

SCALING LIMIT OF A LIMIT ORDER BOOK MODEL VIA THE REGENERATIVE CHARACTERIZATION OF LÉVY TREES

PETER LAKNER¹, JOSH REED¹, AND FLORIAN SIMATOS²

ABSTRACT. We consider the following Markovian dynamic on point processes: either the rightmost atom of the current configuration is removed, or a new atom is added at a random distance from the rightmost atom. Interpreting atoms as limit buy orders, this process was introduced by Lakner et al. [20] to model a one-sided limit order book. We consider this model in the regime where the total number of orders converges to a reflected Brownian motion, and complement the results of Lakner et al. [20] by showing that, in the case where the mean displacement at which a new order is added is positive, the measure-valued process describing the whole limit order book converges to a simple functional of this reflected Brownian motion.

The cornerstone of our approach is the regenerative characterization of Lévy trees proved in Weill [27], which provides an elegant and intuitive proof strategy which we unfold. Moreover, the proofs rely on new results of independent interest on branching random walks with a barrier.

CONTENTS

1. Model and main result	1
2. Notation and preliminary remarks	4
3. Coupling with a branching random walk	7
4. Proof of Theorem 1.1	9
5. Discussion	28
Appendix A. Results on $B(\mathcal{T}_1)$	29
References	36

1. MODEL AND MAIN RESULT

Motivation. A limit order book is a financial trading mechanism that keeps track of orders made by traders, and allows them to execute trades in the future. Typically, a trader places an order to buy a security at a certain level p . If the price of the security is larger than p when the order is placed, then the order is kept in the book and may be fulfilled later, as the price of the security fluctuates and falls below p . Similarly, traders may place sell orders, which gives rise to two-sided order books. Because of the importance of limit order books in financial markets, there has been a lot of research on these models, see for instance the survey by Gould et al. [12].

In recent years, different models have been proposed to study limit order books, from stylized models such as in Yudovina [28] or Lakner et al. [20], to more complex ones proposed in, e.g., Cont et al. [7] or Garèche et al. [10]. Such models may be either discrete,

¹LEONARD N. STERN SCHOOL OF BUSINESS, NEW YORK UNIVERSITY

²EINDHOVEN UNIVERSITY OF TECHNOLOGY, DEN DOLECH 2, 5612 AZ EINDHOVEN, THE NETHERLANDS

E-mail addresses: plakner@stern.nyu.edu, jreed@stern.nyu.edu, f.simatos@tue.nl.

Date: December 9, 2013.

and as such closely stick to the discrete nature of limit order books, or continuous, in order to account for the high frequency at which the number of orders evolves. The main result of the present paper bridges the gap between these two viewpoints, and establishes the convergence, in some high frequency regime, of a discrete model to a continuous one. More precisely, we are interested in the one-sided model proposed in Lakner et al. [20] and derive weak convergence results for the measure-valued process describing the entire limit order book. Similar weak convergence results, but at the process level (typically focusing on the processes tracking the bid and ask prices), have recently been considered in various works, see, e.g., [5, 13, 19].

Besides this (to some extent) practical motivation, we are primarily interested in the model of Lakner et al. [20] because of its connection with a result of Weill [27] on Lévy trees. Lévy trees are, informally speaking, the possible continuous scaling limits of Galton Watson trees. By definition, Galton Watson trees are characterized by a regenerative property which asserts that disjoint subtrees are i.i.d., and Weill [27] showed that a similar (continuous) regenerative property characterizes Lévy trees.

Quite surprisingly, we find that the price process in the model of Lakner et al. [20] satisfies a regenerative property very close to the one characterizing Galton Watson trees (see Lemma 1.2 below, and the discussion following it). This unexpected connection thus suggests an elegant way to study the asymptotic behavior of the price process, namely by showing that the limit of the price process satisfies the continuous regenerative property studied by Weill [27]. Together with minor additional arguments, this allows to get the convergence of the whole measure-valued process.

Although this proof strategy is very attractive at a conceptual level, many technical details need to be solved along the way. A key argument, which is used repeatedly in the proofs, is the coupling established in Simatos [26] between the model studied here and a branching random walk. Leveraging this coupling leads us to derive new results of independent interest on branching random walks with a barrier, which are gathered in the Appendix A. The main ingredients of this coupling needed for our purposes are summarized in Section 3.

Model and main result. Let \mathcal{M} be the set of positive measures on $[0, \infty)$, $\mathbf{z} \in \mathcal{M}$ be the zero measure, δ_a be the Dirac mass at $a \geq 0$ and $\mathcal{M}_F \subset \mathcal{M}$ be the set of finite point measures, i.e., measures $\nu \in \mathcal{M}$ with finite support and of the form $\nu = \sum_p \zeta_p \delta_p$ for some integers ζ_p . For a measure $\nu \in \mathcal{M}$ let $\pi(\nu)$ be the supremum of its support:

$$\pi(\nu) = \sup \{y \geq 0 : \nu([y, \infty)) = 0\}$$

with the convention $\pi(\mathbf{z}) = 0$; $\pi(\nu)$ will be called the *price* of the measure ν , and an atom of $\nu \in \mathcal{M}_F$ will be referred to as an *order*.

We use the canonical notation and denote by $(X_t, t \geq 0)$ the canonical \mathcal{M} -valued process. Let \mathbb{P}_χ be the law of the \mathcal{M}_F -valued Markov process started at $\chi \in \mathcal{M}_F$ and with generator ω given by

$$\omega(f)(\nu) = \lambda \mathbb{E} [f(\nu + \delta_{(\pi(\nu)+J)^+}) - f(\nu)] + \lambda [f(\nu - \delta_{\pi(\nu)}) - f(\nu)] \mathbb{1}_{\{\nu \neq \mathbf{z}\}}$$

where $a^+ = \max(0, a)$ for $a \in \mathbb{R}$, and where $\lambda > 0$ and J , a real-valued random variable, are the only two parameters of the model under consideration. In words, under \mathbb{P}_χ , a new order is added at rate λ and is located at a distance distributed like J to the current price (J will sometimes be referred to as the *displacement* of the newly added order). Note however that an order cannot be placed in the negative half-line, and so an order

with displacement J is placed at $(\pi(v) + J)^+$ (this boundary condition will be discussed in Section 5). On the other hand, provided there is a least one order present, an order currently sitting at the price is removed at rate λ .

