the age of the fragment tagged by i. We call $N^{(i)}$ the number of renewals up to time t. In the following, we treat t as a time parameter. This has nothing to do with the time in which the fragmentation process X evolves.

We observe that, for all t, $\left(B_t^{(1)}, \ldots, B_t^{(q)}\right)$ is exchangeable, meaning that for all σ in the symmetric group of order q, $\left(B_t^{(\sigma(1))}, \ldots, B_t^{(\sigma(q))}\right)$ has the same law as $\left(B_t^{(1)}, \ldots, B_t^{(q)}\right)$. When we look at the fragments of sizes $(\xi_u,\ u \in \mathcal{U}_\varepsilon : A_u \neq \emptyset)$, we have almost the same information as when we look at $\left(B_T^{(1)}, B_T^{(2)}, \ldots, B_T^{(q)}\right)$. We say *almost* because knowing $\left(B_T^{(1)}, B_T^{(2)}, \ldots, B_T^{(q)}\right)$ does not give exactly the number of u in $\mathcal{U}_\varepsilon$ such that A_u is not empty.

In the remainder of Section 2 we define processes that will be useful when we describe the asymptotics of our model (in Section 4).

2.4. Stationary renewal processes $\left(\overline{B}^{(1)}, \overline{B}^{(1),\nu}\right)$

We define $\widetilde{X}$ to be an independent copy of X. We suppose it has q tagged fragments. Therefore it has a mark $\left(\widetilde{\xi}, \widetilde{A}\right)$ and renewal processes $\left(\widetilde{S}_k^{(i)}\right)_{k\geq 0}$ (for all i in $[q]$) defined in the same way as for X. We let $\left(\widetilde{B}^{(1)}, \widetilde{B}^{(2)}\right)$ be the residual lifetimes of the fragments tagged by 1 and 2.

Let $\mu = \int_0^{+\infty} x\pi(\mathrm{d}x)$, and let π_1 be the distribution with density $x \mapsto x/\mu$ with respect to π. We set $\overline{C}$ to be a random variable of law π_1; U to be independent of $\overline{C}$ and uniform on (0, 1); and $\widetilde{S}_{-1} = \overline{C}(1-U)$. The process $\overline{S}_0 = \widetilde{S}_{-1}, \overline{S}_1 = \widetilde{S}_{-1} + \widetilde{S}_0^{(1)}, \overline{S}_2 = \widetilde{S}_{-1} + \widetilde{S}_1^{(1)}, \overline{S}_2 = \widetilde{S}_{-1} + \widetilde{S}_2^{(1)}$, ... is a renewal process with delay π_1 (with waiting times $\overline{S}_0, \overline{S}_1 - \overline{S}_0, \ldots$ all smaller than b by Assumption 2.3). The renewal times are $\overline{S}_0, \overline{S}_1, \overline{S}_2, \ldots$ We set $(\overline{B}_t^{(1)})_{t\geq 0}$ to be the residual lifetime process of this renewal process,

$$\overline{B}_t^{(1)} = \begin{cases} \overline{C}(1-U) - t & \text{if } t < \overline{S}_0, \\ \inf_{n\geq 0}\left\{\overline{S}_n : \overline{S}_n > t\right\} - t & \text{if } t \geq \overline{S}_0; \end{cases}$$

we define $(\overline{C}_t^{(1)})_{t\geq 0}$ as

$$\overline{C}_t^{(1)} = \begin{cases} \overline{C}U + t & \text{if } t < \overline{S}_0, \\ t - \sup_{n\geq 0}\left\{\overline{S}_n : \overline{S}_n \leq t\right\} & \text{if } t \geq \overline{S}_0 \end{cases}$$

(we call it the age process of our renewal process); and we set $\overline{N}_t^{(1)} = \inf\left\{j : \overline{S}_j > t\right\}$.

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

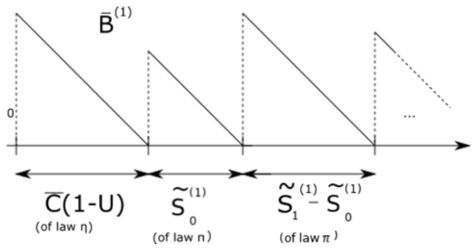

FIGURE 3. Renewal process with delay.

Fact 2.1. *Theorem* 3.3 *on p.* 151 *of* [1] *tells us that* $(\overline{B}_t^{(1)}, \overline{C}_t^{(1)})_{t\geq 0}$ *has the same transition as* $(B_t^{(1)}, C_t^{(1)})_{t\geq 0}$ *defined above, and that* $(\overline{B}_t^{(1)}, \overline{C}_t^{(1)})_{t\geq 0}$ *is stationary. In particular, this means that the law of* $\overline{B}_t^{(1)}$ *does not depend on t.*

Figure 3 provides a graphic representation of $\overline{B}_\cdot^{(1)}$. It might be counter-intuitive to start with $\overline{B}_0^{(1)}$ having a law which is not π in order to get a stationary process, but [1, Corollary 3.6, p. 153] is clear on this point: a delayed renewal process (with waiting time of law π) is stationary if and only if the distribution of the initial delay is η (defined below).

We define a measure η on $\mathbb{R}^+$ by its action on bounded measurable functions:

For all bounded measurable f:

$$\mathbb{R}^+ \to \mathbb{R}, \quad \eta(f) = \frac{1}{\mu}\int_{\mathbb{R}^+} \mathbb{E}(f(Y-s)\mathbf{1}_{\{Y-s>0\}})\,\mathrm{d}s \quad (Y \sim \pi). \tag{2.5}$$

Lemma 2.1. *The measure* η *is the law of* $\overline{B}_t^{(1)}$ *(for any* $t \geq 0$*). It is also the law of* $(\overline{C}_t^{(1)})$ *(for any t).*

Proof. We show the proof for $\overline{B}_t^{(1)}$ only. Let $\xi \geq 0$. We set $f(y) = \mathbf{1}_{y\geq\xi}$, for all y in $\mathbb{R}$. We have, with Y of law π,

$$\begin{aligned}\frac{1}{\mu}\int_{\mathbb{R}^+} \mathbb{E}(f(Y-s)\mathbf{1}_{Y-s>0})\,\mathrm{d}s &= \frac{1}{\mu}\int_{\mathbb{R}^+}\left(\int_0^y \mathbf{1}_{y-s\geq\xi}\,\mathrm{d}s\right)\pi(\mathrm{d}y)\\ &= \frac{1}{\mu}\int_{\mathbb{R}^+}(y-\xi)_+\pi(\mathrm{d}y)\\ &= \int_\xi^{+\infty}\left(1-\frac{\xi}{y}\right)\frac{y}{\mu}\pi(\mathrm{d}y) = \mathbb{P}(\overline{C}(1-U)\geq\xi).\end{aligned}$$ □

We set η_2 to be the law of $(\overline{C}_0^{(1)}, \overline{B}_0^{(1)}) = (\overline{C}U, \overline{C}(1-U))$. The support of η_2 is $\mathcal{C} := \{(u, v) \in [0, b]^2 : a \leq u+v \leq b\}$.

For v in $\mathbb{R}$, we now want to define a process

$$(\overline{C}_t^{(1),v}, \overline{B}_t^{(1),v})_{t\geq v-2b} \text{ having the same transition as}$$

$$(C_t^{(1)}, B_t^{(1)}) \text{ and being stationary.} \tag{2.6}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

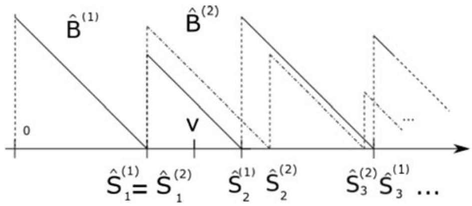

FIGURE 4. Processes $\widehat{B}^{(1),v}$, $\widehat{B}^{(2),v}$.

We set $(\overline{C}^{(1),v}_{v-2b}, \overline{B}^{(1),v}_{v-2b})$ such that it has the law η_2. As we have given its transition, the process $(\overline{C}^{(1),v}_t, \overline{B}^{(1),v}_t)_{t\geq v-2b}$ is well defined in law. In addition, we suppose that it is independent of all the other processes. By Fact 2.1, the process $(\overline{C}^{(1),v}_t, \overline{B}^{(1),v}_t)_{t\geq v-2b}$ is stationary.

We define the renewal times of $\overline{B}^{(1),v}$ by $\overline{S}^{(1),v}_1 = \inf\{t \geq v-2b : \overline{B}^{(1),v}_{t+} \neq \overline{B}^{(1),v}_{t-}\}$, and, by recurrence, $\overline{S}^{(1),v}_k = \inf\{t > \overline{S}^{(1),v}_{k-1} : \overline{B}^{(1),v}_{t+} \neq \overline{B}^{(1),v}_{t-}\}$. We also define, for all t, $\overline{N}^{(1),v}_t = \inf\{j : \overline{S}^{(1),v}_j > t\}$. As will be seen later, the processes $\overline{B}^{(1),v}$ and $\overline{B}^{(2),v}$ are used to define asymptotic quantities (see, for example, Proposition 4.1) and we need them to be defined on an interval $[v, +\infty)$ with v possibly in $\mathbb{R}^-$. The process $\overline{B}^{(2),v}$ is defined below (Section 2.6).

2.5. Tagged fragments conditioned to split up $(\widehat{B}^{(1),v}, \widehat{B}^{(2),v})$

For v in $[0, +\infty)$, we define a process $(\widehat{B}^{(1),v}_t, \widehat{B}^{(2),v}_t)_{t\geq 0}$ such that

$\widehat{B}^{(1),v} = B^{(1)}$ and, with $B^{(1)}$ fixed, $\widehat{B}^{(2),v}$ has the law of $B^{(2)}$ conditioned on

$$\text{for all } u \in \mathcal{U},\ \ 1 \in A_u \Rightarrow [2 \in A_u \Leftrightarrow -\log(\xi_u) \leq v], \tag{2.7}$$

which reads as follows: the tag 2 remains on the fragment bearing the tag 1 until the size of the fragment is smaller than e^{-v}. We observe that, conditionally on $\widehat{B}^{(1),v}_v$ and $\widehat{B}^{(2),v}_v$, $(\widehat{B}^{(1),v}_{v+\widehat{B}^{(1),v}_v+t})_{t\geq 0}$ and $(\widehat{B}^{(2),v}_{v+\widehat{B}^{(2),v}_v+t})_{t\geq 0}$ are independent. We also define $\widehat{C}^{(1),v} = C^{(1)}$. There is an algorithmic way to define $\widehat{B}^{(1),v}$ and $\widehat{B}^{(2),v}$, which is illustrated in Fig. 4. Remember that $\widehat{B}^{(1),v} = B^{(1)}$, and the definition of the mark $(\xi_u, A_u, I_u)_{u\in\mathcal{U}}$ in Section 2.2. We call $(\widehat{S}^{(i)}_j)_{i=1,2;j\geq 1}$ the renewal times of these processes (as before, they can be defined as the times when the right-hand-side and left-hand-side limits are not the same). If $\widehat{S}^{(1)}_j \leq v$ then $\widehat{S}^{(2)}_j = \widehat{S}^{(1)}_j$. If k is such that $\widehat{S}^{(1)}_{k-1} \leq v$ and $\widehat{S}^{(1)}_k > v$, we remember that

$$\exp\left(\widehat{S}^{(1)}_k - \widehat{S}^{(1)}_{k-1}\right) = \widetilde{\xi}_{ui} \tag{2.8}$$

for some u in $\mathcal{U}$ with $|u| = k-1$ and some i in $\mathbb{N}^*$ (because $\widehat{B}^{(1),v} = B^{(1)}$). We have points $Y_1, Y_2 \in [0, 1]$ such that Y_1 and Y_2 are in I_u of length ξ_u. Conditionally on $\{Y_1, Y_2 \in I_u\}$, Y_1 and Y_2 are independent and uniformly distributed on I_u. The interval I_{ui}, of length $\xi_u\widetilde{\xi}_{ui}$, is a sub-interval of I_u such that $Y_1 \in I_{ui}$, because of (2.8). Then, for $r \in \mathbb{N}^*\backslash\{i\}$, we want Y_2 to be in I_{ur} with probability $\widetilde{\xi}_{ur}/(1-\widetilde{\xi}_{ui})$ (because we want $2 \notin A_{ui}$). So we take $\widehat{S}^{(2)}_k = \widehat{S}^{(1)}_{k-1} - \log\widetilde{\xi}_{ur}$ with probability $\widetilde{\xi}_{ur}/(1-\widetilde{\xi}_{ui})$ $(r \in \mathbb{N}^*\backslash\{i\})$.

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

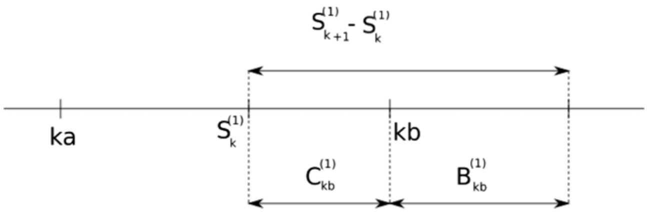

FIGURE 5. $B_{kb}^{(1)}$ and $C_{kb}^{(1)}$.

Fact 2.2.

(i) *The knowledge of the couple* $\left(\widehat{S}^{(1)}_{N_v^{(1)}-1}, \widehat{B}_v^{(1),v}\right)$ *is equivalent to the knowledge of the couple* $\left(\widehat{C}_v^{(1),v}, \widehat{B}_v^{(1),v}\right)$.

(ii) *The law of* $B_v^{(1)}$ *knowing* $C_v^{(1)}$ *is* $\pi - C_v^{(1)}$, *with* π *conditioned to be bigger than* $C_v^{(1)}$; *we call it* $\eta_1\left(\cdots \mid C_v^{(1)}\right)$. *As* $\widehat{B}^{(1),v} = B^{(1)}$ *and* $\widehat{C}^{(1),v} = C^{(1)}$, *we also have that the law of* $\widehat{B}_v^{(1),v}$ *knowing* $\widehat{C}_v^{(1),v}$ *is* $\eta_1\left(\cdots \mid \widehat{C}_v^{(1),v}\right)$.

(iii) *The law of* $\widehat{B}_v^{(2),v}$ *knowing* $\left(\widehat{C}_v^{(1),v}, \widehat{B}_v^{(1),v}\right)$ *does not depend on* v *and we denote it by* $\eta'\left(\cdots \mid \widehat{C}_v^{(1),v}, \widehat{B}_v^{(1),v}\right)$.

The subsequent waiting times $\widehat{S}_{k+1}^{(1)} - \widehat{S}_k^{(1)}, \widehat{S}_{k+1}^{(2)} - \widehat{S}_k^{(2)}, \ldots$ are chosen independently of each other, each of them having the law π. For j equal to 1 or 2 and t in $[0, +\infty)$, we define $\widehat{N}_t^{(j)} = \inf\{i : \widehat{S}_i^{(j)} > t\}$. We observe that, for $t \geq 2b$, $\widehat{N}_t^{(1)}$ is bigger than 2 (because of Assumption 2.3).

2.6. Two stationary processes after a split-up $(\overline{B}^{(1),v}, \overline{B}^{(2),v})$

Let k be an integer, $k \geq 2$, such that

$$k \times (b - a) \geq b. \tag{2.9}$$

Now we state a small lemma that will be useful in what follows. Remember that the process $\left(\overline{C}_t^{(1),v}, \overline{B}_t^{(1),v}\right)_{t \geq v-2b}$ is defined in (2.6). The process $\left(\widehat{C}_t^{(1),v}, \widehat{B}_t^{(1),v}\right)_{t \geq 0}$ is defined in the previous section.

Lemma 2.2. *Let* v *be in* $\mathbb{R}$. *The variables* $\left(\overline{C}_v^{(1),v}, \overline{B}_v^{(1),v}\right)$ *and* $\left(\widehat{C}_{kb}^{(1),kb}, \widehat{B}_{kb}^{(1),kb}\right)$ *have the same support (and it is* $\mathcal{C}$, *defined below Lemma 2.1).*

Proof. The law η_2 is the law of $\left(\overline{C}_0^{(1)}, \overline{B}_0^{(1)}\right)$ (η_2 is defined below Lemma 2.1). As previously stated, the support of η_2 is $\mathcal{C}$; so, by stationarity, the support of $\left(\overline{C}_v^{(1),v}, \overline{B}_v^{(1),v}\right)$ is $\mathcal{C}$.

Keep in mind that $\widehat{B}^{(1),v} = B^{(1)}$, $\widehat{C}^{(1),v} = C^{(1)}$. By Assumption 2.3, the support of $S_k^{(1)}$ is $[ka, kb]$ and the support of $S_{k+1}^{(1)} - S_k^{(1)}$ is $[a, b]$. If $S_{k+1}^{(1)} > kb$ then $B_{kb}^{(1)} = S_{k+1}^{(1)} - S_k^{(1)} - \left(kb - S_k^{(1)}\right)$ and $C_{kb}^{(1)} = kb - S_k^{(1)}$ (see Fig. 5).