Let \mathbb{P}_χ^n be the law of $(\vartheta_n(X_{n^2 t}), t \geq 0)$ under \mathbb{P}_χ , where $\vartheta_n : \mathcal{M} \rightarrow \mathcal{M}$ acts on measures as follows:

$$(1.1) \quad \vartheta_n(v)([y, \infty)) = \frac{1}{n} v([ny, \infty)), \quad y \geq 0.$$

Let $\mathcal{M}_F^n = \vartheta_n(\mathcal{M}_F) = \{\vartheta_n(v) : v \in \mathcal{M}_F\}$. In the sequel we will denote by v^n for $v \in \mathcal{M}_F^n$ the only measure in \mathcal{M}_F such that $\vartheta_n(v^n) = v$. Let in the sequel W be a standard Brownian motion reflected at 0 and $\alpha = (2\lambda)^{1/2}$.

Theorem 1.1. *Assume that $\mathbb{E}(J) > 0$ and that $J \in \{-j^*, -j^* + 1, \dots, 0, 1\}$ for some $j^* \in \mathbb{N}$. Then as $n \rightarrow +\infty$, \mathbb{P}_χ^n converges weakly to \mathbb{P}^∞ under which $\pi \circ X$ is equal in distribution to $\alpha \mathbb{E}(J)W$ and X_t for each $t \geq 0$ is absolutely continuous with respect to Lebesgue measure with density $\mathbb{1}_{\{y \leq \pi(X_t)\}} / \mathbb{E}(J)$, i.e.,*

$$(1.2) \quad X_t([0, y]) = \frac{1}{\mathbb{E}(J)} \min(y, \pi(X_t)), \quad t, y \geq 0.$$

Remark. The behavior when $\mathbb{E}(J) < 0$ is completely different and has been treated in Lakner et al. [20] using stochastic calculus arguments, see Section 5 for more details.

Link with Lévy trees. The following lemma is at the heart of our approach to prove Theorem 1.1. Let in the sequel \mathcal{D} be the set of real-valued càdlàg functions and $\zeta(f) = \inf\{t > 0 : f(t) = 0\}$ for $f \in \mathcal{D}$. We call *excursion*, or *excursion away from 0*, a function $f \in \mathcal{D}$ with $\zeta(f) < +\infty$ and $f(\zeta(f) + t) = 0$ for all $t \geq 0$ (note that we only consider excursions with finite length). We denote by \mathcal{E} the set of excursions. For $a \geq 0$ and $g \leq d$ we say that the function $e = (f((g + t) \wedge d) - a, t \geq 0)$ is an *excursion of f above level a* if e remains non-negative and $e \in \mathcal{E}$.

Lemma 1.2. *Under \mathbb{P}_χ , the sequence of successive excursions of $\pi \circ X$ above level a , for any integer a , are i.i.d., with common distribution the first excursion of $\pi \circ X$ away from 0 under \mathbb{P}_{δ_1} .*

Proof. Consider X under \mathbb{P}_v for any $v \in \mathcal{M}_F$ with $\pi(v) \leq a$ that only puts mass on integers. Then when the first excursion (of $\pi \circ X$) above a begins, the price is at a and an order is added at $a + 1$. Thus if g is the left endpoint of the first excursion above a , X_g must be of the form $X_g = X_{g-} + \delta_{a+1}$ with $\pi(X_{g-}) = a$. This excursion lasts as long as at least one order sits at $a + 1$, and if d is the right endpoint of the first excursion above a , then what happens during the time interval $[g, d]$ above a is independent from X_{g-} and is the same as what happens above 0 during the first excursion of $\pi \circ X$ away from 0 under \mathbb{P}_{δ_1} . Moreover, X_d only puts mass on integers and satisfies $\pi(X_d) \leq a$, so that thanks to the strong Markov property we can iterate this argument. The result therefore follows by induction. \square

Remark. For $a \geq 0$ let $R_a : \mathcal{M} \rightarrow \mathcal{M}$ be defined by $R_a(v)([y, \infty)) = v([a + y, \infty))$, and call $(R_a(X_t), g \leq t \leq d)$ an excursion of X above level a if $(\pi(X_t), g \leq t \leq d)$ is an excursion of $\pi \circ X$ above level a . Then the above proof actually shows that the successive excursions above level a of X are i.i.d., with common distribution the first excursion above 0 of X under \mathbb{P}_{δ_1} .

Lemma 1.2 is at the heart of our proof of Theorem 1.1. Indeed, this regenerative property is strongly reminiscent of Galton Watson branching processes. More precisely, consider a stochastic process $H \in \mathcal{E}$ with finite length and continuous sample paths, that starts at 1, increases or decreases with slope ± 1 and only changes direction at integer times.

For integers $a \geq 0$ and $p > 0$ and conditionally on H having p excursions above level a , let $(e_{a,p}^k, k = 1, \dots, p)$ be these p excursions. Then H is the contour function of a Galton Watson tree if and only if for each a and p , the $(e_{a,p}^k, k = 1, \dots, p)$ are i.i.d. with common distribution H . Indeed, H can always be seen as the contour function of some discrete tree. With this interpretation, the successive excursions above a of H code the subtrees rooted at nodes at depth a in the tree. The $(e_{a,p}^k, k = 1, \dots, p)$ being i.i.d. therefore means that the subtrees rooted at a node at depth a in the tree are i.i.d.: this is precisely the definition of a Galton Watson tree.

The difference between this regenerative property and the regenerative property satisfied by $\pi \circ X$ under \mathbb{P}_z and described in Lemma 1.2 is that, *when conditioned to belong to the same excursion away from 0*, consecutive excursions of $\pi \circ X$ above some level are neither independent, nor identically distributed. If for instance we condition some excursion above level a to be followed by another such excursion within the same excursion away from 0, this biases the number of orders put in $\{0, \dots, a\}$ during the first excursion above a . Typically, one may think that more orders are put in $\{0, \dots, a\}$ in order to increase the chance of the next excursion above a to start soon, i.e., before the end of the current excursion away from 0.

However, this bias is weak and will be washed out in the asymptotic regime we are interested in. Thus it is natural to expect that $\pi \circ X$ under \mathbb{P}_z , properly renormalized, will converge to a process satisfying a continuous version of the discrete regenerative property satisfied by the contour function of Galton Watson trees.

Such a regenerative property has been studied in Weill [27], who has showed that this property characterizes the height process of Lévy trees (see for instance Duquesne and Le Gall [8] for a background on this topic). Thus upon showing that this regenerative property passes to the limit, we will have drastically reduced the possible limit points, and it will remain to show that, among the height processes of Lévy trees, the limit that we have is actually a reflected Brownian motion. From there, a simple argument based on local time considerations allows to conclude that Theorem 1.1 holds.

In summary, our proof of Theorem 1.1 will be divided into four main steps: (1) showing tightness of \mathbb{P}_z^n ; (2) showing, based on Lemma 1.2, that for any accumulation point \mathbf{P} , $\pi \circ X$ under \mathbf{P} satisfies the regenerative property studied in Weill [27] (most of the proof is devoted to this point); (3) arguing that among the height processes of Lévy trees, $\pi \circ X$ under \mathbf{P} must actually be a reflected Brownian motion; (4) showing that X_t under \mathbf{P} has density $\mathbb{1}_{\{y \leq \pi(X_t)\}}/E(J)$ with respect to Lebesgue measure.

Acknowledgements. F. Simatos would like N. Broutin for useful discussions about branching random walks that lead to the results of the appendix.

2. NOTATION AND PRELIMINARY REMARKS

2.1. Additional notation and preliminary remarks. We will write in the sequel \mathbb{P}_x , \mathbb{P} , \mathbb{P}_x^n , \mathbb{P}^n for \mathbb{P}_{δ_x} , \mathbb{P}_z , $\mathbb{P}_{\delta_x}^n$ and \mathbb{P}_z^n , respectively, and denote by \Rightarrow weak convergence. We will

also use \mathbb{P} to denote the probability of other generic random variables (such as when we write $\mathbb{E}(J)$). Let $M(v)$ for $v \in \mathcal{M}$ be the mass of v , i.e., $M(v) = v([0, \infty))$. If $\phi : [0, \infty) \rightarrow [0, \infty)$ is continuous, we will denote by $f_\phi : \mathcal{M} \rightarrow [0, \infty)$ the function defined for $v \in \mathcal{M}$ by $f_\phi(v) = \int \phi dv$.

We will need various local time processes at 0. First of all, $\ell_t = \int_0^t \mathbb{1}_{\{\pi(X_u)=0\}} du$ denotes the Lebesgue measure of the time spent by the price process at 0. For discrete processes, i.e., under \mathbb{P}^n , we will also need the following local time processes at 0 of $M \circ X$ and $\pi \circ X$:

$$L_t^{n,M} = n \int_0^t \mathbb{1}_{\{M(X_u)=0\}} du \quad \text{and} \quad L_t^{n,\pi} = n \int_0^t \mathbb{1}_{\{\pi(X_u)=0\}} du, \quad t \geq 0, n \geq 1.$$

For the continuous processes that will arise as the limit of $\pi \circ X$ and $M \circ X$, we consider the operator \mathcal{L} acting on continuous functions $f : [0, \infty) \rightarrow [0, \infty)$ as follows:

$$\mathcal{L}(f)_t = \lim_{\varepsilon \downarrow 0} \left(\frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{f(u) \leq \varepsilon\}} du \right), \quad t \geq 0.$$

We will only consider \mathcal{L} applied at random processes equal in distribution to βW for some $\beta > 0$, in which case this definition makes sense and indeed leads to a local time process at 0. Note that for any $\beta > 0$ and any f for which $\mathcal{L}(f)$ is well-defined, we have $\mathcal{L}(\beta f) = \beta^{-1} \mathcal{L}(f)$. Moreover, according to Tanaka's formula the canonical semimartingale decomposition of W is given by

$$(2.1) \quad W = \frac{1}{2} \mathcal{L}(W) + \bar{W}$$

where \bar{W} is a standard Brownian motion.

In the sequel we will repeatedly use the fact that the process $\pi \circ X$ under \mathbb{P}^n (or \mathbb{P}) is regenerative at 0, in the sense that successive excursions away from 0 are i.i.d.. Note also that the time durations between successive excursions away from 0 are also i.i.d., independent from the excursions, with common distribution the exponential random variable (with parameter $\lambda \mathbb{P}(J=1)n^2$ under \mathbb{P}^n , and $\lambda \mathbb{P}(J=1)$ under \mathbb{P}).

Moreover, jumps of $\pi \circ X$ under \mathbb{P}^n have size $1/n$, and so if $\pi \circ X$ under \mathbb{P}^n converges weakly to some process P , then P is almost surely continuous (see for instance Theorem 13.4 in Billingsley [4]).

Let θ_t and σ_t for $t \geq 0$ be the shift and stopping operators associated to $\pi \circ X$, i.e., $\theta_t = (\pi(X_{t+s}), s \geq 0)$ and $\sigma_t = (\pi(X_{s \wedge t}), s \geq 0)$. Since by the previous remark, accumulation points of $\pi \circ X$ under \mathbb{P}^n are continuous, these operators are continuous in the following sense (see for instance Lambert et al. [22, Lemma 2.3]).

Lemma 2.1 (Continuity of the shift and stopping operators). *Consider some arbitrary times $T^n, T \geq 0$, possibly random. If, for some probability measure \mathbf{P} on \mathcal{M} , $(\pi \circ X, T^n)$ under \mathbb{P}^n converges weakly to $(\pi \circ X, T)$ under \mathbf{P} , then $(\theta_{T^n}, \sigma_{T^n})$ under \mathbb{P}^n converges weakly to (θ_T, σ_T) under \mathbf{P} .*

We will finally need various random times. For t and $\varepsilon \geq 0$ let

$$G_t = \sup\{s \leq t : \pi(X_s) = 0\}, \quad D_t = \inf\{s \geq t : \pi(X_s) = 0\} \quad \text{and} \quad D_{t,\varepsilon} = \inf\{s \geq t : \pi(X_s) \leq \varepsilon\}.$$

For $0 \leq a \leq b$ we also define $T_b = \inf\{s \geq 0 : \pi(X_s) \geq b\}$ and

$$g_{a,b} = \sup\{s \leq T_b : \pi(X_s) = a\}, \quad d_{a,b} = \inf\{s \geq T_b : \pi(X_s) = a\} \quad \text{and} \quad U_{a,b} = d_{a,b} - g_{a,b}.$$

2.2. An aside on the convergence of hitting times. At several places in the proof of Theorem 1.1 it will be crucial to control the convergence of hitting times. For instance, we will need to show in the fourth step of the proof that if $(X, \pi \circ X) \Rightarrow^n (X, \pi \circ X)$, where by $T_n \Rightarrow^n T$ we mean that T_n under \mathbb{P} converges to T under \mathbf{P} for some probability measure \mathbf{P} on \mathcal{M} , then $D_t \Rightarrow^n D_t$ for any $t \geq 0$. Let us explain why, in order to show that $D_t \Rightarrow^n D_t$, it is enough to show that for any $\eta > 0$,

$$(2.2) \quad \limsup_{n \rightarrow +\infty} \mathbb{P}^n (D_t - D_{t,\varepsilon} \geq \eta) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Let us say that $\pi \circ X$ goes across ε if $\inf_{[D_{t,\varepsilon}, D_{t,\varepsilon} + \eta]} \pi \circ X < \varepsilon$ for every $\eta > 0$, and let $\mathcal{G} = \{\varepsilon > 0 : \pi \circ X \text{ goes across } \varepsilon\}$. Then, the following property holds (see for instance Proposition VI.2.11 in Jacod and Shiryaev [14] or Lemma 3.1 in Lambert and Simatos [21]): if $\mathbf{P}(\varepsilon \in \mathcal{G}) = 1$, then $D_{t,\varepsilon} \Rightarrow^n D_{t,\varepsilon}$.