The support of $S_k^{(1)}$ is $[ka, kb]$ and $kb - ka \geq b$ (see (2.9)), so, as $S_k^{(1)}$ and $S_{k+1}^{(1)} - S_k^{(1)}$ are independent, we get that the support of $\left(C_{kb}^{(1)}, S_{k+1}^{(1)} - S_k^{(1)}\right)$ includes $\{(u, w) \in [0; b]^2 : w \geq$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$\sup(a, u)\}$. Hence, the support of $\big(C^{(1)}_{kb}, B^{(1)}_{kb}\big) = \big(C^{(1)}_{kb}, S^{(1)}_{k+1} - S^{(1)}_{k} - C^{(1)}_{kb}\big)$ includes $\mathcal{C}$. As this support is included in $\mathcal{C}$, we have proved the desired result. □

For v in $\mathbb{R}$, we define a process $\big(\overline{B}^{(2),v}_t\big)_{t\geq v}$. We start with:

$$\overline{B}^{(2),v}_v \text{ has the law } \eta'\big(\cdots \mid \overline{C}^{(1),v}_v, \overline{B}^{(1),v}_v\big) \tag{2.10}$$

(remember that η' is defined in Fact 3.1). This conditioning is correct because the law of $\big(\overline{C}^{(1),v}_v, \overline{B}^{(1),v}_v\big)$ is the law η_2, whose support is included in the support of the law of $\big(\widehat{C}^{(1),kb}_{kb}, \widehat{B}^{(1),kb}_{kb}\big)$, which is η_2 (see the lemma above, and below (2.6)). We then let the process $\big(\overline{B}^{(1),v}_t, \overline{B}^{(2),v}_t\big)_{t\geq v}$ run its course as a Markov process having the same transition as $\big(\widehat{B}^{(1),kb}_{t-v+kb}, \widehat{B}^{(2),kb}_{t-v+kb}\big)_{t\geq v}$. This means that, after time v, $\overline{B}^{(1),v}_t$ and $\overline{B}^{(2),v}_t$ decrease linearly (with slope -1) until they reach 0. When they reach 0, each of these two processes makes a jump of law π, independently of the other one. After that, they decrease linearly, and so on.

Fact 2.3. *The process* $\big(\overline{B}^{(1),v}_t, \overline{B}^{(2),v}_t\big)_{t\geq v}$ *is supposed independent from all the other processes (until now, we have defined its law and said that that* $\overline{B}^{(1),v}$ *is independent from all the other processes).*

3. Rate of convergence in the key renewal theorem

We need the following regularity assumption.

Assumption 3.1. *The probability* $\pi(\mathrm{d}x)$ *is absolutely continuous with respect to the Lebesgue measure (we will write* $\pi(\mathrm{d}x) = \pi(x)\,\mathrm{d}x$*). The density function* $x \mapsto \pi(x)$ *is continuous on* $(0; +\infty)$.

Fact 3.1. *Let* $\theta > 1$ *(*θ *is fixed in the rest of the paper). The density* π *satisfies* $\limsup_{x\to+\infty} \exp(\theta x)\pi(x) < +\infty$.

For φ a non-negative Borel-measurable function on $\mathbb{R}$, we set $S(\varphi)$ to be the set of complex-valued measures ρ (on the Borelian sets) such that $\int_{\mathbb{R}} \varphi(x)|\rho|(\mathrm{d}x) < \infty$, where $|\rho|$ stands for the total variation norm. If ρ is a finite complex-valued measure on the Borelian sets of $\mathbb{R}$, we define $\mathcal{T}\rho$ to be the σ-finite measure with the density

$$v(x) = \begin{cases} \rho((x, +\infty)) & \text{if } x \geq 0, \\ -\rho((-\infty, x]) & \text{if } x < 0. \end{cases}$$

Let F be the cumulative distribution function of π.

We set $B_t = B^{(1)}_t$ (see (2.4) for the definition of $B^{(1)}, B^{(2)}, \ldots$). By [1, Theorem 3.3, p. 151, and Theorem 4.3, p. 156], we know that B_t converges in law to a random variable B_∞ (of law η) and that C_t converges in law to a random variable C_∞ (of law η). The following theorem is a consequence of [21, Theorem 5.1, p. 2429]. It shows there is actually a rate of convergence for these convergences in law.

Theorem 3.1. *Let* $\varepsilon' \in (0, \theta)$*,* $M \in (0, +\infty)$*, and*

$$\varphi(x) = \begin{cases} \mathrm{e}^{(\theta-\varepsilon')x} & \text{if } x \geq 0, \\ 1 & \text{if} x < 0. \end{cases}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

If Y is a random variable of law π, then

$$\sup_{\alpha\,:\,\|\alpha\|_\infty\le M}\left|\mathbb{E}(\alpha(B_t))-\frac{1}{\mu}\int_{\mathbb{R}^+}\mathbb{E}(\alpha(Y-s)\mathbf{1}_{\{Y-s>0\}})\,\mathrm{d}s\right|=o\left(\frac{1}{\varphi(t)}\right) \tag{3.1}$$

as t approaches $+\infty$ outside a set of Lebesgue measure zero (the supremum is taken on α in the set of Borel-measurable functions on $\mathbb{R}$), and

$$\sup_{\alpha\,:\,\|\alpha\|_\infty\le M}\left|\mathbb{E}(\alpha(C_t))-\frac{1}{\mu}\int_{\mathbb{R}^+}\mathbb{E}(\alpha(Y-s)\mathbf{1}_{\{Y-s>0\}})\,\mathrm{d}s\right|=o\left(\frac{1}{\varphi(t)}\right) \tag{3.2}$$

as t approaches $+\infty$ outside a set of Lebesgue measure zero (the supremum is taken on α in the set of Borel-measurable functions on $\mathbb{R}$).

Proof. We give the proof of (3.1); the proof of (3.2) is very similar.

Let $*$ stand for the convolution product. We define the renewal measure $U(\mathrm{d}x)=\sum_{n=0}^{+\infty}\pi^{*n}(\mathrm{d}x)$ (where $\pi^{*0}(\mathrm{d}x)=\delta_0$, the Dirac mass at 0, and $\pi^{*n}=\pi*\pi*\cdots*\pi$, n times). We take i.i.d. variables $X, X_1, X_2, \dots$ of law π. Let $f:\mathbb{R}\to\mathbb{R}$ be a measurable function such that $\|f\|_\infty\le M$. We have, for all $t\ge 0$,

$$\begin{aligned}\mathbb{E}(f(B_t))&=\mathbb{E}\left(\sum_{n=0}^{+\infty}f(X_1+X_2+\cdots+X_{n+1}-t)\mathbf{1}_{\{X_1+\cdots+X_n\le t<X_1+\cdots+X_{n+1}\}}\right)\\&=\int_0^t\mathbb{E}(f(s+X-t)\mathbf{1}_{\{s+X-t>0\}})U(\mathrm{d}s).\end{aligned}$$

We set

$$g(t)=\begin{cases}\mathbb{E}(f(X-t)\mathbf{1}_{\{X-t>0\}}) & \text{if } t\ge 0,\\ 0 & \text{if } t<0.\end{cases}$$

We observe that $\|g\|_\infty\le M$. We have, for all $t\ge 0$,

$$|\mathbb{E}(f(X-t)\mathbf{1}_{\{X-t>0\}})|\le\|f\|_\infty\mathbb{P}(X>t)\le\|f\|_\infty\mathrm{e}^{-(\theta-\varepsilon'/2)t}\mathbb{E}\big(\mathrm{e}^{(\theta-\varepsilon'/2)X}\big).$$

We have, by Fact 3.1, $\mathbb{E}\big(\mathrm{e}^{(\theta-\varepsilon'/2)X}\big)<\infty$. The function φ is sub-multiplicative and is such that

$$\lim_{x\to-\infty}\frac{\log(\varphi(x))}{x}=0\le\lim_{x\to+\infty}\frac{\log(\varphi(x))}{x}=\theta-\varepsilon'.$$

The function g is in $L^1(\mathbb{R})$, and the function $g.\varphi$ is in $L^\infty(\mathbb{R})$. We have $g(x)\varphi(x)\to 0$ as $|x|\to\infty$,

$$\varphi(t)\int_t^{+\infty}|g(x)|\,\mathrm{d}x\underset{t\to+\infty}{\longrightarrow}0,\qquad \varphi(t)\int_{-\infty}^{t}|g(x)|\,\mathrm{d}x\underset{t\to-\infty}{\longrightarrow}0,$$

and $\mathcal{T}^{\circ 2}(\pi)\in S(\varphi)$.

Let us now take a function α such that $\|\alpha\|_\infty\le M$. We set

$$\widehat{\alpha}(t)=\begin{cases}\mathbb{E}\big(\alpha(X-t)\mathbf{1}_{\{X-t\ge 0\}}\big) & \text{if } t\ge 0,\\ 0 & \text{if } t<0.\end{cases}$$

Then we have $\|\widehat{\alpha}\|_\infty\le M$ and, computing as above for f, $\mathbb{E}(\alpha(B_t))=\widehat{\alpha}*U(t)$.

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

In the case where f is a constant equal to M, we have $\|g\|_\infty = M$. So, by [21, Theorem 5.1] (applied to the case $f \equiv M$), we have proved the desired result. □

Corollary 3.1. *There exists a constant Γ_1 bigger than 1 such that, for any bounded measurable function F on $\mathbb{R}$ such that $\eta(F) = 0$, for t outside a set of Lebesgue measure zero,*

$$|\mathbb{E}(F(B_t))| \leq \|F\|_\infty \times \frac{\Gamma_1}{\varphi(t)}, \tag{3.3}$$

$$|\mathbb{E}(F(C_t))| \leq \|F\|_\infty \times \frac{\Gamma_1}{\varphi(t)}. \tag{3.4}$$

for t outside a set of Lebesgue measure zero.

Proof. We provide the proof of (3.3) only; the proof of (3.4) is very similar.

We take $M = 1$ in Theorem 3.1. Keep in mind that η is defined in (2.5). By the above theorem, there exists a constant Γ_1 such that, for all measurable functions α such that $\|\alpha\|_\infty \leq 1$,

$$|\mathbb{E}(\alpha(B_t)) - \eta(\alpha)| \leq \frac{\Gamma_1}{\varphi(t)} \text{ (for } t \text{ outside a set of Lebesque measure zero).} \tag{3.5}$$

Let us now take a bounded measurable F such that $\eta(F) = 0$. By (3.5), we have, for t outside a set of Lebesgue measure zero,

$$\left|\mathbb{E}\left(\frac{F(B_t)}{\|F\|_\infty}\right) - \eta\left(\frac{F}{\|F\|_\infty}\right)\right| \leq \frac{\Gamma_1}{\varphi(t)}$$

$$|\mathbb{E}(F(B_t))| \leq \|F\|_\infty \times \frac{\Gamma_1}{\varphi(t)}.$$ □

4. Limits of symmetric functionals

4.1. Notation

We fix $q \in \mathbb{N}^*$, and set $\mathcal{S}_q$ to be the symmetric group of order q. A function $F : \mathbb{R}^q \to \mathbb{R}$ is symmetric if, for all $\sigma \in \mathcal{S}_q$ and all $(x_1, \ldots, x_q) \in \mathbb{R}^q$,

$$F(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(q)}) = F(x_1, x_2, \ldots, x_q).$$

For $F : \mathbb{R}^q \to \mathbb{R}$, we define a symmetric version of F by

$$F_{\text{sym}}(x_1, \ldots, x_q) = \frac{1}{q!} \sum_{\sigma \in \mathcal{S}_q} F(x_{\sigma(1)}, \ldots, x_{\sigma(q)}) \qquad \text{for all } (x_1, \ldots, x_q) \in \mathbb{R}^q. \tag{4.1}$$

We set $\mathcal{B}_{\text{sym}}(q)$ to be the set of bounded, measurable, symmetric functions F on $\mathbb{R}^q$, and we set $\mathcal{B}^0_{\text{sym}}(q)$ to be the F of $\mathcal{B}_{\text{sym}}(q)$ such that

$$\int_{x_1} F(x_1, x_2, \ldots, x_q)\eta(\mathrm{d}x_1) = 0 \qquad \text{for all } (x_2, \ldots, x_q) \in \mathbb{R}^{q-1}.$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

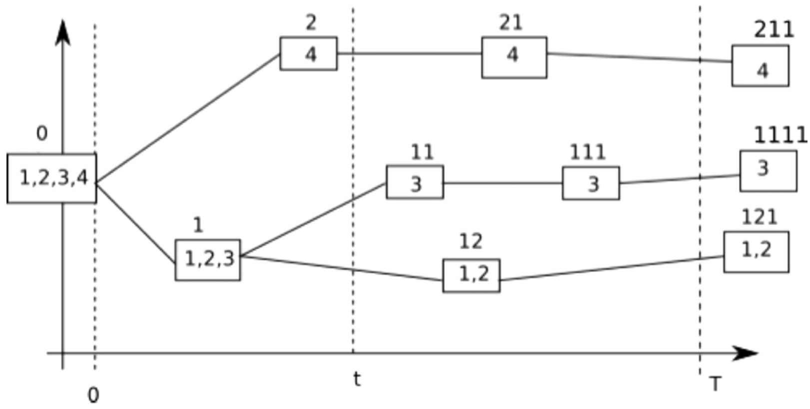

FIGURE 6. Example tree and marks.

Suppose that k is in $[q]$ and $l \geq 1$. For t in $[0, T]$, we consider the following collections of nodes of $\mathcal{U}$ (remember that $T = -\log \varepsilon$, and $\mathcal{U}$ and $\mathbf{m}(\cdot)$ are defined in Section 2.1):

$$\mathcal{T}_1 = \{u \in \mathcal{U} \backslash \{0\} : A_u \neq \emptyset,\ \xi_{\boldsymbol{m}(u)} \geq \varepsilon\} \cup \{0\},$$

$$S(t) = \{u \in \mathcal{T}_1 :\ -\log(\xi_{\boldsymbol{m}(u)}) \leq t,\ -\log(\xi_u) > t\} = \mathcal{U}_{e^{-t}}, \tag{4.2}$$

$$L_t = \sum_{u \in S(t) : A_u \neq \emptyset} (\#A_u - 1). \tag{4.3}$$

We set $\mathcal{L}_1$ to be the set of leaves in the tree $\mathcal{T}_1$. For t in $[0, T]$ and i in $[q]$, there exists one and only one u in $S(t)$ such that $i \in A_u$. We call it $u\{t, i\}$. Under Assumption 2.3, there exists a constant bounding the numbers of vertices of $\mathcal{T}_1$ almost surely.

Let us consider the example in Fig. 6. Here, we have a graphical representation of a realization of $\mathcal{T}_1$. Each node u of $\mathcal{T}_1$ is written above a rectangular box in which we read A_u; the right side of the box has the coordinate $-\log(\xi_u)$ on the x-axis. For simplicity, the node (1,1) is designated by 11, the node (1,2) by 12, and so on. In this example:

$$\mathcal{T}_1 = \{(0), (1), (2), (1, 1), (2, 1), (1, 2), (1, 1, 1), (2, 1, 1), (1, 1, 1, 1), (1, 2, 1)\},$$

$$\mathcal{L}_1 = \{(2, 1, 1), (1, 1, 1, 1), (1, 2, 1)\},$$

$$A_{(1)} = \{1, 2, 3\},\ A_{(1,2)} = \{1, 2\},\ \ldots,$$

$$S(t) = \{(1, 2), (1, 1), (2, 1)\},$$

$$u\{t, 1\} = (1, 2),\ u\{t, 2\} = (1, 2),\ u\{t, 3\} = (1, 1),\ u\{t, 4\} = (2, 1).$$

For $k, l \in \mathbb{N}$ and $t \in [0, T]$, we define the event

$$C_{k,l}(t) = \left\{ \sum_{u \in S(t)} \mathbf{1}_{\#A_u = 1} = k,\ \sum_{u \in S(t)} (\#A_u - 1) = l \right\}.$$

For example, in Fig. 6, we are in the event $C_{2,1}(t)$.

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

We define $\mathcal{T}_2 = \{u \in \mathcal{T}_1 \backslash \{0\} : \#A_{\boldsymbol{m}(u)} \geq 2\} \cup \{0\}$, $m_2 : u \in \mathcal{T}_2 \mapsto (\xi_u, \inf\{i, i \in A_u\})$. For example, in Fig. 6, $\mathcal{T}_2 = \{(0), (1), (2), (1, 1), (1, 2), (1, 2, 1)\}$. Let α be in $(0, 1)$.

Fact 4.1. *We can always suppose that* $(1-\alpha)T > b$ *because we are interested in* T *going to infinity. So, in the following, we suppose* $(1-\alpha)T > b$.

For any t, we can compute $\sum_{u\in S(t)} (\#A_u - 1)$ if we know $\sum_{u\in S(t)} \mathbf{1}_{\#A_u=1}$ and $\#S(t)$. As $T - \alpha T > b$, any u in $S(\alpha T)$ satisfies $\#A_u \geq 2$ if and only if u is the mother of some v in $\mathcal{T}_2$. So we deduce that $C_{k,l}(\alpha T)$ is measurable with respect to $(\mathcal{T}_2, m_2)$. We set, for all u in $\mathcal{T}_2$,

$$T_u = -\log(\xi_u). \tag{4.4}$$

For any i in $[q]$, $t \mapsto u\{t, i\}$ is piecewise constant and the ordered sequence of its jump times is $S_1^{(i)} < S_2^{(i)} < \cdots$ (the $S_{\dots}^{(i)}$ are defined in Section 2.3). We simply have that $1, \mathrm{e}^{-S_1^{(i)}}, \mathrm{e}^{-S_2^{(i)}}, \dots$ are the successive sizes of the fragment supporting the tag i. For example, in Fig. 6, we have

$$S_1^{(1)} = -\log(\xi_1), \quad S_2^{(1)} = -\log(\xi_{(1,2)}), \quad S_3^{(1)} = -\log(\xi_{(1,2,1)}), \dots \tag{4.5}$$

Let $\mathcal{L}_2$ be the set of leaves u in the tree $\mathcal{T}_2$ such that the set A_u has a single element n_u. For example, in Fig. 6, $\mathcal{L}_2 = \{(2), (1, 1)\}$. We observe that $\#\mathcal{L}_1 = q \Leftrightarrow \#\mathcal{L}_2 = q$, and thus

$$\{\#\mathcal{L}_1 = q\} \in \sigma(\mathcal{L}_2). \tag{4.6}$$

We summarize the definition of n_u:

$$\#A_u = 1 \Rightarrow A_u = \{n_u\}. \tag{4.7}$$

For q even ($q = 2p$) and for all t in $[0, T]$, we define the events

$$G_t = \{\text{for all } i \in [p], \text{ there exists } u_i \in \mathcal{U} : \xi_{u_i} < \mathrm{e}^{-t}, \ \xi_{\boldsymbol{m}(u_i)} \geq \mathrm{e}^{-t}, \ A_{u_i} = \{2i-1, 2i\}\},$$

$$\text{for all } i \in [p], \ G_{i,i+1}(t) = \{\text{there exists } u \in S(t) : \{2i-1, 2i\} \subset A_u\}.$$

We set, for all t in $[0, T]$, $\mathcal{F}_{S(t)} = \sigma(S(t), (\xi_u, A_u)_{u\in S(t)})$.

4.2. Intermediate results

The reader must keep in mind that $T = -\log(\varepsilon)$, (2.3), and that δ is defined in Assumption 2.3. The set $\mathcal{B}^0_{\mathrm{sym}}(q)$ is defined in Section 4.1.

Lemma 4.1. *We suppose that* F *is in* $\mathcal{B}^0_{\mathrm{sym}}(q)$ *and that* F *is of the form* $F = (f_1 \otimes f_2 \otimes \cdots \otimes f_q)_{\mathrm{sym}}$, *with* $f_1, f_2, \dots, f_q \in \mathcal{B}^0_{\mathrm{sym}}(1)$. *Let* A *be in* $\sigma(\mathcal{L}_2)$. *For any* α *in* $]0, 1[$, k *in* $[q]$, *and* l *in* $\{0, 1, \dots, (q-k-1)_+\}$, *we have*

$$\left|\mathbb{E}(\mathbf{1}_{C_{k,l}(\alpha T)} \mathbf{1}_A F(B_T^{(1)}, B_T^{(2)}, \dots, B_T^{(q)}))\right| \leq \|F\|_\infty \Gamma_1^q C_{\mathrm{tree}}(q) \left(\frac{1}{\delta}\right)^q \varepsilon^{q/2}$$

(for a constant $C_{\mathrm{tree}}(q)$ *defined in the proof, and* Γ_1 *defined in Corollary 3.1), and*

$$\varepsilon^{-q/2} \mathbb{E}\left(\mathbf{1}_{C_{k,l}(\alpha T)} \mathbf{1}_A F(B_T^{(1)}, B_T^{(2)}, \dots, B_T^{(q)})\right) \underset{\varepsilon \to 0}{\longrightarrow} 0.$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Proof. Let A be in $\sigma(\mathcal{L}_2)$. We have

$$\begin{aligned}&\mathbb{E}\big(\mathbf{1}_{C_{k,l}(\alpha T)}\mathbf{1}_A F\big(B_T^{(1)}, B_T^{(2)}, \dots, B_T^{(q)}\big)\big)\\&= \mathbb{E}\left(\mathbf{1}_A \sum_{f\,:\,\mathcal{T}_2\to\mathcal{P}([q])\text{ s.t.}\dots} \mathbb{E}\big(F\big(B_T^{(1)}, B_T^{(2)}, \dots, B_T^{(q)}\big)\mathbf{1}_{A_u=f(u),\text{ for all }u\in\mathcal{T}_2} \mid \mathcal{L}_2, \mathcal{T}_2, m_2\big)\right)\end{aligned}$$

($\mathcal{P}$ defined in Section 1.5), where we sum on the $f : \mathcal{T}_2 \to \mathcal{P}([q])$ such that

$$\begin{cases} f(u) = \sqcup_{v\,:\,\boldsymbol{m}(v)=u} f(v) & \text{for all } u \text{ in } \mathcal{T}_2,\\ \sum_{u\in S(\alpha T)} \mathbf{1}_{\#f(u)=1} = k \text{ and } \sum_{u\in S(\alpha T)} (\#f(u)-1) = l. \end{cases} \tag{4.8}$$

We remind the reader that $\sqcup$ is defined in Section 1.5 (disjoint union), $\boldsymbol{m}(\cdot)$ is defined in Section 2.1 (mother), and $S(\dots)$ is defined in (4.2). Here, we mean that we sum over the f compatible with a description of tagged fragments.

If $u \in \mathcal{L}_2$ and $T_u < T$, then, conditionally on $\mathcal{T}_2$ and m_2, $B_T^{(n_u)}$ is independent of all the other variables and has the same law as $B_{T-T_u}^{(1)}$ (T_u is defined in (4.4), n_u in (4.7)). Thus, using Theorem 3.1 and Corollary 3.1, we get, for any $\varepsilon' \in (0, \theta-1)$, $u \in \mathcal{L}_2$,

$$\big|\mathbb{E}\big(f_{n_u}\big(B_T^{(n_u)}\big)|\mathcal{L}_2, \mathcal{T}_2, m_2\big)\big| \le \Gamma_1 \|f_{n_u}\|_\infty \mathrm{e}^{-(\theta-\varepsilon')(T-T_u)_+}$$
$$\text{for } T - T_u \notin Z_0,\ \text{ where } Z_0 \text{ is of Lebesgue measure zero.}$$

Thus, we get

$$\begin{aligned}&\frac{|\mathbb{E}\big(\mathbf{1}_{C_{k,l}(\alpha T)}\mathbf{1}_A F\big(B_T^{(1)}, B_T^{(2)}, \dots, B_T^{(q)}\big)\big)|}{\|F\|_\infty \Gamma_1^q} \quad \left\{\begin{array}{l}\text{(since } F \text{ is of the form } F = (f_1\otimes\cdots\otimes\\ f_q)_{\text{sym}}, \text{ since, conditionally on } u\in\mathcal{L}_2, \text{ the}\\ \text{distribution of } T_u \text{ is absolutely continuous}\\ \text{with respect to the Lebesgue measure)}\end{array}\right.\\
&\le \mathbb{E}\left(\sum_{f:\mathcal{T}_2\to\mathcal{P}([q])\text{ s.t.}\dots}\left[\prod_{u\in\mathcal{L}_2} \mathrm{e}^{-(\theta-\varepsilon')(T-T_u)_+}\mathbf{1}_A\mathbb{E}\big(\mathbf{1}_{A_u=f(u),\text{for all }u\in\mathcal{T}_2} \mid \mathcal{L}_2, \mathcal{T}_2, m_2\big)\right]\right)\\
&\qquad\qquad\text{(because of Assumption 2.3, and because } \theta-\varepsilon' > 1)\\
&\le \mathbb{E}\left(\sum_{f\,:\,\mathcal{T}_2\to\mathcal{P}([q])\text{ s.t.}\dots}\left[\prod_{u\in\mathcal{L}_2} \mathrm{e}^{-(T-T_{\boldsymbol{m}(u)})-\log(\delta)}\mathbf{1}_A\mathbb{E}\big(\mathbf{1}_{A_u=f(u),\text{for all }u\in\mathcal{T}_2} \mid \mathcal{L}_2, \mathcal{T}_2, m_2\big)\right]\right)\\
&\qquad\qquad\text{(because of (2.1); see full proof in Section A)}\\
&\le \mathbb{E}\left(\sum_{f\,:\,\mathcal{T}_2\to\mathcal{P}([q])\text{ s.t.}\dots}\mathbf{1}_A\left[\prod_{u\in\mathcal{L}_2} \mathrm{e}^{-(T-T_{\boldsymbol{m}(u)})-\log(\delta)} \prod_{u\in\mathcal{T}_2\backslash\{0\}} \mathrm{e}^{-(\#f(u)-1)(T_u-T_{\boldsymbol{m}(u)})}\right]\right).\end{aligned}$$

For a fixed ω and a fixed f, we have

$$\prod_{u\in\mathcal{L}_2} \mathrm{e}^{-(T-T_{\boldsymbol{m}(u)})-\log(\delta)} \prod_{u\in\mathcal{T}_2\backslash\{0\}} \mathrm{e}^{-(\#f(u)-1)(T_u-T_{\boldsymbol{m}(u)})} = \left(\frac{1}{\delta}\right)^{\#\mathcal{L}_2} \exp\left(-\int_0^T a(s)\,\mathrm{d}s\right),$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

where, for all s,

$$a(s) = \sum_{u\in\mathcal{L}_2\backslash\{0\}\,:\,T_{\mathbf{m}(u)}\le s<T} \mathbf{1}_{\#f(u)=1} + \sum_{u\in\mathcal{T}_2\backslash\{0\}\,:\,T_{\mathbf{m}(u)}\le s\le T_u} (\#f(u)-1)$$

$$\text{(if } u\in\mathcal{T}_2\backslash\mathcal{L}_2,\ \mathbf{1}_{\#f(u)=1}=0)$$

$$= \sum_{u\in\mathcal{T}_2\backslash\{0\}\,:\,T_{\mathbf{m}(u)}\le s<T} \mathbf{1}_{\#f(u)=1} + \sum_{u\in\mathcal{T}_2\backslash\{0\}\,:\,T_{\mathbf{m}(u)}\le s\le T_u} (\#f(u)-1)$$

$$(S(\cdot) \text{ defined in (4.2)})$$

$$\ge \sum_{u\in S(s)} \mathbf{1}_{\#f(u)=1} + \sum_{u\in S(s)} (\#f(u)-1).$$

We observe that, under (4.8),

$$a(t) \ge \left\lceil \frac{q}{2} \right\rceil \quad \text{for all } t, \qquad a(\alpha T)\ge k+l,$$

and, if t is such that $\sum_{u\in S(t)} \mathbf{1}_{\#f(u)=1}=k'$ and $\sum_{u\in S(t)}(\#f(u)-1)=l'$ for some integers k', l', then, for all $s\ge t$,

$$a(s)\ge k' + \left\lceil \frac{q-k'}{2} \right\rceil .$$

We observe that, under Assumption 2.3, there exists a constant which bounds $\#\mathcal{T}_2$ almost surely (because, for all u in $\mathcal{U}\backslash\{0\}$, $-\log(\xi_u)+\log(\xi_{\mathbf{m}(u)})\ge a$), and so there exists a constant $C_{tree}(q)$ which bounds $\#\{f:\mathcal{T}_2\to\mathcal{P}([q])\}$ almost surely. So, we have

$$\begin{aligned}
&\big|\mathbb{E}\big(\mathbf{1}_A F(B_T^{(1)}, B_T^{(2)}, \dots, B_T^{(q)})\big)\big| \\
&\le \|F\|_\infty \Gamma_1^q \mathbb{E}\left(\sum_{f:\mathcal{T}_2\to\mathcal{P}([q])\text{ s.t.}\dots} \mathbf{1}_A \left(\frac{1}{\delta}\right)^{\#\mathcal{L}_2} \mathrm{e}^{-\lceil q/2\rceil \alpha T} \exp\left\{-\left(k+\left\lceil\frac{q-k}{2}\right\rceil\right)(T-\alpha T)\right\}\right) \\
&\le \|F\|_\infty \Gamma_1^q C_{\text{tree}}(q) \left(\frac{1}{\delta}\right)^q \mathrm{e}^{-\lceil q/2\rceil \alpha T} \exp\left\{-\left(k+\left\lceil\frac{q-k}{2}\right\rceil\right)(1-\alpha)T\right\}.
\end{aligned} \tag{4.9}$$

Since $k\ge 1$, $k+\lceil (q-k)/2\rceil > q/2$, and so we have proved the desired result (remember that $T=-\log\varepsilon$). □

Remark 4.1. If we replaced Assumption 2.3 by Assumption 2.4, we would have difficulties adapting the above proof. In the second line of (4.9), the $1/\delta$ becomes $\mathrm{e}^{T_u-T_{\mathbf{m}(u)}}$. In addition, the tree $\mathcal{T}_2$ is no longer a.s. finite. So, the expectation on the second line of (4.9) could certainly be bounded, but for a high price (a lot more computations, maybe assumptions on the tails of π, and so on). This is why we stick with Assumption 2.3.

Remember that L_t ($t\ge 0$) is defined in (4.3).

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Lemma 4.2. *Let k be an integer, $k \geq q/2$, and let $\alpha \in [q/(2k), 1]$. Then we have $\mathbb{P}(L_{\alpha T} \geq k) \leq K_1(q)\varepsilon^{q/2}$, where*

$$K_1(q) = \sum_{i\in[q]} \frac{q!}{(q-i)!} \times i^{q-i}.$$

Let k be an integer, $k > q/2$, and let $\alpha \in (q/(2k), 1)$. Then

$$\varepsilon^{-q/2}\mathbb{P}(L_{\alpha T} \geq k) \underset{\varepsilon\to 0}{\longrightarrow} 0.$$

(We remind the reader that $T = -\log(\varepsilon)$.)

Proof. Let k be an integer, $k \geq q/2$, and let $\alpha \in [q/(2k), 1]$. Remember that $S(\cdot)$ is defined in (4.2). Observe that $\#S(\alpha T) = i$ if and only if $L_{\alpha T} = q - i$ (see (4.3)). We use the decomposition

$$\begin{aligned}\{L_{\alpha T} \geq k\} &= \{L_{\alpha T} \in \{k, k+1, \dots, q-1\}\}\\ &= \cup_{i\in[q-k]}\{\#S(\alpha T) = i\}\\ &= \cup_{i\in[q-k]} \cup_{m:[i]\hookrightarrow[q]} (F(i, m) \cap \{\#S(\alpha T) = i\})\end{aligned}$$

(remember that '$\hookrightarrow$' means we are summing on injections; see Section 1.5), where

$$\begin{aligned}F(i, m) = \{&i_1, i_2 \in [i] \text{ with } i_1 \neq i_2 \Rightarrow\\ &\text{there exists } u_1, u_2 \in S(\alpha T), u_1 \neq u_2, m(i_1) \in A_{u_1}, m(i_2) \in A_{u_2}\}.\end{aligned}$$

(To make the above equations easier to understand, observe that if $\#S(\alpha T) = i$, we have, for each $j \in [i]$, an index $m(j)$ in A_u for some $u \in S(\alpha T)$, and we can choose m such that we are in the event $F(i, m)$). Suppose we are in the event $F(i, m)$. For $u \in S(\alpha T)$ and for all j in $[i]$ such that $m(j) \in A_u$, we define (remember $|u|$ and $\mathbf{m}$ are defined in Section 2.1)

$$T^{(j)}_{|u|} = -\log(\xi_u),\ T^{(j)}_{|u|-1} = -\log(\xi_{\mathbf{m}(u)}),\ \dots,\ T^{(j)}_1 = -\log(\xi_{\mathbf{m}^{\circ(|u|-1)}(u)}), T^{(j)}_0 = 0,$$

with $l(j) = |u|$, $v(j) = u$. We have

$$\begin{aligned}\mathbb{P}(L_{\alpha T} \geq k) &\leq \sum_{i\in[q-k]}\sum_{m:[i]\hookrightarrow[q]} \mathbb{P}(F(i, m) \cap \{\#S(\alpha T) = i\})\\ &= \sum_{i\in[q-k]}\sum_{m:[i]\hookrightarrow[q]} \mathbb{E}\Big(\mathbf{1}_{F(i,m)}\mathbb{E}\Big(\mathbf{1}_{\#S(\alpha T)=i} \mid F(i, m), \big(T^{(j)}_p\big)_{j\in[i],p\in[l(j)]}, (v(j))_{j\in[i]}\Big)\Big)\\ &\qquad \text{(below, we sum over the partitions } \mathcal{B} \text{ of } [q]\backslash m([i]) \text{ into } i \text{ subsets } \mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_i)\\ &= \sum_{i\in[q-k]}\sum_{m:[i]\hookrightarrow[q]}\sum_{\mathcal{B}} \mathbb{E}\left(\mathbf{1}_{F(i,m)}\mathbb{E}\left(\prod_{j\in[i]}\prod_{r\in\mathcal{B}_j} \mathbf{1}_{r\in A_{v(j)}} \mid F(i, m), \big(T^{(j)}_p\big)_{j\in[i],p\in[l(j)]}, (v(j))_{j\in[i]}\right)\right)\\ &\qquad \text{(as } Y_1, \dots, Y_q \text{ defined in Section 2.2 are independent)}\end{aligned}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$= \sum_{i\in[q-k]} \sum_{m:[i]\hookrightarrow[q]} \sum_{\mathcal{B}} \mathbb{E}\left(\mathbf{1}_{F(i,m)} \prod_{j\in[i]} \prod_{r\in\mathcal{B}_j} \mathbb{E}\big(\mathbf{1}_{r\in A_{v(j)}} \mid F(i,m), \big(T_p^{(j)}\big)_{j\in[i],p\in[l(j)]}, (v(j))_{j\in[i]}\big)\right)$$

(because of (2.1) and (4.4))

$$= \sum_{i\in[q-k]} \sum_{m:[i]\hookrightarrow[q]} \sum_{\mathcal{B}} \mathbb{E}\left(\mathbf{1}_{F(i,m)} \prod_{j\in[i]} \prod_{r\in\mathcal{B}_j} \prod_{s=1}^{l(j)} \exp\left(-T_s^{(j)} + T_{s-1}^{(j)}\right)\right) \text{ (as } v(j)\in S(\alpha T))$$

$$\le \sum_{i\in[q-k]} \sum_{m:[i]\hookrightarrow[q]} \sum_{\mathcal{B}} \prod_{j\in[i]} \prod_{r\in\mathcal{B}_j} \mathrm{e}^{-\alpha T}$$

$$= \sum_{i\in[q-k]} \sum_{m:[i]\hookrightarrow[q]} \sum_{\mathcal{B}} \mathrm{e}^{-\alpha(q-i)T}$$

$$\le \sum_{i\in[q-k]} \sum_{m:[i]\hookrightarrow[q]} \sum_{\mathcal{B}} \mathrm{e}^{-k\alpha T} \le \mathrm{e}^{-k\alpha T} \sum_{i\in[q]} \frac{q!}{(q-i)!} i^{q-i}.$$

If we suppose that $k > q/2$ and $\alpha \in (q/(2k), 1)$, then

$$\exp\left(\frac{qT}{2}\right) \exp(-k\alpha T) \underset{T\to+\infty}{\longrightarrow} 0. \qquad \square$$

Immediate consequences of the two lemmas above are the following corollaries.