On the other hand, the complement \mathcal{G}^c of \mathcal{G} is precisely the set of discontinuities of the process $(D_{t,\varepsilon}, \varepsilon > 0)$. Since $(D_{t,\varepsilon}, \varepsilon > 0)$ is càglàd, as the left-continuous inverse of the process $(\inf_{[t, t+s]} \pi \circ X, s \geq 0)$, the set $\{\varepsilon > 0 : \mathbf{P}(\varepsilon \in \mathcal{G}^c) > 0\}$ is at most countable, see for instance Billingsley [4, Section 13]. Gathering these two observations, we see that $D_{t,\varepsilon} \Rightarrow^n D_{t,\varepsilon}$ for all $\varepsilon > 0$ outside a countable set. Then, writing for any $\varepsilon, \eta > 0$

$$\mathbb{P}^n (D_t \geq x) = \mathbb{P}^n (D_t \geq x, D_t - D_{t,\varepsilon} \geq \eta) + \mathbb{P}^n (D_t \geq x, D_t - D_{t,\varepsilon} < \eta)$$

gives

$$\mathbb{P}^n (D_t \geq x) \leq \mathbb{P}^n (D_t - D_{t,\varepsilon} \geq \eta) + \mathbb{P}^n (D_{t,\varepsilon} \geq x - \eta).$$

Since $D_{t,\varepsilon} \Rightarrow^n D_{t,\varepsilon}$ for all ε outside a countable set, and since for those ε we have $\mathbb{P}^n (D_{t,\varepsilon} \geq x - \eta) \rightarrow \mathbf{P} (D_{t,\varepsilon} \geq x - \eta)$ for all η 's outside a countable set, we obtain for all $\varepsilon, \eta > 0$ outside a countable set

$$\limsup_{n \rightarrow +\infty} \mathbb{P}^n (D_t \geq x) \leq \limsup_{n \rightarrow +\infty} \mathbb{P}^n (D_t - D_{t,\varepsilon} \geq \eta) + \mathbf{P} (D_{t,\varepsilon} \geq x - \eta).$$

Next, observe that $D_{t,\varepsilon} \rightarrow D_t$ as $\varepsilon \rightarrow 0$, \mathbf{P} -almost surely. Indeed, $D_{t,\varepsilon}$ decreases as $\varepsilon \downarrow 0$, and its limit D' must satisfy $t \leq D' \leq D_t$, since $t \leq D_{t,\varepsilon} \leq D_t$, and also $\pi(X_{D'}) = 0$, since $\pi(X_{D_{t,\varepsilon}}) \leq \varepsilon$ and $\pi \circ X$ is \mathbf{P} -almost surely continuous. Thus letting first $\varepsilon \rightarrow 0$ and then $\eta \rightarrow 0$ in the previous display, we obtain by (2.2)

$$\limsup_{n \rightarrow +\infty} \mathbb{P}^n (D_t \geq x) \leq \mathbf{P} (D_t \geq x)$$

which shows that $D_t \Rightarrow^n D_t$ by the Portmanteau theorem. This reasoning, detailed for D_t and used in the proof of Lemma 4.6, will also be used in Section 4.4 to control the asymptotic behavior of the hitting times T_b , $g_{a,b}$ and $d_{a,b}$.

We will also use the following useful property: if $\pi \circ X$ and D_t converge weakly, then the convergence actually holds jointly. The reasoning goes as follows. If $\pi \circ X$ and D_t under \mathbb{P}^n converge to $\pi \circ X$ and D_t under \mathbf{P} , then $(\pi \circ X, D_t)$ under \mathbb{P}^n is tight (we always consider the product topology). Let (P', D') be any accumulation point.

Since projections are continuous, P' is equal in distribution to $\pi \circ X$ under \mathbf{P} , in particular it is almost surely continuous, and D' is equal in distribution to D_t under \mathbf{P} , in particular it is almost surely $\geq t$. Further, assume using Skorohod's representation theorem that (P^n, D_t^n) is a version of $(\pi \circ X, D_t)$ under \mathbb{P}^n which converges almost surely to (P', D') . Since $P_{D_t^n}^n = 0$ and P' is continuous, we get $P'_{D'} = 0$ and thus, since $D' \geq t$, $\inf\{s \geq t : P'_s = 0\} \leq D'$. Since these two random variables are both equal in distribution to D_t under \mathbf{P} , they must be (almost surely) equal. This shows that (P', D') is equal in distribution to $(\pi \circ X, D_t)$ under \mathbf{P} and thus uniquely identifies accumulation points.

This reasoning applies to all the hitting times considered in this paper, in particular to T_b , $g_{a,b}$ and $d_{a,b}$. Thus, once we will have shown the convergence of $\pi \circ X$ and, say, T_b , then we will typically be in position to use Lemma 2.1 and deduce the convergence of θ_{T_b} and σ_{T_b} .

2.3. Convention. In the sequel we will need to derive numerous upper and lower bounds, where only the asymptotic behavior up to a multiplicative constant matters. It will therefore be convenient to denote by C a strictly positive and finite constant that may change from line to line, and even within the same line, but which is only allowed to depend on λ and the law of J .

3. COUPLING WITH A BRANCHING RANDOM WALK

Let \mathbb{T} be the set of *colored, labelled, rooted and oriented* trees. Trees in \mathbb{T} are endowed with the lexicographic order. Thus in addition to its genealogical structure, each edge of a tree $T \in \mathbb{T}$ has a real-valued label and each node has one of three colors: either white, green or red.