Corollary 4.1. *If q is odd and if $F \in \mathcal{B}^0_{\mathrm{sym}}(q)$ is of the form $F = (f_1 \otimes \cdots \otimes f_q)_{\mathrm{sym}}$, then*

$$\varepsilon^{-q/2}\mathbb{E}\big(F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\big) \underset{\varepsilon\to0}{\longrightarrow} 0.$$

($\mathcal{B}^0_{\mathrm{sym}}$ and $\mathcal{L}_1$ are defined in Section 4.1.)

Proof. We take $\alpha \in ((q/2)\lceil q/2\rceil^{-1}, 1)$. We observe that, for k in $[q]$, t in $(0, T)$,

$$\sum_{u\in S(t)} \mathbf{1}_{\#A_u=1} = k \Rightarrow \sum_{u\in S(t)} (\#A_u - 1) \in \{0, 1, \ldots, (q-k-1)_+\},$$

and (L_t is defined in (4.3)) $\sum_{u\in S(t)} \mathbf{1}_{\#A_u=1} = 0 \Rightarrow L_t \ge \lceil q/2\rceil$. So, we can use the decomposition

$$\begin{aligned}
&\varepsilon^{-q/2}\big|\mathbb{E}\big(F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\big)\big| \\
&= \Bigg|\varepsilon^{-q/2} \sum_{k\in[q]} \sum_{l\in\{0,1,\ldots,(q-k-1)_+\}} \mathbb{E}\big(\mathbf{1}_{C_{k,l}(\alpha T)}\mathbf{1}_{\#\mathcal{L}_1=q}F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\big) \\
&\qquad + \varepsilon^{-q/2}\mathbb{E}\big(\mathbf{1}_{L_{\alpha T}\ge\lceil q/2\rceil}\mathbf{1}_{\#\mathcal{L}_1=q}F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\big)\Bigg| \underset{\varepsilon\to0}{\longrightarrow} 0
\end{aligned} \tag{4.10}$$

(by (4.6) and Lemmas 4.1 and 4.2).

($\mathcal{L}_1$ and $\mathcal{L}_2$ are defined in Section 4.1.) $\square$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Corollary 4.2. *Suppose* $F \in \mathcal{B}^0_{\text{sym}}(q)$ *is of the form* $F = (f_1 \otimes \cdots \otimes f_q)_{\text{sym}}$. *Let* $A \in \sigma(\mathcal{L}_2)$. *Then*

$$\left|\mathbb{E}\big(F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\mathbf{1}_A\big)\right| \leq \|F\|_\infty \varepsilon^{q/2}\left\{K_1(q) + \Gamma_1^q C_{\text{tree}}(q)\left(\frac{1}{\delta}\right)^q (q+1)^2\right\}.$$

Proof. We get, as in (4.10),

$$\begin{aligned}
&\left|\mathbb{E}\big(F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\mathbf{1}_A\big)\right| \\
&\quad = \left|\mathbb{E}\left(F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\mathbf{1}_A\left(\mathbf{1}_{L_{\alpha T} \geq q/2} + \sum_{k' \in [q]} \sum_{0 \leq l \leq (q-k'-1)_+} \mathbf{1}_{C_{k',l}(\alpha T)}\right)\right)\right| \\
&\qquad\qquad \text{(from Lemmas 4.1 and 4.2)} \\
&\quad \leq \|F\|_\infty \varepsilon^{q/2}\left\{K_1(q) + \Gamma_1^q C_{\text{tree}}(q)\left(\frac{1}{\delta}\right)^q \sum_{k' \in [q]} 1 + (q-k'-1)_+\right\},
\end{aligned}$$

and $\sum_{k' \in [q]} 1 + (q-k'-1)_+ \leq (q+1)^2$ (see Section B for a detailed proof). □

We now want to find the limit of $\varepsilon^{-q/2}\mathbb{E}\big(\mathbf{1}_{L_T \leq q/2}\mathbf{1}_{\#\mathcal{L}_1 = q}F\big(B_T^{(1)}, \ldots, B_T^{(q)}\big)\big)$ when ε goes to 0, for q even. First we need a technical lemma.

For any i, the process $\big(B_t^{(i)}\big)$ has a stationary law (see [1, Theorem 3.3 p. 151]). Let B_∞ be a random variable having this stationary law η (it has already appeared in Section 3). We can always suppose that it is independent of all the other variables.

Fact 4.2. *From now on, when we have an* α *in* $(0, 1)$, *we suppose that* $\alpha T - \log(\delta) < (T + \alpha T)/2$ *and* $(T + \alpha T)/2 - \log(\delta) < T$ *(this is true if* T *is large enough). (The constant* δ *is defined in Assumption 2.3.)*

Lemma 4.3. *Let* f_1, f_2 *be in* $\mathcal{B}^0_{\text{sym}}(1)$. *Let* α *belong to* $(0, 1)$, *and* ε' *belong to* $(0, \theta - 1)$ *(*θ *is defined in Fact 3.1). We have*

$$\int_{-\infty}^{-\log(\delta)} \mathrm{e}^{-v}\left|\mathbb{E}\big(f_1\big(\overline{B}_0^{(1),v}\big)f_2\big(\overline{B}_0^{(2),v}\big)\big)\right| \,\mathrm{d}v < \infty, \tag{4.11}$$

and, almost surely, for T large enough,

$$\begin{aligned}
&\Bigg|\mathrm{e}^{T-\alpha T - B_{\alpha T}^{(1)}}\mathbb{E}\big(f_1 \otimes f_2\big(B_{T,}^{(1)} B_T^{(2)}\big)\mathbf{1}_{G_{1,2}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\big) \\
&\qquad - \int_{-\infty}^{-\log(\delta)} \mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v \leq \overline{B}_0^{(1),v}} f_1\big(\overline{B}_0^{(1),v}\big)f_2\big(\overline{B}_0^{(2),v}\big)\big)\,\mathrm{d}v\Bigg| \\
&\qquad\qquad \leq \Gamma_2\|f_1\|_\infty\|f_2\|_\infty \exp\left(-(T-\alpha T)\left(\frac{\theta - \varepsilon' - 1}{2}\right)\right),
\end{aligned}$$

where

$$\Gamma_2 = \frac{\Gamma_1^2}{\delta^{2+2(\theta-\varepsilon')}(2(\theta-\varepsilon')-1)} + \frac{\Gamma_1}{\delta^{\theta-\varepsilon'}} + \frac{\Gamma_1^2}{\delta^{2(\theta-\varepsilon')}(2(\theta-\varepsilon')-1)}.$$

(The processes $B^{(1)}$, $B^{(2)}$, $\overline{B}^{(1),v}$, *and* $\overline{B}^{(2),v}$ *are defined in Sections 2.3, 2.4, and 2.6.)*

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Proof. We have, for all s in $\left[\alpha T + B^{(1)}_{\alpha T}, T\right]$ (because of (2.1) and (4.4)),

$$\mathbb{P}\left(u\{s, 2\} = u\{s, 1\} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}\right) = \exp\left[-\left(s + B^{(1)}_s - \left(\alpha T + B^{(1)}_{\alpha T}\right)\right)\right]$$

(we remind the reader that $u\{s, 1\}$, $G_{1,2}$, are defined in Section 4.1, below (4.3)). Let us introduce the breaking time $\tau_{1,2}$ between 1 and 2 as a random variable having the following property: conditionally on $\mathcal{F}_{S(\alpha T)}$, $G_{\alpha T}$, and $\left(S^{(1)}_j\right)_{j\geq 1}$, $\tau_{1,2}$ has the density

$$s \in \mathbb{R} \mapsto \mathbf{1}_{[\alpha T + B^{(1)}_{\alpha T}, +\infty)}(s) \mathrm{e}^{-(s-(\alpha T + B^{(1)}_T))}$$

(this is a translation of an exponential law). We have the equalities $\alpha T + B^{(1)}_{\alpha T} = S^{(1)}_{j_0}$ for some j_0, and $T + B^{(1)}_T = S^{(1)}_{i_0}$ for some i_0. Here, we need to comment on the definitions of Section 4.1. In Fig. 6 we have $-\log(\xi_{(1,2)}) = S^{(1)}_2$ (as in (4.5)), $S(S^{(1)}_2) = \{(1, 2, 1), (1, 1, 1), (2, 1)\}$, and $u\{-\log(\xi_{(1,2)}), 1\} = \{(1, 2, 1)\}$. It is important to understand this example before reading what follows. The breaking time $\tau_{1,2}$ has the following interesting property (for all $k \geq j_0$):

$$\mathbb{P}\left(u\left\{S^{(1)}_k, 2\right\} \neq u\left\{S^{(1)}_k, 1\right\} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}\right)$$
$$= \mathbb{P}\left(\tau_{1,2} \in \left[\alpha T + B^{(1)}_{\alpha T}, S^{(1)}_k\right] \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}\right).$$

Just because we can, we impose, for all $k \geq j_0$, conditionally on $\mathcal{F}_{S(\alpha T)}$, $G_{\alpha T}$, and $\left(S^{(1)}_j\right)_{j\geq 1}$,

$$\left\{u\left\{S^{(1)}_k, 2\right\} \neq u\left\{S^{(2)}_k, 1\right\}\right\} = \left\{\tau_{1,2} \in \left[\alpha T + B^{(1)}_{\alpha T}, S^{(1)}_k\right]\right\}.$$

Now, let v be in $\left[\alpha T + B^{(1)}_{\alpha T}, T + B^{(1)}_T\right]$. We observe that, for all v in $\left[\alpha T + B^{(1)}_{\alpha T}, T + B^{(1)}_T\right]$,

$$\mathbb{E}\left(f_1 \otimes f_2\left(B^{(1)}_{T,} B^{(2)}_T\right) \mathbf{1}_{G_{1,2}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}, \tau_{1,2} = v\right)$$
$$= \mathbb{E}\left(f_1 \otimes f_2\left(\widehat{B}^{(1),v}_{T,} \widehat{B}^{(2),v}_T\right) \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}\right)$$

(because of (2.7)).

And so,

$$\mathbb{E}\left(f_1 \otimes f_2\left(B^{(1)}_{T,} B^{(2)}_T\right) \mathbf{1}_{G_{1,2}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\right)$$
$$= \mathbb{E}\left(\mathbb{E}\left(f_1 \otimes f_2\left(B^{(1)}_{T,} B^{(2)}_T\right) \mathbf{1}_{G_{1,2}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}\right) \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\right)$$
$$= \mathbb{E}\left(\mathbb{E}\left(\mathbb{E}\left(f_1 \otimes f_2\left(B^{(1)}_{T,} B^{(2)}_T\right) \mathbf{1}_{G_{1,2}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}, \tau_{1,2}\right) \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \left(S^{(1)}_j\right)_{j\geq 1}\right) \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\right)$$

(keep in mind that $\widehat{B}^{(1),v} = B^{(1)}$ for all v)

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$= \mathbb{E}\Big(\mathbb{E}\Big(\int_{\alpha T + B^{(1)}_{\alpha T}}^{T + B^{(1)}_T} \mathrm{e}^{-\left(v - \alpha T - \widehat{B}^{(1),v}_{\alpha T}\right)} \mathbb{E}\big(f_1 \otimes f_2\big(B^{(1)}_{T,} B^{(2)}_T\big) \mathbf{1}_{G_{1,2}(T)^{c}} \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \big(S^{(1)}_j\big)_{j \geq 1}, \tau_{1,2} = v\big)\, \mathrm{d}v$$

$$| \, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \big(S^{(1)}_j\big)_{j \geq 1}\Big) \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)$$

$$= \mathbb{E}\Big(\mathbb{E}\Big(\int_{\alpha T + B^{(1)}_{\alpha T}}^{T + B^{(1)}_T} \mathrm{e}^{-(v - \alpha T - \widehat{B}^{(1),v}_{\alpha T})} \mathbb{E}\Big(f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big) \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \big(S^{(1)}_j\big)_{j \geq 1}\Big)\, \mathrm{d}v$$

$$| \, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}, \big(S^{(1)}_j\big)_{j \geq 1}\Big) \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)$$

$$= \mathbb{E}\Big(\int_{\alpha T + B^{(1)}_{\alpha T}}^{T + B^{(1)}_T} \mathrm{e}^{-(v - \alpha T - \widehat{B}^{(1),v}_{\alpha T})} f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big)\, \mathrm{d}v \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big).$$

Let us split the above integral into two parts and multiply them by $\mathrm{e}^{T - \alpha T - B^{(1)}_{\alpha T}}$. For the first part:

$$\Big|\mathrm{e}^{T - \alpha T - B^{(1)}_{\alpha T}} \mathbb{E}\Big(\int_{\alpha T + B^{(1)}_{\alpha T}}^{(T+\alpha T)/2} \mathrm{e}^{-\left(v - \alpha T - \widehat{B}^{(1),v}_{\alpha T}\right)} f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big)\, \mathrm{d}v \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)\Big|$$

$$= \mathrm{e}^{T - \alpha T - B^{(1)}_{\alpha T}} \Big|\mathbb{E}\Big(\int_{\alpha T + B^{(1)}_{\alpha T}}^{(T+\alpha T)/2} \mathrm{e}^{-\left(v - \alpha T - \widehat{B}^{(1),v}_{\alpha T}\right)} \mathbb{E}\big(f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big) | \widehat{B}^{(1),v}_v, \widehat{B}^{(2),v}_v, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\big) \mathrm{d}v$$

$$| \, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)\Big|$$

(using the fact that $\widehat{B}^{(1),v}_T$ and $\widehat{B}^{(2),v}_T$ are independent conditionally on $\big\{\widehat{B}^{(1),v}_v, \widehat{B}^{(2),v}_v, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\big\}$ if $T \geq v - \log(\delta)$, we get, by Theorem 3.1, Corollary 3.1, and Fact 4.2)

$$\leq \mathrm{e}^{T - \alpha T - B^{(1)}_{\alpha T}} \mathbb{E}\Big(\int_{\alpha T + B^{(1)}_{\alpha T}}^{(T+\alpha T)/2} \mathrm{e}^{-(v - \alpha T - \widehat{B}^{(1)}_{\alpha T})}$$

$$\big(\Gamma_1 \|f_1\|_\infty \mathrm{e}^{-(\theta - \varepsilon')(T - v - \widehat{B}^{(1),v}_v)_+} + \Gamma_1 \|f_2\|_\infty \mathrm{e}^{-(\theta - \varepsilon')(T - v - \widehat{B}^{(2),v}_v)_+}\big)\, \mathrm{d}v \,|\, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)$$

(using Assumption 2.3)

$$\leq \Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty \mathrm{e}^{(T - \alpha T - \log(\delta))} \int_{\alpha T}^{(T+\alpha T)/2} \mathrm{e}^{-(v - \alpha T + \log(\delta))} \mathrm{e}^{-2(\theta - \varepsilon')(T - v + \log(\delta))}\, \mathrm{d}v$$

$$= \frac{\Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty}{\delta^{2 + 2(\theta - \varepsilon')}} \mathrm{e}^{(T - 2(\theta - \varepsilon')T)} \left[\frac{\mathrm{e}^{(2(\theta - \varepsilon') - 1)v}}{2(\theta - \varepsilon') - 1}\right]_{\alpha T}^{(T+\alpha T)/2}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$\leq \frac{\Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty}{\delta^{2+2(\theta-\varepsilon')}} \frac{\exp\left(-(2(\theta-\varepsilon')-1)T + (2(\theta-\varepsilon')-1)(T+\alpha T)/2\right)}{2(\theta-\varepsilon')-1}$$

$$= \frac{\Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty}{\delta^{2+2(\theta-\varepsilon')}} \frac{\exp\left(-(2(\theta-\varepsilon')-1)((T-\alpha T)/2)\right)}{2(\theta-\varepsilon')-1}. \tag{4.12}$$

For the second part, minus some other terms:

$$\Bigg| \underbrace{\mathrm{e}^{T-\alpha T-B^{(1)}_{\alpha T}} \mathbb{E}\Big(\int_{(T+\alpha T)/2}^{T+B^{(1)}_T} \mathrm{e}^{-(v-\alpha T-B^{(1)}_{\alpha T})} f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big)\,\mathrm{d}v \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)}_{\text{second part}}$$

$$- \underbrace{\int_{(T+\alpha T)/2}^{T-\log(\delta)} \mathrm{e}^{-(v-T)} \mathbb{E}\Big(\mathbf{1}_{v\leq T+\overline{B}^{(1),v}_T} f_1\big(\overline{B}^{(1),v}_T\big) f_2\big(\overline{B}^{(2),v}_T\big)\Big)\,\mathrm{d}v}_{(\heartsuit)} \Bigg|$$

$$= \Bigg| \mathrm{e}^{T-\alpha T-B^{(1)}_{\alpha T}} \mathbb{E}\Big(\int_{(T+\alpha T)/2}^{T-\log(\delta)} \mathrm{e}^{-(v-\alpha T-B^{(1)}_{\alpha T})} \mathbf{1}_{v\leq T+B^{(1)}_T} f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big)\,\mathrm{d}v \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)$$

$$- \mathrm{e}^{T-\alpha T-B^{(1)}_{\alpha T}} \mathbb{E}\Big(\int_{(T+\alpha T)/2}^{T-\log(\delta)} \mathrm{e}^{-(v-\alpha T-B^{(1)}_{\alpha T})} \mathbf{1}_{v\leq T+\overline{B}^{(1),v}_T} f_1\big(\overline{B}^{(1),v}_T\big) f_2\big(\overline{B}^{(2),v}_T\big)\,\mathrm{d}v\Big) \Bigg|$$

$$= \mathrm{e}^{T-\alpha T-B^{(1)}_{\alpha T}} \Bigg| \int_{(T+\alpha T)/2}^{T-\log(\delta)} \mathrm{e}^{-(v-\alpha T-B^{(1)}_{\alpha T})} \mathbb{E}\Big(\mathbb{E}\Big(\mathbf{1}_{v\leq T+B^{(1)}_T} f_1\big(\widehat{B}^{(1),v}_T\big) f_2\big(\widehat{B}^{(2),v}_T\big)$$

$$\mid \widehat{C}^{(1),v}_v, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big) \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\Big)\,\mathrm{d}v$$

$$- \int_{(T+\alpha T)/2}^{T-\log(\delta)} \mathrm{e}^{-(v-\alpha T-B^{(1)}_{\alpha T})} \mathbb{E}\Big(\mathbb{E}\Big(\mathbf{1}_{v\leq T+\overline{B}^{(1),v}_T} f_1\big(\overline{B}^{(1),v}_T\big) f_2\big(\overline{B}^{(2),v}_T\big) \mid \overline{C}^{(1),v}_v\Big)\Big)\,\mathrm{d}v\Bigg|. \tag{4.13}$$

We observe that, for all v in $[(T+\alpha T)/2, T-\log(\delta)]$, once $\widehat{C}^{(1),v}_v$ is fixed, we can make a simulation of $\widehat{B}^{(1),v}_T = B^{(1)}_T$ and $\widehat{B}^{(2),v}_T$ (these processes are independent of $\mathcal{F}_{S(\alpha T)}, G_{\alpha T}$ conditionally on $\widehat{C}^{(1),v}_v$). Indeed, we draw $\widehat{B}^{(1),v}_v$ conditionally on $\widehat{C}^{(1),v}_v$ (with law $\eta_1\big(\cdot \mid \widehat{C}^{(1),v}_v\big)$ defined in Fact 2.2), then we draw $\widehat{B}^{(2),v}_v$ conditionally on $\widehat{B}^{(1),v}_v$ and $\widehat{C}^{(1),v}_v$ (with law $\eta'\big(\cdot \mid \widehat{B}^{(1),v}_v, \widehat{C}^{(1),v}_v\big)$, see Fact 2.2). Then, $\big(\widehat{B}^{(1),v}_t\big)_{t\geq v}$ and $\big(\widehat{B}^{(2),v}_t\big)_{t\geq v}$ run their courses as independent Markov processes, until we get $\widehat{B}^{(1),v}_T, \widehat{B}^{(2),v}_T$.

In the same way (for all v in $[(T+\alpha T)/2, T-\log(\delta)]$), we observe that the process $\big(\overline{C}^{(1),v}, \overline{B}^{(1),v}\big)$ starts at time $v-2b$ and has the same transition as $\big(C^{(1)}, B^{(1)}\big)$ (see (2.6)). By Assumption 2.1, the following time exists: $S = \sup\big\{t : v-b \leq t \leq v,\ \overline{C}^{(1),v}_t = 0\big\}$. We then have $v - S = \overline{C}^{(1),v}_v$. When $\overline{C}^{(1),v}_v$ is fixed, this entails that $\overline{B}^{(1),v}_v$ has the law $\eta_1\big(\cdot \mid \overline{C}^{(1),v}_v\big)$. We have $\overline{B}^{(2),v}_v$ of law $\eta'\big(\cdot \mid \overline{C}^{(1),v}_v, \overline{B}^{(1),v}_v\big)$ (by (2.10)). As before, we then let

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

the process $\left(\overline{B}_t^{(1),v}, \overline{B}_t^{(2),v}\right)_{t\geq v}$ run its course as a Markov process having the same transition as $\left(\widehat{B}_{t-v+kb}^{(1),kb}, \widehat{B}_{t-v+kb}^{(2),kb}\right)_{t\geq v}$ until we get $\overline{B}_T^{(1),v}, \overline{B}_T^{(2),v}$.

So we get that (for all v in $[(T+\alpha T)/2, T-\log(\delta)]$)

$$\mathbb{E}\left(\mathbf{1}_{v\leq T+B_T^{(1)}} f_1\left(\widehat{B}_T^{(1),v}\right) f_2\left(\widehat{B}_T^{(2),v}\right) \mid \widehat{C}_v^{(1),v}, \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\right) = \Psi\left(\widehat{C}_v^{(1),v}\right),$$

$$\mathbb{E}\left(\mathbf{1}_{v\leq T+\overline{B}_T^{(1),v}} f_1\left(\overline{B}_T^{(1),v}\right) f_2\left(\overline{B}_T^{(2),v}\right) \mid \overline{C}_v^{(1),v}\right) = \Psi\left(\overline{C}_v^{(1),v}\right) \overset{\text{law}}{=} \Psi(C_\infty)$$

for some function Ψ, the same in both lines, such that $\|\Psi\|_\infty \leq \|f_1\|_\infty \|f_2\|_\infty$ (where C_∞ is defined in Section 3). So, by Theorem 3.1 and Corollary 3.1 applied on the time interval $\left[\alpha T + B_{\alpha T}^{(1)}, v\right]$, the quantity in (4.14) can be bounded (remember that $\widehat{C}^{(1),v} = C^{(1)}$, Section 2.5) by

$$e^{T-\alpha T-B_{\alpha T}^{(1)}} \int_{(T+\alpha T)/2}^{T-\log(\delta)} e^{-(v-\alpha T-B_{\alpha T}^{(1)})} \Gamma_1 \|f_1\|_\infty \|f_2\|_\infty e^{-(\theta-\varepsilon')(v-\alpha T-B_{\alpha T}^{(1)})}\, dv.$$

(Coming from Corollary 3.1 there is an integral over a set of Lebesgue measure zero in the above bound, but this term vanishes.) The above bound can in turn be bounded by

$$\Gamma_1 \|f_1\|_\infty \|f_2\|_\infty \delta^{-(\theta-\varepsilon')} e^T \int_{(T+\alpha T)/2}^{T-\log(\delta)} e^{(\theta-\varepsilon')\alpha T} e^{-(\theta-\varepsilon'+1)v}\, dv$$

(as $\theta-\varepsilon'+1>1$)

$$\leq \Gamma_1 \|f_1\|_\infty \|f_2\|_\infty \delta^{-(\theta-\varepsilon')} e^{T+\alpha T(\theta-\varepsilon')} \exp\left[-(\theta-\varepsilon'+1)\left(\frac{T+\alpha T}{2}\right)\right]$$

$$= \Gamma_1 \|f_1\|_\infty \|f_2\|_\infty \delta^{-(\theta-\varepsilon')} \exp\left[-(\theta-\varepsilon'-1)\left(\frac{T-\alpha T}{2}\right)\right]. \tag{4.14}$$

We have

$$\int_{\frac{T+\alpha T}{2}}^{T-\log(\delta)} e^{-(v-T)} \mathbb{E}\left(\mathbf{1}_{v\leq T+\overline{B}_T^{(1),v}} f_1\left(\overline{B}_T^{(1),v}\right) f_2\left(\overline{B}_T^{(2),v}\right)\right) dv$$

(as $\left(\overline{B}_T^{(1),v}, \overline{B}_T^{(2),v}\right)$ and $\left(\overline{B}_0^{(1),v-T}, \overline{B}_0^{(2),v-T}\right)$ have same law)

$$= \int_{\frac{T+\alpha T}{2}}^{T-\log(\delta)} e^{-(v-T)} \mathbb{E}\left(\mathbf{1}_{v-T\leq \overline{B}_0^{(1),v-T}} f_1\left(\overline{B}_0^{(1),v-T}\right) f_2\left(\overline{B}_0^{(2),v-T}\right)\right) dv$$

(change of variable $v' = v-T$)

$$= \mathbb{E}\left(\int_{-\left(\frac{T-\alpha T}{2}\right)}^{-\log(\delta)} e^{-v'} \mathbf{1}_{v'\leq \overline{B}_0^{(1),v'}} f_1\left(\overline{B}_0^{(1),v'}\right) f_2\left(\overline{B}_0^{(1),v'}\right) dv'\right) \tag{4.15}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

and

$$\int_{-\infty}^{-\frac{(T-\alpha T)}{2}} \mathrm{e}^{-v} \big|\mathbb{E}\big(f_1\big(\overline{B}_0^{(1),v}\big)f_2\big(\overline{B}_0^{(2),v}\big)\big)\big|\, \mathrm{d}v$$

(since $\overline{B}_0^{(1),v}$ and $\overline{B}_0^{(2),v}$ are independent conditionally on $\overline{B}_v^{(1),v}$, $\overline{B}_v^{(2),v}$ if $v - \log(\delta) \leq 0$; using Theorem 3.1 and Corollary 3.1)

$$\leq \int_{-\infty}^{-\frac{(T-\alpha T)}{2}} \mathrm{e}^{-v}\Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty \mathbb{E}\big(\mathrm{e}^{-(\theta-\varepsilon')(-v-\overline{B}_v^{(1),v})_+}\mathrm{e}^{-(\theta-\varepsilon')(-v-\overline{B}_v^{(2),v})_+}\big)\, \mathrm{d}v$$

(again, coming from Corollary 3.1 there is an integral over a set of Lebesgue measure zero in the above bound, but this term vanishes)

$$\leq \int_{-\infty}^{-\frac{(T-\alpha T)}{2}} \mathrm{e}^{-v}\Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty \mathrm{e}^{-2(\theta-\varepsilon')(-v+\log(\delta))}\, \mathrm{d}v$$

$$= \frac{\Gamma_1^2 \|f_1\|_\infty \|f_2\|_\infty}{\delta^{2(\theta-\varepsilon')}} \frac{\exp(-(2(\theta-\varepsilon')-1)(T-\alpha T)/2)}{2(\theta-\varepsilon')-1}. \tag{4.16}$$

Equations (4.16) and (4.17) give us (4.11). Equations (4.13) and (4.15)–(4.17) give us the desired result (see below to understand the puzzle).

$$\mathrm{e}^{T-\alpha T-B_{\alpha T}^{(1)}}\mathbb{E}\big(f_1 \otimes f_2\big(B_{T,}^{(1)}B_T^{(2)}\big)\mathbf{1}_{G_{1,2}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\big)$$

↙ First part (in (4.13)) ↓ Small

↓ Second part (in (4.14))

↓ Close to some term (♡) (see (4.14), (4.15))

↓ (♡) close to

$$-\int_{-\infty}^{-\log(\delta)} \mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v\leq \overline{B}_0^{(1),v}} f_1\big(\overline{B}_0^{(1),v}\big)f_2\big(\overline{B}_0^{(2),v}\big)\big)\, \mathrm{d}v$$

(see (4.16), (4.17)) □

Lemma 4.4. *Let $k \in \{0, 1, 2, \ldots, p\}$. We suppose that q is even and $q = 2p$. Let $\alpha \in (q/(q+2), 1)$. We suppose that $F = f_1 \otimes f_2 \otimes \cdots \otimes f_q$, with $f_1, \ldots, f_q$ in $\mathcal{B}^0_{\mathrm{sym}}(1)$. Then,*

$$\varepsilon^{-q/2}\mathbb{E}\big(F\big(B_{T,}^{(1)}\ldots, B_T^{(q)}\big)\mathbf{1}_{G_{\alpha T}}\mathbf{1}_{\#\mathcal{L}_1=q}\big)$$

$$\underset{\varepsilon\to 0}{\longrightarrow} \prod_{i=1}^{p} \int_{-\infty}^{-\log(\delta)} \mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v\leq \overline{B}_0^{(1),v}} f_{2i-1}\big(\overline{B}_0^{(1),v}\big)f_{2i}\big(\overline{B}_0^{(2),v}\big)\big)\, \mathrm{d}v. \tag{4.17}$$

(Remember that $T = -\log\varepsilon$.)

Proof. By Fact 4.1, we have $T > \alpha T - \log(\delta)$. We have (remember the definitions just before Section 4.2)

$$G_{\alpha T} \cap \{\#\mathcal{L}_1 = q\} = G_{\alpha T} \cap \bigcap_{1\leq i\leq p} G_{2i-1,2i}(T)^{\mathrm{c}}.$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

We have (remember $T = -\log(\varepsilon)$)

$$
\begin{aligned}
&\varepsilon^{-q/2}\mathbb{E}\big(F\big(B_{T,}^{(1)}\dots, B_T^{(q)}\big)\mathbf{1}_{G_{\alpha T}}\mathbf{1}_{\#\mathcal{L}_1=q}\big)\\
&= \mathrm{e}^{pT}\mathbb{E}\left(\mathbf{1}_{G_{\alpha T}}\mathbb{E}\left(\prod_{i=1}^{p} f_{2i-1}\otimes f_{2i}\big(B_T^{(2i-1)}, B_T^{(2i)}\big)\mathbf{1}_{G_{2i-1,2i}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{\alpha T}\right)\right)\\
&\qquad\qquad \text{(as } \big(B_T^{(1)}, B_T^{(2)}, \mathbf{1}_{G_{1,2}(T)}\big), \big(B_T^{(3)}, B_T^{(4)}, \mathbf{1}_{G_{3,4}(T)}\big), \dots \text{ are independent}\\
&\qquad\qquad \text{conditionally on } \mathcal{F}_{S(\alpha T)}, G_{\alpha T} \text{ due to Fact 4.2)}\\
&= \mathbb{E}\left(\mathbf{1}_{G_{\alpha T}}\prod_{i=1}^{p}\mathrm{e}^{T}\mathbb{E}\big(f_{2i-1}\otimes f_{2i}\big(B_T^{(2i-1)}, B_T^{(2i)}\big)\mathbf{1}_{G_{2i-1,2i}(T)^{\mathrm{c}}} \mid \mathcal{F}_{S(\alpha T)}, G_{2i-1,2i}(\alpha T)\big)\right)\\
&\qquad\qquad \text{(by Lemma 4.3 and as } \big(B^{(1)}, \dots, B^{(q)}\big) \text{ is exchangeable)}\\
&= \mathbb{E}\left(\mathbf{1}_{G_{\alpha T}}\prod_{i=1}^{p}\mathrm{e}^{\alpha T + B_{\alpha T}^{(2i-1)}}\left(\int_{-\infty}^{-\log(\delta)}\mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v\le \overline{B}_0^{(1),v}} f_{2i-1}\big(\overline{B}_0^{(1),v}\big) f_{2i}\big(\overline{B}_0^{(2),v}\big)\big)\,\mathrm{d}v + R_{2i-1,2i}\right)\right),
\end{aligned}
\tag{4.18}
$$

with (a.s.)

$$
|R_{2i-1,2i}| \le \Gamma_2\|f_{2i-1}\|_\infty\|f_{2i}\|_\infty \mathrm{e}^{-(T-\alpha T)(\theta-\varepsilon'-1)/2}. \tag{4.19}
$$