In the sequel we write $v \in T$ to mean that v is a node of T , and we denote by $\emptyset \in T$ the root of T , by $|T|$ its size (the total number of nodes) and by $h(T)$ its height. Nodes inherit labels in the usual way, i.e., the root has some label and the label of a node that is not the root is obtained recursively by adding to the label of its parent the label on the edge between them. If $v \in T$ we write $\psi(v, T)$ for the label of v (in T), $|v|$ for the depth of v (so that $|\emptyset| = 1$ and $h(T) = \sup_{v \in T} |v|$) and $v_k \in T$ for $k = 1, \dots, |v|$ for the node at depth k on the path from the root to v (so that $v_1 = \emptyset$ and $v_{|v|} = v$). Also, $\psi^*(T) = \sup_{v \in T} \psi(v, T)$ is the largest label in T , $\gamma(T)$ is the green node in T with largest label, with $\gamma(T) = \emptyset$ if T has no green node and in case several nodes have the largest label, $\gamma(T)$ is the last one, and $\Gamma(T) \in \mathcal{M}_F$ is the point measure that records the labels of green nodes:

$$\Gamma(T) = \sum_{v \in T: v \text{ is green}} \delta_{\psi(v, T)}.$$

We say that a node $v \in T$ is *killed* if the label of v is $<$ than the label of the root, and if the label of every other node on the path from the root to v has a label \geq to the one of the root. Let $\mathcal{K}(T) \subset T$ be the set of *killed nodes*:

$$\mathcal{K}(T) = \left\{ v \in T : \min_{1 \leq k \leq |v|-1} \psi(v_k, T) \geq \psi(\emptyset, T) \text{ and } \psi(v, T) < \psi(\emptyset, T) \right\}$$

and consider $B(T) \in \mathbb{T}$ the tree obtained from T by removing all the descendants of the killed nodes (but keeping the killed nodes themselves), and $B_+(T)$ the tree obtained from $B(T)$ by applying the map $x \mapsto x^+$ to the label of every node in $B(T)$. Note that since $B(T)$ is a subtree of T , we always have $\psi^*(B(T)) \leq \psi^*(T)$.

Let $\Phi : \mathbb{T} \rightarrow \mathbb{T}$ be the operator acting on a tree $T \in \mathbb{T}$ as follows. If T has no green node then $\Phi(T) = T$. Else, Φ changes the color of one node in T according to the following rule:

- if $\gamma(T)$ has at least one white child, then its first white child becomes green;
- if $\gamma(T)$ has no white child, then $\gamma(T)$ becomes red.

Let Φ_k be the k th iterate of Φ , i.e., Φ_0 is the identity map and $\Phi_{k+1} = \Phi \circ \Phi_k$, and let also $\tau(T) = \inf\{k \geq 0 : \psi^*(\Phi_k(T)) < \psi(\emptyset, T)\}$. We will sometimes refer to the process $(\Phi_k(T), k = 0, \dots, \tau(T))$ as the *exploration of the tree* T .

Consider a tree $T \in \mathbb{T}$ such that all the nodes are white, except for the root which is green. For such a tree, the dynamic of Φ is such that $\tau(T)$ corresponds to the first time at which the nodes of $B(T) \setminus \mathcal{K}(T)$ are red, the nodes of $\mathcal{K}(T)$ are green and the other

nodes are still white. It has taken one iteration of Φ to make the nodes of $\mathcal{K}(T)$ green, and two to make the nodes of $B(T) \setminus \mathcal{K}(T)$ red (first each of them had to be made green), except for the root which was already green to start with. Thus for such a tree we have $\tau(T) = 2|B(T)| - |\mathcal{K}(T)| - 1$.

Let finally \mathcal{T}_x for $x \in \mathbb{R}$ be the following random tree:

- its genealogical structure is a (critical) Galton Watson tree with geometric offspring distribution with parameter $1/2$;
- $\psi(\emptyset, \mathcal{T}_x) = x$ and labels on the edges are i.i.d., independent from the genealogical structure, and with common distribution J ;
- all nodes are white, except for the root which is green.

Because of the last property and the preceding remark, we have

$$(3.1) \quad \tau(\mathcal{T}_x) = 2|B(\mathcal{T}_x)| - |\mathcal{K}(\mathcal{T}_x)| - 1.$$

Note that since $J \leq 1$, we have $\psi^*(\mathcal{T}_1) \leq h(\mathcal{T}_1)$, and in particular $\psi^*(B(\mathcal{T}_1)) \leq h(\mathcal{T}_1)$. The following result is a slight variation of Theorem 2 in Simatos [26], where the same model in discrete-time and without the boundary condition (i.e., an order may be added in the negative half-line) was studied. The intuition behind this coupling is to create a genealogy between orders in the book, a newly added order being declared the child of the order corresponding to the current price, see Section 3.1 in Simatos [26] for more details.

Theorem 3.1. [Theorem 2 in [26]] *Let a be any integer and $g < d$ be the endpoints of the first excursion of $\pi \circ X$ above level a . Then the process $(X_t - X_{g-}, g \leq t \leq d)$ under \mathbb{P} and embedded at jump epochs is equal in distribution to the process $(\Gamma \circ \Phi_k \circ B_+(\mathcal{T}_{a+1}), k = 0, \dots, \tau(\mathcal{T}_{a+1}))$.*

Ambient tree. Thanks to this coupling, we can see any piece of path of X corresponding to an excursion of the price process above some level a as the exploration of some random tree \mathcal{T}_{a+1} : we will sometimes refer to this tree as the *ambient tree*. Note that the ambient tree of an excursion above a , say e , is a subtree of the ambient tree of the excursion above $a - 1$ containing e . Moreover, the remark following Lemma 1.2 implies that the ambient trees corresponding to successive excursions above some given level are i.i.d..

Exploration time. Theorem 3.1 gives, via (3.1), the number of steps needed to explore the ambient tree, say \mathcal{T} . However, we are interested in X in continuous time. Since jumps in X under \mathbb{P} occur at rate 2λ , independent from everything else, the length of the corresponding excursion is given by $\mathcal{S}(\tau(\mathcal{T}))$, where, here and in the sequel, \mathcal{S} is a random walk with step distribution the exponential random variable with parameter 2λ , independent from the ambient tree \mathcal{T} .