We introduce the events (for $t\in[0, T]$, with $u\{\cdot\}$ defined below (4.3))

$$
O_t = \{\#\{u\{t, 2i-1\}, 1\le i\le p\} = p\},
$$

and the tribes (for i in $[q]$, $t\in[0, T]$) $\mathcal{F}_{t,i} = \sigma(u\{t, i\}, \xi_{u\{t,i\}})$. As $G_{\alpha T} = O_{\alpha T}\cap \bigcap_{1\le i\le p}\{u\{\alpha T, 2i-1\} = u\{\alpha T, 2i\}\}$, we have

$$
\begin{aligned}
&\mathbb{E}\left(\mathbf{1}_{G_{\alpha T}}\prod_{i=1}^{p}\mathrm{e}^{B_{\alpha T}^{(2i-1)}+\alpha T}\right)\\
&\quad= \mathbb{E}\left(\mathbf{1}_{O_{\alpha T}}\prod_{i=1}^{p}\mathrm{e}^{B_{\alpha T}^{(2i-1)}+\alpha T}\mathbb{E}\left(\prod_{i=1}^{p}\mathbf{1}_{u\{\alpha T,2i-1\}=u\{\alpha T,2i\}} \mid \vee_{1\le i\le p}\mathcal{F}_{\alpha T,2i-1}\right)\right)\\
&\qquad\qquad \text{(by Proposition 2.1 and (2.1))}\\
&\quad= \mathbb{E}(\mathbf{1}_{O_{\alpha T}}).
\end{aligned}
\tag{4.20}
$$

We then observe that

$$
O_{\alpha T}^{\mathrm{c}} = \cup_{i\in[p]}\cup_{j\in[p], j\neq i}\{u\{\alpha T, 2i-1\} = u\{\alpha T, 2j-1\}\},
$$

and, for $i\neq j$,

$$
\begin{aligned}
\mathbb{P}(u\{\alpha T, 2i-1\} = u\{\alpha T, 2j-1\}) &= \mathbb{E}\big(\mathbb{E}\big(\mathbf{1}_{u\{\alpha T,2i-1\}=u\{\alpha T,2j-1\}} \mid \mathcal{F}_{\alpha T,2i-1}\big)\big)\\
&= \mathbb{E}\big(\mathrm{e}^{-\alpha T - B_{\alpha T}^{(2i-1)}}\big) \qquad \text{(by Proposition 2.1 and (2.1))}\\
&\le \mathrm{e}^{-\alpha T - \log(\delta)}. \qquad \text{(because of Assumption (2.3))}
\end{aligned}
$$

So, $\mathbb{P}(O_{\alpha T}) \underset{\varepsilon\to 0}{\longrightarrow} 1$. This gives us enough material to finish the proof of (4.18).

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Indeed, starting from (4.19), we have

$$\begin{aligned}
&\varepsilon^{-q/2}\mathbb{E}\big(F\big(B^{(1)}_{T,},\ldots,B^{(q)}_T\big)\mathbf{1}_{G_{\alpha T}}\mathbf{1}_{\#\mathcal{L}_1=q}\big)\\
&=\mathbb{E}\left(\mathbf{1}_{G_{\alpha T}}\prod_{i=1}^{p}\mathrm{e}^{\alpha T+B^{(2i-1)}_{\alpha T}}\right)\prod_{i=1}^{p}\left(\int_{-\infty}^{-\log(\delta)}\mathrm{e}^{-v}\mathbb{E}(\mathbf{1}_{v\le\overline{B}^{(1),v}_0}f_{2i-1}\big(\overline{B}^{(1),v}_0\big)f_{2i}\big(\overline{B}^{(2),v}_0\big))\,\mathrm{d}v\right)\\
&\quad+\mathbb{E}\left(\mathbf{1}_{G_{\alpha T}}\prod_{i=1}^{p}\mathrm{e}^{\alpha T+B^{(2i-1)}_{\alpha T}}R_{2i-1,2i}\right)=:(\mathrm{I})+(\mathrm{II}).
\end{aligned}$$

By (4.21),

$$\begin{aligned}
(\mathrm{I})&=\mathbb{P}(O_{\alpha t})\prod_{i=1}^{p}\left(\int_{-\infty}^{-\log(\delta)}\mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v\le\overline{B}^{(1),v}_0}f_{2i-1}\big(\overline{B}^{(1),v}_0\big)f_{2i}\big(\overline{B}^{(2),v}_0\big)\big)\,\mathrm{d}v\right)\\
&\underset{\varepsilon\to0}{\longrightarrow}\prod_{i=1}^{p}\left(\int_{-\infty}^{-\log(\delta)}\mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v\le\overline{B}^{(1),v}_0}f_{2i-1}\big(\overline{B}^{(1),v}_0\big)f_{2i}\big(\overline{B}^{(2),v}_0\big)\big)\,\mathrm{d}v\right).
\end{aligned}$$

And, by (4.20),

$$(\mathrm{II})\le\mathbb{P}(O_{\alpha t})\prod_{i=1}^{p}\big(\Gamma_2\|f_{2i-1}\|_\infty\|f_{2i}\|_\infty\mathrm{e}^{-(T-\alpha T)(\theta-\varepsilon'-1)/2}\big)\underset{\varepsilon\to0}{\longrightarrow}0.\qquad\square$$

4.3. Convergence result

For f and g bounded measurable functions, we set

$$V(f,g)=\int_{-\infty}^{-\log(\delta)}\mathrm{e}^{-v}\mathbb{E}\big(\mathbf{1}_{v\le\overline{B}^{(1),v}_0}f\big(\overline{B}^{(1),v}_0\big)g\big(\overline{B}^{(2),v}_0\big)\big)\,\mathrm{d}v.\tag{4.21}$$

For q even, we set $\mathcal{I}_q$ to be the set of partitions of $[q]$ into subsets of cardinality 2. We have

$$\#\mathcal{I}_q=\frac{q!}{(q/2)!2^{q/2}}.\tag{4.22}$$

For I in $\mathcal{I}_q$ and t in $[0,T]$, we introduce

$$G_{t,I}=\{\text{for all }\{i,j\}\in I,\ \text{there exists }u\in\mathcal{U}\text{ such that }\xi_u<\mathrm{e}^{-t},\ \xi_{\boldsymbol{m}(u)}\ge\mathrm{e}^{-t},\ A_u=\{i,j\}\}.$$

For t in $[0, T]$, we define $\mathcal{P}_t=\cup_{I\in\mathcal{I}_q}G_{t,I}$. The above event can be understood as ‘at time t, the dots are paired on different fragments’. As before, the reader has to keep in mind that $T=-\log(\varepsilon)$, see (2.3).

Proposition 4.1. *Let q be in $\mathbb{N}^*$. Let $F=(f_1\otimes\cdots\otimes f_q)_{\mathrm{sym}}$ with $f_1,\ldots,f_q$ in $\mathcal{B}^0_{\mathrm{sym}}(1)$ ($(\cdot)_{\mathrm{sym}}$ defined in (4.1)). If q is even ($q=2p$) then*

$$\varepsilon^{q/2}\mathbb{E}\big(F\big(B^{(1)}_T,\ldots,B^{(q)}_T\big)\mathbf{1}_{\#\mathcal{L}_1=q}\big)\underset{\varepsilon\to0}{\longrightarrow}\sum_{I\in\mathcal{I}_q}\prod_{\{a,b\}\in I}V(f_a,f_b).\tag{4.23}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Proof. Let α be in $(q/(q+2), 1)$. We have

$$\varepsilon^{-q/2}\mathbb{E}\big(F\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\big) = \varepsilon^{-q/2}\mathbb{E}\big(F\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}(\mathbf{1}_{\mathcal{P}_{\alpha T}} + \mathbf{1}_{\mathcal{P}^{c}_{\alpha T}})\big).$$

Remember that the events of the form $C_{k,l}(t)$, L_t are defined in Section 4.1. The set $\mathcal{P}^{c}_{\alpha T}$ is a disjoint union of sets of the form $C_{k,l}(\alpha T)$ (with $k \geq 1$) and $\{L_{\alpha T} > q/2\}$ (this can be understood heuristically by: 'if the dots are not paired on fragments then some of them are alone on their fragment, or none of them is alone on a fragment and some are a group of at least three on a fragment'). As before, the event $\{\#\mathcal{L}_1 = q\}$ is measurable with respect to $\mathcal{L}_2$ (see (4.6)). So, by Lemmas 4.1 and 4.2,

$$\lim_{\varepsilon\to 0} \varepsilon^{-q/2}\mathbb{E}\big(F\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\mathbf{1}_{\mathcal{P}^{c}_{\alpha T}}\big) = 0.$$

We compute:

$$\varepsilon^{-q/2}\mathbb{E}\big(F\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\mathbf{1}_{\mathcal{P}_{\alpha T}}\big) = \varepsilon^{-q/2}\mathbb{E}\Bigg(F\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\sum_{I_q\in\mathcal{I}_q}\mathbf{1}_{G_{\alpha T, I_q}}\Bigg)$$

(as F is symmetric and $\big(B_T^{(1)}, \dots, B_T^{(q)}\big)$ is exchangeable)

$$= \frac{q!}{2^{q/2}(q/2)!}\varepsilon^{-q/2}\mathbb{E}\big(F\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\mathbf{1}_{G_{\alpha T}}\big)$$

$$= \frac{q!\varepsilon^{-q/2}}{2^{q/2}(q/2)!}\frac{1}{q!}\sum_{\sigma\in\mathcal{S}_q}\mathbb{E}\big((f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(q)})\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1=q}\mathbf{1}_{G_{\alpha T}}\big)$$

(by Lemma 4.4)

$$\underset{\varepsilon\to 0}{\longrightarrow} \frac{1}{2^{q/2}(q/2)!}\sum_{\sigma\in\mathcal{S}_q}\prod_{i=1}^{p} V(f_{\sigma(2i-1)}, f_{\sigma(2i)}) = \sum_{I\in\mathcal{I}_q}\prod_{\{a,b\}\in I} V(f_a, f_b).$$ □

5. Results

We are interested in the probability measure γ_T defined by its action on bounded measurable functions $F : [0, 1] \to \mathbb{R}$ by

$$\gamma_T(F) = \sum_{u\in\mathcal{U}_\varepsilon} \xi_u F\left(\frac{\xi_u}{\varepsilon}\right).$$

We define, for all $q \in \mathbb{N}^*$ and F from $[0, 1]^q$ to $\mathbb{R}$,

$$\gamma_T^{\otimes q}(F) = \sum_{a : [q]\to\mathcal{U}_\varepsilon} \xi_{a(1)}\cdots\xi_{a(q)}F\left(\frac{\xi_{a(1)}}{\varepsilon}, \dots, \frac{\xi_{a(q)}}{\varepsilon}\right),$$

$$\gamma_T^{\odot q}(F) = \sum_{a : [q]\hookrightarrow\mathcal{U}_\varepsilon} \xi_{a(1)}\cdots\xi_{a(q)}F\left(\frac{\xi_{a(1)}}{\varepsilon}, \dots, \frac{\xi_{a(q)}}{\varepsilon}\right),$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

where the last sum is taken over all the injective applications a from $[q]$ to $\mathcal{U}_\varepsilon$. We set

$$\Phi(F) : (y_1, \dots, y_q) \in \mathbb{R}^+ \mapsto F(\mathrm{e}^{-y_1}, \dots, \mathrm{e}^{-y_q}).$$

The law $\gamma^{\otimes q}$ is the law of q fragments picked in $\mathcal{U}_\varepsilon$ with replacement. For each fragment, the probability of being picked is its size. The measure $\gamma^{\odot q}$ is not a law: $\gamma^{\odot q}(F)$ is an expectation over q fragments picked in $\mathcal{U}_\varepsilon$ with replacement (for each fragment, the probability of being picked is its size); in this expectation, we multiply the integrand by zero if two fragments are the same (and by one otherwise). The definition of Section 2.2 says that we can define the tagged fragment by painting colored dots on the stick [0, 1] (q dots of different colors, these are the $Y_1, \dots, Y_q$) and then by looking on which fragments of $\mathcal{U}_\varepsilon$ we have these dots. So, we get (remember $T = -\log \varepsilon$)

$$\begin{aligned} \mathbb{E}\big(\gamma_T^{\otimes q}(F)\big) &= \mathbb{E}\big(\Phi(F)\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\big), \\ \mathbb{E}\big(\gamma_T^{\odot q}(F)\big) &= \mathbb{E}\big(\Phi(F)\big(B_T^{(1)}, \dots, B_T^{(q)}\big)\mathbf{1}_{\#\mathcal{L}_1 = q}\big). \end{aligned} \tag{5.1}$$

We define, for all bounded continuous $f : \mathbb{R}^+ \to \mathbb{R}$,

$$\gamma_\infty(f) = \eta(\Phi(f)). \tag{5.2}$$

Proposition 5.1. (Law of large numbers.) *We remind the reader that we have Fact 3.1, and that we are under Assumptions 2.1, 2.2, 2.3, and 3.1. Let f be a continuous function from* [0, 1] *to* $\mathbb{R}$. *Then*

$$\gamma_T(f) \xrightarrow[T \to +\infty]{\text{a.s.}} \gamma_\infty(f).$$

(Remember $T = -\log \varepsilon$.)

Proof. We take a bounded measurable function $f : [0, 1] \to \mathbb{R}$. We define $\overline{f} = f - \eta(\Phi(f))$. We take an integer $q \geq 2$. We introduce the notation

$$\text{for all } g : \mathbb{R}^+ \to \mathbb{R} \text{ and all } (x_1, \dots, x_q) \in \mathbb{R}^q, \quad g^{\otimes q}(x_1, \dots, x_q) = g(x_1)g(x_2)\dots g(x_q).$$

We have

$$\begin{aligned} \mathbb{E}((\gamma_T(f) - \eta(\Phi(f)))^q) &= \mathbb{E}\big(\big(\gamma_T(\overline{f})\big)^q\big) \\ &= \mathbb{E}\big(\gamma_t^{\otimes q}\big(\overline{f}^{\otimes q}\big)\big) \qquad \text{(as } \big(B^{(1)}, \dots, B^{(q)}\big) \text{ is exchangeable)} \\ &= \mathbb{E}\big(\gamma_t^{\otimes q}\big(\big(\overline{f}^{\otimes q}\big)_{\text{sym}}\big)\big) \\ &\leq \|\overline{f}\|_\infty^q \varepsilon^{q/2} \left\{ K_1(q) + \Gamma_1^q C_{\text{tree}}(q) \left(\frac{1}{\delta}\right)^q (q+1)^2 \right\}. \\ & \qquad \text{(by Corollary 4.2)} \end{aligned}$$

We now take sequences $(T_n = \log(n))_{n \geq 1}$, $(\varepsilon_n = 1/n)_{n \geq 1}$. We then have, for all n and for all $\iota > 0$,

$$\mathbb{P}([\gamma_{T_n}(f) - \eta(\Phi(f))]^4 \geq \iota) \leq \frac{\|\overline{f}\|_\infty^4}{\iota n^2} \left\{ K_1(4) + \Gamma_1^4 C_{\text{tree}}(4) \left(\frac{1}{\delta}\right)^4 \times 25 \right\}.$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

So, by the Borell–Cantelli lemma,

$$\gamma_{T_n}(f) \xrightarrow[n\to+\infty]{\text{a.s.}} \eta(\Phi(f)). \tag{5.3}$$

We now have a little more work to do to get to the result. Let n be in $\mathbb{N}^*$. We use the decomposition (where $\mathcal{U}_\varepsilon$ is defined in Section 2.3 and $\sqcup$ stands for 'disjoint union', defined in Section 1.5) $\mathcal{U}_{\varepsilon_n} = \mathcal{U}_{\varepsilon_n}^{(1)} \sqcup \mathcal{U}_{\varepsilon_n}^{(2)}$, where $\mathcal{U}_{\varepsilon_n}^{(1)} = \mathcal{U}_{\varepsilon_n} \cap \mathcal{U}_{\varepsilon_{n+1}} = \mathcal{U}_{\varepsilon_{n+1}}, \mathcal{U}_{\varepsilon_n}^{(2)} = \mathcal{U}_{\varepsilon_n} \backslash \mathcal{U}_{\varepsilon_{n+1}}$. For u in $\mathcal{U}_{\varepsilon_n} \backslash \mathcal{U}_{\varepsilon_{n+1}}$, we set $\boldsymbol{d}(u) = \{v : u = \mathbf{m}(v)\}$ ($\mathbf{m}$ is defined in Section 2.1) and we observe that, for all u (T_u defined in (4.4)),

$$\sum_{v\in \boldsymbol{d}(u)} \xi_v = \xi_u. \tag{5.4}$$

We can then write

$$\sum_{u\in\mathcal{U}_{\varepsilon_n}} \xi_u f\left(\frac{\xi_u}{\varepsilon_n}\right) = \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u f(n\xi_u) + \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \xi_u f(n\xi_u).$$