More generally, we will need to control the time needed to explore certain regions of \mathcal{T} , which will translate to controlling $\mathcal{S}(\beta)$ for some random times β defined in terms of \mathcal{T} , and thus independent from \mathcal{S} . As it turns out, the random variables β that need be considered have a heavy tail distribution. Since on the other hand jumps of \mathcal{S} are light-tailed, the approximation $\mathbb{P}(\mathcal{S}(\beta) \geq y) \approx \mathbb{P}(\beta \geq 2\lambda y)$ will accurately describe the situation. Let us make this approximation rigorous: for the upper bound, we write

$$\mathbb{P}(\mathcal{S}(\beta) \geq y) \leq \mathbb{P}(\beta \geq \lambda y) + \mathbb{P}(\mathcal{S}(\beta) \geq y, \beta \leq \lambda y) \leq \mathbb{P}(\beta \geq \lambda y) + \mathbb{P}(\mathcal{S}(\lambda y) \geq y).$$

Then, a large deviations bound shows that $\mathbb{P}(\mathcal{S}(y) \geq y) \leq e^{-\bar{\mu}y}$ with $\bar{\mu} = (1 - \log 2)\lambda$. Carrying out a similar reasoning for the lower bound, we get

$$(3.2) \quad \mathbb{P}(\beta \geq 4\lambda y) - e^{-\underline{\mu}y} \leq \mathbb{P}(\mathcal{S}(\beta) \geq y) \leq \mathbb{P}(\beta \geq \lambda y) + e^{-\bar{\mu}y}$$

with $\underline{\mu} = (2 \log 2 - 1)\lambda$.

4. PROOF OF THEOREM 1.1

We decompose the proof of Theorem 1.1 into several steps. The coupling of Theorem 1.1 makes it possible to translate many questions on \mathbb{P}^n to questions on $B(\mathcal{T}_1)$, and in order to keep the focus of the proof on \mathbb{P}^n , we postpone to the Appendix A the proofs of the various results on $B(\mathcal{T}_1)$ which we need along the way.

At a high level, it is useful to keep in mind that, since $\mathbb{E}(J) > 0$, the law of large numbers prevails and the approximation $\psi(v, \mathcal{T}_1) \approx \mathbb{E}(J)|v|$ describes accurately enough (for our purposes) the labels in the tree $B(\mathcal{T}_1)$. In some sense, most of the randomness of $B(\mathcal{T}_1)$ lies in its genealogical structure, and the results of the Appendix A aim at justifying this approximation.

Note that similar results than the ones we need here are known in a more general setting, but for the tree without the barrier, i.e., for \mathcal{T}_1 instead of $B(\mathcal{T}_1)$, see, e.g., Durrett et al. [9] and Kesten [18].

We begin with a preliminary lemma: recall that W is a reflected Brownian motion, that $\alpha = (2\lambda)^{1/2}$ and that $\mathcal{L}(\beta W) = \beta^{-1}\mathcal{L}(W)$ for any $\beta > 0$.

Lemma 4.1. *As $n \rightarrow +\infty$, $(M \circ X, L^{n,M}, L^{n,\pi})$ under \mathbb{P}^n converges weakly to*

$$(4.1) \quad (\alpha W, \mathcal{L}(\alpha W), \mathcal{L}(\alpha \mathbb{E}(J)W)).$$

Moreover,

$$(4.2) \quad \sup \{ \varepsilon^{-1/2} \mathbb{E}_v^n (L_\varepsilon^{n,M}) : n \geq 1, 0 < \varepsilon < 1, v \in \mathcal{M}_F \} < +\infty.$$

Proof. By definition, $M \circ X$ under \mathbb{P} is a critical $M/M/1$ queue with input rate λ , which is well-known to converge under \mathbb{P}^n to αW . Further, $\lambda L^{n,M}$ is the finite variation process that appears in its canonical (semimartingale) decomposition, and standard arguments show that it converges, jointly with $M \circ X$, to the finite variation process that appears in the canonical decomposition of αW , equal to $(\alpha/2)\mathcal{L}(W)$ by (2.1). Dividing by λ we see that $L^{n,M}$ under \mathbb{P}^n converges to $(\alpha/(2\lambda))\mathcal{L}(W) = \mathcal{L}(\alpha W)$. This shows that $(M \circ X, L^{n,M})$ under \mathbb{P}^n converges weakly to $(\alpha W, \mathcal{L}(\alpha W))$.

We now show that $L^{n,\pi}$ under \mathbb{P}^n converges weakly to $(1/\mathbb{E}(J))\mathcal{L}(\alpha W)$ jointly with $M \circ X$ and $L^{n,M}$. Since $L_t^{n,M}/L_t^{n,\pi}$ under \mathbb{P}^n is equal to $\int_0^{n^2 t} \mathbb{1}_{\{M(X_u)=0\}} du / \ell_{n^2 t}$ under \mathbb{P} , it is enough to show that $\int_0^y \mathbb{1}_{\{M(X_u)=0\}} du / \ell_y \rightarrow \mathbb{E}(J)$ as $y \rightarrow +\infty$, \mathbb{P} -almost surely. Indeed, this would imply that $L^{n,\pi}$ under \mathbb{P}^n converges in the sense of finite-dimensional distributions to $(1/\mathbb{E}(J))\mathcal{L}(\alpha W)$ (jointly with $M \circ X$ and $L^{n,M}$), and so, since $L^{n,\pi}$ and $\mathcal{L}(W)$ are continuous and increasing, Theorem VI.2.15 in Jacod and Shiryaev [14] would imply the desired functional convergence result.

Let $Q = M \circ X \circ \ell^{-1}$, where ℓ^{-1} stands for the right-continuous inverse of ℓ . The composition with ℓ^{-1} makes Q evolve only when the price is at 0. Under \mathbb{P} and while the price is at 0, the dynamic of Q is as follows:

- Q increases by one at rate $\lambda \mathbb{P}(J \leq 0)$ (which corresponds to an order with a displacement ≤ 0 being added) and decreases by one at rate λ , provided $Q > 0$ (which corresponds to an order being removed);
- when an order with displacement > 0 is added, which happens at rate $\lambda \mathbb{P}(J = 1)$, the price makes an excursion away from 0. When it comes back to 0, Q resumes evolving and, by the coupling, a random number of orders distributed like $|\mathcal{K}(\mathcal{T}_1)|$ and independent from everything else have been added at 0.