There exists n_1 such that, for n bigger than n_1, $\mathrm{e}^{-a} < \varepsilon_{n+1}/\varepsilon_n$ (remember Assumption 2.3). We suppose $n \geq n_1$; we then have, for all u in $\mathcal{U}_{\varepsilon_n}^{(2)}$, $\varepsilon_n > \xi_u \geq \varepsilon_{n+1}$ and, for any v in $\mathbf{d}(u)$, $\xi_v \leq \varepsilon_n \mathrm{e}^{-a}$, $\xi_v < \varepsilon_{n+1}$. So we get

$$\sum_{u\in\mathcal{U}_{\varepsilon_{n+1}}} \xi_u f\left(\frac{\xi_u}{\varepsilon_{n+1}}\right) = \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u f((n+1)\xi_u) + \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v\in\boldsymbol{d}(u)} \xi_v f((n+1)\xi_v). \tag{5.5}$$

Thus we have, for $n \geq n_1$,

$$\begin{aligned}\left| \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v\in\boldsymbol{d}(u)} \xi_v f((n+1)\xi_v) - \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \xi_u f(n\xi_u) \right| \\ \leq |\gamma_{T_{n+1}}(f) - \gamma_{T_n}(f)| + \left| \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u f((n+1)\xi_u) - \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u f(n\xi_u) \right|. \end{aligned} \tag{5.6}$$

If we take $f = \mathrm{Id}$, the terms in the equation above can be bounded:

$$\begin{aligned} &\left| \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v\in\boldsymbol{d}(u)} \xi_v f((n+1)\xi_v) - \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \xi_u f(n\xi_u) \right| \\ &\geq \left| \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \left(\xi_u f(n\xi_u) - \sum_{v\in\boldsymbol{d}(u)} \xi_v f(n\xi_v) \right) \right| - \left| \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v\in\boldsymbol{d}(u)} (\xi_v f(n\xi_v) - \xi_v f((n+1)\xi_v)) \right| \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by Assumption 2.3)} \\ &\geq \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \left(\xi_u f(n\xi_u) - \sum_{v\in\boldsymbol{d}(u)} \xi_v f(n\xi_u) e^{-a} \right) - \left| \sum_{u\in\mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v\in\boldsymbol{d}(u)} (\xi_v f(n\xi_v) - \xi_v f((n+1)\xi_v)) \right| \end{aligned}$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$\geq \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}} \xi_u (1 - \mathrm{e}^{-a}) \frac{n}{n+1} - \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v \in \boldsymbol{d}(u)} \xi_v \frac{1}{n+1}, \tag{5.7}$$

(by (5.4))

$$|\gamma_{T_{n+1}}(f) - \gamma_{T_n}(f)| + \left| \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u f((n+1)\xi_u) - \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u f(n\xi_u) \right|$$

$$\leq |\gamma_{T_{n+1}}(f) - \gamma_{T_n}(f)| + \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u \frac{1}{n}. \tag{5.8}$$

Let $\iota > 0$. We fix ω in Ω. By (5.3), almost surely, there exists n_2 such that, for $n \geq n_2$, $|\gamma_{T_{n+1}}(f) - \gamma_{T_n}(f)| < \iota$. For $n \geq n_1 \vee n_2$, we can then write

$$\sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}} \xi_u \leq \frac{n+1}{n(1-\mathrm{e}^{-a})} \left(\iota + \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}} \sum_{v \in \boldsymbol{d}(u)} \xi_v \frac{1}{n+1} + \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}} \xi_u \frac{1}{n} \right)$$

(by (5.6), (5.7), and (5.8))

$$\leq \frac{n+1}{n(1-\mathrm{e}^{-a})} \left(\iota + \frac{1}{n} \right). \tag{5.9}$$

(by (5.4))

Let $n \geq n_1 \vee n_2$ and $t \in (T_n, T_{n+1})$. We can use the decomposition

$$\mathcal{U}_{\varepsilon_n} = \mathcal{U}_{\varepsilon_n}^{(1)}(t) \sqcup \mathcal{U}_{\varepsilon_n}^{(2)}(t), \quad \text{where } \mathcal{U}_{\varepsilon_n}^{(1)}(t) = \mathcal{U}_{\varepsilon_n} \cap \mathcal{U}_{\mathrm{e}^{-t}} = \mathcal{U}_{\mathrm{e}^{-t}},\ \mathcal{U}_{\varepsilon_n}^{(2)}(t) = \mathcal{U}_{\varepsilon_n} \backslash \mathcal{U}_{\mathrm{e}^{-t}}. \tag{5.10}$$

For u in $\mathcal{U}_{\varepsilon_n} \backslash \mathcal{U}_{\varepsilon_n}^{(1)}(t)$, we set $\boldsymbol{d}(u, t) = \{v \in \mathcal{U}_{\mathrm{e}^{-t}} : u = \mathbf{m}(v)\}$. As $n \geq n_1$, $\mathbf{d}(u, t) = \mathbf{d}(u)$ and we have

$$\sum_{v \in \boldsymbol{d}(u,t)} \xi_v = \xi_u. \tag{5.11}$$

Similar to (5.5), we have

$$\sum_{u \in \mathcal{U}_{e^{-t}}} \xi_u f\left(\mathrm{e}^t \xi_u\right) = \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}(t)} \xi_u f\left(\mathrm{e}^t \xi_u\right) + \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}(t)} \sum_{v \in \boldsymbol{d}(u,t)} \xi_v f\left(\mathrm{e}^t \xi_v\right).$$

We fix f continuous from [0, 1] to $\mathbb{R}$; there exists $n_3 \in \mathbb{N}^*$ such that, for all $x, y \in [0, 1]$, $|x - y| \leq 1/n_3 \Rightarrow |f(x) - f(y)| < \iota$. Suppose that $n \geq n_1 \vee n_2 \vee n_3$. Then, using (5.10) and (5.11), we have, for all $t \in [T_n, T_{n+1}]$,

$$|\gamma_t(f) - \gamma_{T_n}(f)| = \left| \sum_{u \in \mathcal{U}_{e^{-t}}} \xi_u f\left(\mathrm{e}^t \xi_u\right) - \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}(t)} \xi_u f(n\xi_u) - \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}(t)} \xi_u f(n\xi_u) \right|$$

$$= \left| \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}(t)} \xi_u f\left(\mathrm{e}^t \xi_u\right) + \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}(t)} \sum_{v \in \boldsymbol{d}(u,t)} \xi_v f\left(\mathrm{e}^t \xi_v\right) - \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(1)}(t)} \xi_u f(n\xi_u) - \sum_{u \in \mathcal{U}_{\varepsilon_n}^{(2)}(t)} \xi_u f(n\xi_u) \right|$$

$$\leq \left| \sum_{u \in \mathcal{U}^{(1)}_{\varepsilon_n}(t)} \xi_u f\left(\mathrm{e}^t \xi_u\right) - \sum_{u \in \mathcal{U}^{(1)}_{\varepsilon_n}(t)} \xi_u f(n\xi_u) \right| + 2 \sum_{u \in \mathcal{U}^{(2)}_{\varepsilon_n}(t)} \xi_u \|f\|_\infty$$

$$\leq \sum_{u \in \mathcal{U}^{(1)}_{\varepsilon_n}(t)} \xi_u \iota + 2 \sum_{u \in \mathcal{U}^{(2)}_{\varepsilon_n}(t)} \xi_u \|f\|_\infty$$

$$\leq \iota + 2\|f\|_\infty \frac{n+1}{n(1-\mathrm{e}^{-a})}\left(\iota + \frac{1}{n}\right). \tag{5.12}$$

(using (5.9), and since $\mathcal{U}^{(2)}_{\varepsilon_n}(t) \subset \mathcal{U}^{(2)}_{\varepsilon_n}$)

Equations (5.3) and (5.12) prove the desired result. □

The set $\mathcal{B}^0_{\mathrm{sym}}(1)$ is defined in Section 4.1.

Theorem 5.1. (Central limit theorem.) *We remember we have Fact 3.1, and we are under Assumptions 2.1, 2.2, 2.3, and 3.1. Let $q \in \mathbb{N}^*$. For functions $f_1, \ldots, f_q$ that are continuous and in $\mathcal{B}^0_{\mathrm{sym}}(1)$, we have*

$$\varepsilon^{-q/2}(\gamma_T(f_1), \ldots, \gamma_T(f_q)) \underset{T\to+\infty}{\xrightarrow{\text{law}}} \mathcal{N}(0, (K(f_i, f_j))_{1\leq i,j\leq q}) \qquad (\varepsilon = \mathrm{e}^{-T}).$$

(K is given in (5.13).)

Proof. Let $f_1, \ldots, f_q$, $\mathcal{B}^0_{\mathrm{sym}}(1)$, $v_1, \ldots, v_q \in \mathbb{R}$.

First, we develop the following product (remember that for $u \in \mathcal{U}_\varepsilon$, $\xi_u/\varepsilon < 1$ a.s.):

$$\prod_{u \in \mathcal{U}_\varepsilon} \left(1 + \sqrt{\varepsilon}\frac{\xi_u}{\varepsilon}(iv_1 f_1 + \cdots + iv_q f_q)\left(\frac{\xi_u}{\varepsilon}\right)\right)$$

$$= \exp\left(\sum_{u \in \mathcal{U}_\varepsilon} \log\left[1 + \sqrt{\varepsilon}\,\mathrm{Id} \times (iv_1 f_1 + \cdots + iv_q f_q)\left(\frac{\xi_u}{\varepsilon}\right)\right]\right)$$

(for ε small enough)

$$= \exp\left(\sum_{u \in \mathcal{U}_\varepsilon}\sum_{k\geq 1} \frac{(-1)^{k+1}}{k}\varepsilon^{k/2}(\mathrm{Id} \times (iv_1 f_1 + \cdots + iv_q f_q))^k\left(\frac{\xi_u}{\varepsilon}\right)\right)$$

$$= \exp\left(\frac{1}{\sqrt{\varepsilon}}\gamma_T(iv_1 f_1 + \cdots + iv_q f_q) + \frac{1}{2}\gamma_T(\mathrm{Id} \times (v_1 f_1 + \cdots + v_q f_q)^2) + R_\varepsilon\right),$$

where

$$R_\varepsilon = \sum_{k\geq 3}\sum_{u \in \mathcal{U}_\varepsilon} \frac{(-1)^{k+1}}{k}\varepsilon^{k/2-1}\xi_u\left(\frac{\xi_u}{\varepsilon}\right)^{k-1}(iv_1 f_1 + \cdots + iv_q f_q)^k\left(\frac{\xi_u}{\varepsilon}\right)$$

$$= \sum_{k\geq 3} \frac{(-1)^{k+1}}{k}\varepsilon^{k/2-1}\gamma_T((\mathrm{Id}\,)^{k-1}(iv_1 f_1 + \cdots + iv_q f_q)^k),$$

$$|R_\varepsilon| \leq \sum_{k\geq 3} \frac{\varepsilon^{k/2-1}}{k}(|v_1|\|f_1\|_\infty + \cdots + |v_q|\|f_q\|_\infty)^k = O(\sqrt{\varepsilon}).$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

We have, for some constant C (using $x \in \mathbb{R} \Rightarrow |e^{ix}| = 1$),

$$
\begin{aligned}
&\mathbb{E}\bigg(\bigg|\exp\bigg(\frac{1}{\sqrt{\varepsilon}}\gamma_T(iv_1f_1+\cdots+iv_qf_q)+\frac{1}{2}\gamma_T(\mathrm{Id}\times(v_1f_1+\cdots+v_qf_q)^2)+R_\varepsilon\bigg)\\
&\qquad -\exp\bigg(\frac{1}{\sqrt{\varepsilon}}\gamma_T(iv_1f_1+\cdots+iv_qf_q)+\frac{1}{2}\eta(\Phi(\mathrm{Id}\times(v_1f_1+\cdots+v_qf_q)^2))\bigg)\bigg|\bigg)\\
&\leq \mathbb{E}\bigg(C\bigg|\frac{1}{2}\gamma_T(\mathrm{Id}\times(v_1f_1+\cdots+v_qf_q)^2)-\frac{1}{2}\eta(\Phi(\mathrm{Id}\times(v_1f_1+\cdots+v_qf_q)^2))+R_\varepsilon\bigg|\bigg)\\
&\underset{\varepsilon\to 0}{\longrightarrow} 0. \qquad\qquad \text{(by Proposition 5.1)}
\end{aligned}
$$

Second, we develop the same product in a different manner. We have (the order on $\mathcal{U}$ is defined in Section 2.1),

$$
\begin{aligned}
&\prod_{u\in\mathcal{U}_\varepsilon}\bigg(1+\sqrt{\varepsilon}\frac{\xi_u}{\varepsilon}(iv_1f_1+\cdots+iv_qf_q)\Big(\frac{\xi_u}{\varepsilon}\Big)\bigg)\\
&=\sum_{k\geq 0}\varepsilon^{-k/2}i^k\sum_{1\leq j_1,\dots,j_k\leq q}v_{j_1}\cdots v_{j_k}\sum_{\substack{u_1,\dots,u_k\in\mathcal{U}_\varepsilon\\ u_1<\cdots<u_k}}\xi_{u_1}\cdots\xi_{u_k}f_{j_1}\Big(\frac{\xi_{u_1}}{\varepsilon}\Big)\cdots f_{j_k}\Big(\frac{\xi_{u_k}}{\varepsilon}\Big)\\
&\qquad\qquad \text{(a detailed proof can be found in Section D)}\\
&=\sum_{k\geq 0}\varepsilon^{-k/2}i^k\sum_{1\leq j_1,\dots,j_k\leq q}v_{j_1}\cdots v_{j_k}\frac{1}{k!}\gamma_T^{\odot k}(f_{j_1}\otimes\cdots\otimes f_{j_k}).
\end{aligned}
$$

We have, for all k,

$$
\begin{aligned}
&\bigg|\varepsilon^{-k/2}\sum_{1\leq j_1,\dots,j_k\leq q}v_{j_1}\cdots v_{j_k}\frac{1}{k!}\mathbb{E}\big(\gamma_T^{\odot k}(f_{j_1}\otimes\cdots\otimes f_{j_k})\big)\bigg|\\
&\qquad\leq \varepsilon^{-k/2}\times\frac{q^k\sup(|v_1|,\dots,|v_q|)^k\sup(\|f_1\|_\infty,\dots,\|f_q\|_\infty)^k}{k!}.
\end{aligned}
$$

So, by Corollary 4.1, Proposition 4.1, and (5.1), we get

$$
\begin{aligned}
&\mathbb{E}\bigg(\prod_{u\in\mathcal{U}_\varepsilon}\bigg(1+\sqrt{\varepsilon}\frac{\xi_u}{\varepsilon}(iv_1f_1+\cdots+iv_qf_q)\Big(\frac{\xi_u}{\varepsilon}\Big)\bigg)\bigg)\\
&\underset{\varepsilon\to 0}{\longrightarrow}\sum_{\substack{k\geq 0\\ k\text{ even}}}(-1)^{k/2}\sum_{1\leq j_1,\dots,j_k\leq q}\frac{1}{k!}\sum_{I\in I_k}\prod_{\{a,b\}\in I}V(v_{j_a}f_{j_a},v_{j_b}f_{j_b})\\
&\qquad\qquad \text{(a detailed proof can be found in Section C)}
\end{aligned}
$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$= \sum_{\substack{k\geq 0 \\ k \text{ even}}} \frac{(-1)^{k/2}}{2^{k/2}(k/2)!} \sum_{1\leq j_1,\ldots,j_k\leq q} V(v_{j_1}f_{j_1}, v_{j_2}f_{j_2}) \cdots V(v_{j_{k-1}}f_{j_{k-1}}, v_{j_k}f_{j_k})$$

(using (4.23))

$$= \sum_{\substack{k\geq 0 \\ k \text{ even}}} \frac{(-1)^{k/2}}{2^{k/2}(k/2)!} \left(\sum_{1\leq j_1,j_2\leq q} v_{j_1}v_{j_2}V(f_{j_1}, f_{j_2}) \right)^{k/2}$$

$$= \exp\left(-\frac{1}{2} \sum_{1\leq j_1,j_2\leq q} v_{j_1}v_{j_2}V(f_{j_1}, f_{j_2}) \right).$$

In conclusion, we have

$$\mathbb{E}\left(\exp\left(\frac{1}{\sqrt{\varepsilon}} \gamma_T(iv_1f_1 + \cdots + iv_qf_q) \right) \right)$$
$$\underset{\varepsilon\to 0}{\longrightarrow} \exp\left(-\frac{1}{2}\eta(\Phi(\mathrm{Id} \times (v_1f_1 + \cdots + v_qf_q)^2)) - \frac{1}{2} \sum_{1\leq j_1,j_2\leq q} v_{j_1}v_{j_2}V(f_{j_1}, f_{j_2}) \right).$$

So we get the desired result with, for all f, g,

$$K(f, g) = \eta(\Phi(\mathrm{Id} \times fg) + V(f, g)) \tag{5.13}$$

(V is defined in (4.22)). □

Appendix A. Detailed proof of a bound appearing in the proof of Lemma 4.1

Lemma A.1. *We have, for any f appearing in the proof of Lemma 4.1,*

$$\mathbb{E}(\mathbf{1}_{A_u=f(u),\text{ for all } u\in\mathcal{T}_2} \mid \mathcal{L}_2, \mathcal{T}_2, m_2) \leq \prod_{u\in\mathcal{T}_2\backslash\{0\}} \mathrm{e}^{-(\#f(u)-1)(T_u - T_{\boldsymbol{m}(u)})}.$$

Proof. We want to show this by recurrence on the cardinality of $\mathcal{T}_2$.