Thus we see that Q under \mathbb{P} is stochastically equivalent to a $G/M/1$ single-server queue, with two independent Poisson flows of arrivals: customers arrive either one by one at rate $\lambda \mathbb{P}(J \leq 0)$, or by batch of size distributed according to $|\mathcal{K}(\mathcal{T}_1)|$ at rate $\lambda \mathbb{P}(J = 1)$. Then, customers have i.i.d. service requirements following an exponential distribution with parameter λ . In particular, the load of this queue is $\mathbb{P}(J \leq 0) + \mathbb{P}(J = 1) \mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|)$ which by (A.4) is equal to $1 - \mathbb{E}(J)$. Since $\mathbb{E}(J) > 0$, Q is positive recurrent and in particular, the long-term average idle time is equal to one minus the load, i.e.,

$$(4.3) \quad \frac{1}{y} \int_0^y \mathbb{1}_{\{Q_u=0\}} du \xrightarrow{y \rightarrow +\infty} \mathbb{E}(J), \quad \mathbb{P} - \text{almost surely.}$$

Fix on the other hand some $y > 0$: then

$$\int_0^y \mathbb{1}_{\{M(X_u)=0\}} du = \int_0^y \mathbb{1}_{\{M(X_u)=0\}} \mathbb{1}_{\{\pi(X_u)=0\}} du = \int_0^y \mathbb{1}_{\{Q(\ell_u)=0\}} d\ell_u = \int_0^{\ell_y} \mathbb{1}_{\{Q_u=0\}} du$$

which, combined with (4.3), proves that $\int_0^y \mathbb{1}_{\{M(X_u)=0\}} du / \ell_y \rightarrow \mathbb{E}(J)$ and achieves the proof of the convergence of $(M \circ X, L^{n,M}, L^{n,\pi})$.

It remains to prove (4.2): since $M \circ X$ spends more time at 0 when started empty (this can be easily seen with a coupling argument), we have $\mathbb{E}_\nu^n(L_\varepsilon^{n,M}) \leq \mathbb{E}^n(L_\varepsilon^{n,M})$ which gives the uniformity in ν . Further, since as mentioned previously $M \circ X - \lambda L^{n,M}$ is a martingale, we have $\mathbb{E}^n(L_\varepsilon^{n,M}) = \lambda^{-1} \mathbb{E}^n(M(X_\varepsilon))$. Since $M \circ X$ is a reflected critical random walk with jump size $\pm 1/n$ and jump rates λn^2 , one easily proves that $\mathbb{E}^n(M(X_\varepsilon)^2) \leq 2\lambda\varepsilon$ which gives the desired result by Cauchy-Schwarz inequality. \square

4.1. First step : tightness of \mathbb{P}^n . To show the tightness of \mathbb{P}^n , it is enough to show that $M \circ X$ under \mathbb{P}^n is tight, and that for each continuous ϕ which is infinitely differentiable with a compact support, $f_\phi \circ X$ under \mathbb{P}^n is tight (recall that $f_\phi(\nu) = \int \phi d\nu$), see for instance Theorem 2.1 in Roelly-Coppoletta [25]. The tightness of $M \circ X$ is a direct consequence of Lemma 4.1, and so it remains to show the tightness of $f_\phi \circ X$. First of all, note that jumps of $f_\phi \circ X$ under \mathbb{P}^n are upper bounded by $\sup|\phi|/n$, and so we only need to control the oscillations of this process (see for instance the Corollary on page 179 in Billingsley [4]). Using standard arguments, we see that the process $f_\phi \circ X$ is a special semimartingale with canonical decomposition

$$\int \phi dX_t = \int \phi dX_0 + \int_0^t \Omega_n(f_\phi)(X_u) du + Z_t,$$

where Ω_n is the generator of \mathbb{P}^n , given for any $\nu \in \mathcal{M}_F^n$ and any function $f : \mathcal{M} \rightarrow \mathbb{R}$ by

$$\begin{aligned} \Omega_n(f)(\nu) &= \lambda n^2 \mathbb{E} [f(\nu + n^{-1} \delta_{(\pi(\nu) + J/n)^+}) - f(\nu)] \\ &\quad + \lambda n^2 [f(\nu - n^{-1} \delta_{\pi(\nu)}) - f(\nu)] \mathbb{1}_{\{M(\nu) > 0\}}, \end{aligned}$$

and Z is a local martingale with increasing process given by $\langle Z \rangle_t = \int_0^t \Omega_n^2(f_\phi)(X_u) du$, with

$$\begin{aligned} \Omega_n^2(f)(v) &= \lambda n^2 \mathbb{E} \left[\left(f(v + n^{-1} \delta_{(\pi(v) + J/n)^+}) - f(v) \right)^2 \right] \\ &\quad + \lambda n^2 \left[f(v - n^{-1} \delta_{\pi(v)}) - f(v) \right]^2 \mathbb{1}_{\{M(v) > 0\}} \end{aligned}$$

(see, e.g., Lemma VIII.3.68 in Jacod and Shiryaev [14]). In particular, we have

$$\begin{aligned} \Omega_n(f_\phi)(v) &= \lambda n \mathbb{E} \left[\phi((\pi(v) + J/n)^+) \right] - \lambda n \phi(\pi(v)) \mathbb{1}_{\{M(v) > 0\}} \\ &= \lambda n \mathbb{E} \left[\phi((\pi(v) + J/n)^+) - \phi(v) \right] + \lambda n \phi(0) \mathbb{1}_{\{M(v) = 0\}} \end{aligned}$$

and

$$\Omega_n^2(f_\phi)(v) = \lambda \mathbb{E} \left[\phi((\pi(v) + J/n)^+)^2 \right] - \lambda \phi(\pi(v))^2 \mathbb{1}_{\{M(v) > 0\}},$$

from which it follows that

$$\left| \Omega_n(f_\phi)(v) \right| \leq j^* \lambda \sup |\phi'| + n \lambda |\phi(0)| \mathbb{1}_{\{M(v) = 0\}} \quad \text{and} \quad \left| \Omega_n^2(f_\phi)(v) \right| \leq 2 \lambda \sup \phi^2.$$

Thus there exists a finite constant C' , that only depends on λ , the law of J and ϕ , such that for any finite stopping time V and any $\varepsilon > 0$ we have

$$\mathbb{E}^n \left(\left| \int_V^{V+\varepsilon} \Omega_n(f_\phi)(X_u) du \right| \right) \leq C' \varepsilon + C' \mathbb{E}^n \left(L_{V+\varepsilon}^{n,M} - L_V^{n,M} \right) \leq C' \varepsilon + C' \varepsilon^{1/2},$$

where the last inequality follows from (4.2) combined with the strong Markov property at time V . Similarly, $\mathbb{E}^n (|\langle Z \rangle_{V+\varepsilon} - \langle Z \rangle_V|) \leq C' \varepsilon$ and these upper bounds imply the tightness of $f_\phi \circ X$ by standard arguments for the tightness of a sequence of semimartingales, see for instance Theorem VI.4.18 in Jacod and Shiryaev [14], or Theorem 2.3 in Roelly-Coppoletta [25].