If $\#\mathcal{T}_2 = 1$, then $\mathcal{T}_2 = \{0\}$ and the claim is true.

Suppose now that $\#\mathcal{T}_2 = k$ and the claim is true up to the cardinality $k-1$. There exists v in $\mathcal{T}_2$ such that (v, i) is not in $\mathcal{T}_2$, for any i in $\mathbb{N}^*$. We set $\mathcal{T}_2{}' = \mathcal{T}_2\backslash\{v\}$, $\mathcal{L}_2{}' = \mathcal{L}_2\backslash\{v\}$, $m_2{}' : u \in \mathcal{T}_2{}' \to (\xi_u, \inf\{i, i \in A_u\})$. We set $f(v) = \{i_1, \ldots, i_p\}$ (with $i_1 < \cdots < i_p$), $f(\boldsymbol{m}(v)) = \{i_1, \ldots, i_p, i_{p+1}, \ldots, i_q\}$ (with $i_{p+1} < \cdots < i_q$). We suppose that $m_2(v) = (\xi_v, i_1)$, because if $m_2(v) = (\xi_v, j)$ with $j \neq i_1$ then $A_v \neq f(v)$ for all ω, and then the left-hand side of the inequality above is zero. We have,

$$\mathbb{E}(\mathbf{1}_{A_u=f(u),\text{ for all } u\in\mathcal{T}_2} \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$
$$= \mathbb{E}(\mathbf{1}_{A_u=f(u),\text{ for all } u\in\mathcal{T}_2{}'}\mathbb{E}(\mathbf{1}_{A_v=f(v)} \mid \mathcal{L}_2, \mathcal{T}_2, m_2, (A_u, u \in \mathcal{T}_2{}')) \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$= \mathbb{E}(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'}\mathbb{E}(\mathbf{1}_{i_1,\ldots,i_p\in A_v}\mathbf{1}_{i_{p+1},\ldots,i_q\notin A_v} \mid \mathcal{L}_2, \mathcal{T}_2, m_2, (A_u, u\in\mathcal{T}_2{}')) \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$

(remember we condition on m_2, so the $\mathbf{1}_{i_1,\ldots,i_p}$ can be replaced by $\mathbf{1}_{i_2,\ldots,i_p}$)

$$= \mathbb{E}(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'}\mathbb{E}(\mathbf{1}_{i_2,\ldots,i_p\in A_v}\mathbf{1}_{i_{p+1},\ldots,i_q\notin A_v} \mid \mathcal{L}_2, \mathcal{T}_2, m_2, (A_u, u\in\mathcal{T}_2{}')) \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$

$$\le \mathbb{E}(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'}\mathbb{E}(\mathbf{1}_{i_2,\ldots,i_p\in A_v} \mid \mathcal{L}_2, \mathcal{T}_2, m_2, (A_u, u\in\mathcal{T}_2{}')) \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$

(because the (Y_j) introduced in Section 2.2 are independent)

$$= \mathbb{E}\left(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'}\prod_{r=2}^{p}\mathbb{E}(\mathbf{1}_{i_r\in A_v} \mid \mathcal{L}_2, \mathcal{T}_2, m_2, (A_u, u\in\mathcal{T}_2{}')) \mid \mathcal{L}_2, \mathcal{T}_2, m_2\right)$$

(because of (2.1); if $v\in\mathcal{L}_2$ then $\prod_{r=2}^{p}\ldots$ is empty and thus $=1$)

$$= \mathbb{E}\left(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'}\prod_{r=2}^{p}\widetilde{\xi}_v \mid \mathcal{L}_2, \mathcal{T}_2, m_2\right) \qquad \text{(by (4.4) and Proposition 2.1)}$$

$$= \mathrm{e}^{-(\#f(v)-1)(T_v - T_{\boldsymbol{m}(v)})}\mathbb{E}(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'} \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$

$$= \mathrm{e}^{-(\#f(v)-1)(T_v - T_{\boldsymbol{m}(v)})}\mathbb{E}(\mathbb{E}(\mathbf{1}_{A_u = f(u),\text{ for all } u\in\mathcal{T}_2{}'} \mid \mathcal{L}_2{}', \mathcal{T}_2{}', m_2{}') \mid \mathcal{L}_2, \mathcal{T}_2, m_2)$$

(by recurrence)

$$\le \prod_{u\in\mathcal{T}_2\setminus\{0\}} \mathrm{e}^{-(\#f(u)-1)(T_u - T_{\boldsymbol{m}(u)})}. \qquad \square$$

Appendix B. Detailed proof of a bound appearing in the proof of Corollary 4.2

Lemma B.1. *Let q be in $\mathbb{N}$. Then $\sum_{k'\in[q]} 1 + (q-k'-1)_+ \le (q+1)^2$.*

Proof. We have

$$\sum_{k'\in[q]} 1 + (q-k'-1)_+ = q + \sum_{k'\in[q-2]} (q-k'-1) = q + \sum_{i=1}^{q-2} i \le \frac{q(q+1)}{2} \le (q+1)^2. \qquad \square$$

Appendix C. Detailed proof of an equality appearing in the proof of Theorem 5.1

Lemma C.1. *Let $q\in\mathbb{N}^*$. Suppose we have q functions $g_1, \ldots, g_q$ in $\mathcal{B}^0_{\mathrm{sym}}(1)$. Then, for all k even (k in $\mathbb{N}$),*

$$\sum_{1\le j_1,\ldots,j_k\le q}\ \sum_{I\in\mathcal{I}_k}\ \prod_{\{a,b\}\in I} V(g_{j_a}, g_{j_b}) = \frac{k!}{2^{k/2}(k/2)!}\sum_{1\le j_1,\ldots,j_k\le q} V(g_{j_1}, g_{j_2})\cdots V(g_{j_{k-1}}, g_{j_k}).$$

Proof. We set

$$\sum_{1\le j_1,\ldots,j_k\le q}\ \sum_{I\in\mathcal{I}_k}\ \prod_{\{a,b\}\in I} V(g_{j_a}, g_{j_b}) = (\mathrm{I}), \qquad \sum_{1\le j_1,\ldots,j_k\le q} V(g_{j_1}, g_{j_2})\cdots V(g_{j_{k-1}}, g_{j_k}) = (\mathrm{II}).$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

Suppose, for some k, we have $i_1, \ldots, i_k \in [q]$, all distinct. There exist N_1 and N_2 such that:

- the term (I) has N_1 terms $V(g_{i_1}, g_{i_2}) \cdots V(g_{i_{k-1}}, g_{i_k})$ (up to permutations; that is, we consider that $V(g_{i_3}, g_{i_4})V(g_{i_2}, g_{i_1}) \cdots V(g_{i_{k-1}}, g_{i_k})$ and $V(g_{i_1}, g_{i_2}) \cdots V(g_{i_{k-1}}, g_{i_k})$ are the same term);
- the term (II) has N_2 terms $V(g_{i_1}, g_{i_2}) \cdots V(g_{i_{k-1}}, g_{i_k})$ (again, up to permutations).

These numbers N_1 and N_2 do not depend on $i_1, \ldots, i_k$. In the case where the indexes $i_1, \ldots, i_k$ are not distinct, we can easily find the number of terms equal to $V(g_{i_1}, g_{i_2}) \cdots V(g_{i_{k-1}}, g_{i_k})$ in terms (I) and (II). For example, if $i_2 = i_1$ and $i_1, i_3, \ldots, i_k$ are distinct, then

- the term (I) has $2N_1$ terms $V(g_{i_1}, g_{i_2}) \cdots V(g_{i_{k-1}}, g_{i_k})$;
- the term (II) has $2N_2$ terms $V(g_{i_1}, g_{i_2}) \cdots V(g_{i_{k-1}}, g_{i_k})$

(we multiply simply by the number of σ in $\mathcal{S}_k$ such that $(i_1, i_2, \ldots, i_k) = (i_{\sigma(1)}, i_{\sigma(2)}, \ldots, i_{\sigma(k)})$). We do not need to know N_1 and N_2, but we need to know N_1/N_2. By taking $V(g, f)$ to be 1 for all g, f, we see that $N_1/N_2 = \#\mathcal{I}_k = k!/(2^{k/2}(k/2)!)$. □

Appendix D. Detailed proof of an equality appearing in the proof of Theorem 5.1

Lemma D.1. *We have $f_1, \ldots, f_q$, $\mathcal{B}^0_{\text{sym}}(1)$, $k \in \mathbb{N}$, and $v_1, \ldots, v_q \in \mathbb{R}$. Then*

$$\sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \gamma_T^{\odot k}(f_{j_1} \otimes \cdots \otimes f_{j_k})$$

$$= k! \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \sum_{\substack{u_1, \ldots, u_k \in \mathcal{U}_\varepsilon \\ u_1 < \cdots < u_k}} \xi_{u_1} \cdots \xi_{u_k} f_{j_1}\left(\frac{\xi_{u_1}}{\varepsilon}\right) \cdots f_{j_k}\left(\frac{\xi_{u_k}}{\varepsilon}\right).$$

Proof. We have

$$\sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \gamma_T^{\odot k}(f_{j_1} \otimes \cdots \otimes f_{j_k})$$

$$= \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \sum_{a:[k] \hookrightarrow \mathcal{U}_\varepsilon} \xi_{a(1)} \cdots \xi_{a(k)} f_{j_1}\left(\frac{\xi_{a(1)}}{\varepsilon}\right) \cdots f_{j_k}\left(\frac{\xi_{a(k)}}{\varepsilon}\right)$$

(for all injections a, there is exactly one $\sigma_a \in \mathcal{S}_k$ such that $a(\sigma_a(1)) < \cdots < a(\sigma_a(k))$)

$$= \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \sum_{a:[k] \hookrightarrow \mathcal{U}_\varepsilon} \xi_{a(\sigma_a(1))} \cdots \xi_{a(\sigma_a(k))} f_{j_{\sigma_a(1)}}\left(\frac{\xi_{a(\sigma_a(1))}}{\varepsilon}\right) \cdots f_{j_{\sigma_a(k)}}\left(\frac{\xi_{a(\sigma_a(k))}}{\varepsilon}\right)$$

(for $\tau \in \mathcal{S}_k$, we set $\mathcal{E}(\tau) = \{a : [k] \hookrightarrow \mathcal{U}_\varepsilon : \sigma_a = \tau\}$)

$$= \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \sum_{\tau \in \mathcal{S}_k} \sum_{a \in \mathcal{E}(\tau)} \xi_{a(\tau(1))} \cdots \xi_{a(\tau(k))} f_{j_{\tau(1)}}\left(\frac{\xi_{a(\tau(1))}}{\varepsilon}\right) \cdots f_{j_{\tau(k)}}\left(\frac{\xi_{a(\tau(k))}}{\varepsilon}\right)$$

$$= \sum_{\tau \in \mathcal{S}_k} \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \sum_{a \in \mathcal{E}(\tau)} \xi_{a(\tau(1))} \cdots \xi_{a(\tau(k))} f_{j_{\tau(1)}}\left(\frac{\xi_{a(\tau(1))}}{\varepsilon}\right) \cdots f_{j_{\tau(k)}}\left(\frac{\xi_{a(\tau(k))}}{\varepsilon}\right)$$

https://doi.org/10.1017/jpr.2022.114 Published online by Cambridge University Press

$$= \sum_{\tau \in \mathcal{S}_k} \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_{\tau(1)}} \cdots v_{j_{\tau(k)}} \sum_{a \in \mathcal{E}(\tau)} \xi_{a(\tau(1))} \cdots \xi_{a(\tau(k))} f_{j_{\tau(1)}}\left(\frac{\xi_{a(\tau(1))}}{\varepsilon}\right) \cdots f_{j_{\tau(k)}}\left(\frac{\xi_{a(\tau(k))}}{\varepsilon}\right)$$

$$= \sum_{\tau \in \mathcal{S}_k} \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_{\tau(1)}} \cdots v_{j_{\tau(k)}} \sum_{\substack{u_1, \ldots, u_k \in \mathcal{U}_\varepsilon \\ u_1 < \cdots < u_k}} \xi_{u_1} \cdots \xi_{u_k} f_{j_{\tau(1)}}\left(\frac{\xi_{u_1}}{\varepsilon}\right) \cdots f_{j_{\tau(k)}}\left(\frac{\xi_{u_k}}{\varepsilon}\right).$$

The application ('$\hookrightarrow$' means that an application is injective)

$$(a : [k] \to [q],\ \tau : [k] \hookrightarrow [k]) \xrightarrow{\Theta} a \circ \tau$$

is such that, for all $b : [k] \to [q]$, $\#\Theta^{-1}(\{b\}) = k!$. So the above quantity is equal to

$$k! \sum_{1 \le j_1, \ldots, j_k \le q} v_{j_1} \cdots v_{j_k} \sum_{\substack{u_1, \ldots, u_k \in \mathcal{U}_\varepsilon \\ u_1 < \cdots < u_k}} \xi_{u_1} \cdots \xi_{u_k} f_{j_1}\left(\frac{\xi_{u_1}}{\varepsilon}\right) \cdots f_{j_k}\left(\frac{\xi_{u_k}}{\varepsilon}\right).$$

□

References

[1] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd edn (Appl. Math. (New York) **51**). Springer, New York.

[2] Bertoin, J. (2002). Self-similar fragmentations. *Ann. Inst. H. Poincaré Prob. Statist.* **38**, 319–340.

[3] Bertoin, J. (2006). *Random Fragmentation and Coagulation Processes* (Cambridge Studies Adv. Math. **102**). Cambridge University Press.

[4] Bertoin, J. and Martínez, S. (2005). Fragmentation energy. *Adv. Appl. Prob.* **37**, 553–570.

[5] Bond, F. C. (1952). The third theory of comminution. *Trans. AIME Mining Eng.* **193**, 484–494.

[6] Charles, R. J. (1957). Energy–size reduction relationships in comminution. *Trans. AIME Mining Eng.* **208**, 80–88.

[7] Dawson, D. A. and Zheng, X. (1991). Law of large numbers and central limit theorem for unbounded jump mean-field models. *Adv. Appl. Math.* **12**, 293–326.

[8] de la Peña, V. H. and Giné, E. (1999). *Decoupling*. Springer, New York.

[9] Del Moral, P., Patras, F. and Rubenthaler, S. (2009). Tree-based functional expansions for Feynman–Kac particle models. *Ann. Appl. Prob.* **19**, 778–825.

[10] Del Moral, P., Patras, F. and Rubenthaler, S. (2011). Convergence of *U*-statistics for interacting particle systems. *J. Theor. Prob.* **24**, 1002–1027.

[11] Del Moral, P., Patras, F. and Rubenthaler, S. (2011). A mean field theory of nonlinear filtering. In *The Oxford Handbook of Nonlinear Filtering*. Oxford University Press, pp. 705–740.

[12] Devoto, D. and Martnez, S. (1998). Truncated Pareto law and oresize distribution of ground rocks. *Math. Geology* **30**, 661–673.

[13] Dynkin, E. B. and Mandelbaum, A. (1983). Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* **11**, 739–745.

[14] Fontbona, J., Krell, N. and Martnez, S. (2010). Energy efficiency of consecutive fragmentation processes. *J. Appl. Prob.* **47**, 543–561.

[15] Harris, S. C., Knobloch, R. and Kyprianou, A. E. (2010). Strong law of large numbers for fragmentation processes. *Ann. Inst. H. Poincaré Prob. Statist.* **46**, 119–134.

[16] Hoffmann, M. and Krell, N. (2011). Statistical analysis of self-similar conservative fragmentation chains. *Bernoulli* **17**, 395–423.

[17] Lee, A. J. (1990). *U-statistics* (Statistics: Textbooks and Monographs **110**). Marcel Dekker Inc., New York.

[18] Meleard, S. (1998). Convergence of the fluctuations for interacting diffusions with jumps associated with Boltzmann equations. *Stoch. Stoch. Rep.* **63**, 195–225.

[19] Perrier, E. M. and Bird, N. R. (2002). Modelling soil fragmentation: The pore solid fractal approach. *Soil and Tillage Res.* **64**, 91–99.

[20] Rubenthaler, S. (2016). Central limit theorem through expansion of the propagation of chaos for Bird and Nanbu systems. *Ann. Fac. Sci. Toulouse Math. (6)* **25**, 829–873.

[21] SGIBNEV, M. S. (2002). Stone's decomposition of the renewal measure via Banach-algebraic techniques. *Proc. Amer. Math. Soc.* **130**, 2425–2430.
[22] TURCOTTE, D. L. (1986). Fractals and fragmentation. *J. Geophys. Res.* **91**, 1921–1926.
[23] UCHIYAMA, K. (1988). Fluctuations in a Markovian system of pairwise interacting particles. *Prob. Theory Relat. Fields* **79**, 289–302.
[24] WALKER, W. H., LEWIS, W. K., MCADAMS, W. H. AND GILLILAND, E. R. (1967). *Principles of Chemical Engineering*. McGraw-Hill, New York.
[25] WEISS, N. L. (1985). *SME Mineral Processing Handbook*. Society of Mining Engineers of the American Institute of Mining, Metallurgical, and Petroleum Engineers, New York.