We now know that \mathbb{P}^n is tight: it remains to identify accumulation points. In the rest of this section we consider \mathbf{P} an arbitrary accumulation point and assume without loss of generality that \mathbb{P}^n converges weakly to \mathbf{P} . We will denote by \Rightarrow^n weak convergence under \mathbb{P}^n to \mathbf{P} , i.e., we will write $Y_n \Rightarrow^n Z$ if Y_n under \mathbb{P}^n converges to Z under \mathbf{P} . For instance, since \mathbb{P}^n converges weakly to \mathbf{P} we have $f_\phi \circ X \Rightarrow^n f_\phi \circ X$ for every continuous function $\phi \geq 0$ with a compact support, see for instance Theorem 16.16 in Kallenberg [16]. Also, as noted in Section 2.1 the process $\pi \circ X$ under \mathbf{P} is almost surely continuous, and since the jumps of $f_\phi \circ X$ under \mathbb{P}^n are upper bounded by $n^{-1} \sup |\phi|$, the same argument shows that the process $f_\phi \circ X$ under \mathbf{P} is also almost surely continuous.

4.2. Second step : joint convergence. We now show that X under \mathbb{P}^n actually converges jointly with its mass, price and local time processes.

Lemma 4.2. *The following joint convergence holds:*

$$(X, M \circ X, \pi \circ X, L^{n,M}, L^{n,\pi}) \Rightarrow^n (X, M \circ X, \pi \circ X, \mathcal{L}(M \circ X), \mathcal{L}(\mathbb{E}(J)M \circ X)).$$

Moreover, it holds \mathbf{P} -almost surely that $M(X_t) \geq \pi(X_t)$ for every $t \geq 0$.

Proof. Lemma 4.1 and the first step imply that the sequence $(X, M \circ X, L^{n,M}, L^{n,\pi})$ under \mathbb{P}^n is tight. Let $(X', M', \mathcal{L}(M'), \mathcal{L}(\mathbb{E}(J)M'))$ be any accumulation point (which is necessarily of this form by Lemma 4.1), and assume in the rest of the proof, using Skorohod's representation theorem, that X^n is a version of X under \mathbb{P}^n (and $L^{n,M}$ and $L^{n,\pi}$ are defined in terms of X^n) such that $(X^n, M \circ X^n, L^{n,M}, L^{n,\pi}) \rightarrow (X', M', \mathcal{L}(M'), \mathcal{L}(\mathbb{E}(J)M'))$

almost surely. Then, in order to prove the joint convergence, we only have to prove that $M' = M \circ X'$ and that $\pi(X_t^n) \rightarrow \pi(X_t')$ for every $t \geq 0$.

We will use the following key observation: under \mathbb{P} and provided that $M(X_t) > 0$, we have $X_t(\{p\}) \geq 1$ for any $p \leq \pi(X_t)$. It follows that, \mathbb{P}^n -almost surely, $\pi(X_t) \leq M(X_t)$ for every $t \geq 0$ and so $\pi(X_t^n) \leq M(X_t^n)$. Note that this implies the desired inequality $\pi(X_t) \leq M(X_t)$ under \mathbf{P} , once we will have proved that $M' = M \circ X$ and that $\pi(X_t^n) \rightarrow \pi(X_t')$.

So fix some $t \geq 0$ and let us first show that $M'_t = M(X'_t)$. Let $K \geq M'_t + 1$ and ϕ be any decreasing, continuous function with $\phi(x) = 1$ for $x \leq K$ and $\phi(x) = 0$ for $x \geq K + 1$, so that $\int \phi dX_t^n \rightarrow \int \phi dX'_t$ as $n \rightarrow +\infty$. On the other hand, since $M(X_t^n) \rightarrow M'_t$ we have $M(X_t^n) \leq M'_t + 1$, and in particular $\pi(X_t^n) \leq K$, for n large enough. Since $\phi(x) = 1$ for $x \leq K$, we have $\int \phi dX_t^n = M(X_t^n)$ for those n , and since the left-hand side of this equality converges to $\int \phi dX'_t$ while the right-hand side converges to M'_t , we obtain $\int \phi dX'_t = M'_t$. Letting $K \rightarrow +\infty$ we obtain $M(X'_t) = M'_t$.

Let us now prove that $\pi(X_t^n) \rightarrow \pi(X'_t)$. Since $\pi(X_t^n) \leq M(X_t^n)$, the sequence $(\pi(X_t^n), n \geq 1)$ is bounded and any accumulation point is upper bounded by $M(X'_t)$. Consider any such accumulation point p , and assume without loss of generality that $\pi(X_t^n) \rightarrow p \leq M(X'_t)$: we have to show that $p = \pi(X'_t)$. Fix some $b > a > p$ and let ϕ be any continuous function with compact support in $[a, b]$: then $\int \phi dX_t^n \rightarrow \int \phi dX'_t$ and since $\pi(X_t^n) \rightarrow p$ and $a > p$, we have $\int \phi dX_t^n = 0$ for n large enough. Thus $\int \phi dX'_t = 0$ which shows that $\pi(X'_t) \leq p$. To show the reverse inequality, consider $a < b < p$ such that $X_t^n([a, b]) \rightarrow X'_t([a, b])$ (this holds for every $a < b < p$ outside a countable set). Then for n large enough we have $a < b < \pi(X_t^n)$ and so a consequence of the key observation made at the beginning of the proof is that $X_t^n([a, b]) \geq b - a$. This shows that $X'_t([a, b]) \geq b - a$ and so $\pi(X'_t) \geq a$. Letting $a \uparrow p$ achieves the proof. \square

Since under \mathbb{P}^n , orders are added at a distance at most j^*/n from the current price, it follows readily from the previous result that X under \mathbf{P} only evolves locally around its price, in the sense that if $y \geq 0$ and $0 \leq g \leq d$ are such that $\pi(X_t) > y$ for $g \leq t \leq d$, then for $t \in [g, d]$ the measures X_t and X_g restricted to $[0, y]$ are equal. Actually we will only need the following weaker property.

Corollary 4.3. *The following property holds \mathbf{P} -almost surely. Let $t, y \geq 0$ such that $\pi(X_t) > y$, and let g be the left endpoint of the excursion of $\pi \circ X$ above y straddling t . Then $X_t([0, y]) = X_g([0, y])$.*

4.3. Third step : \mathbf{P} is regenerative at 0. So far, we have used the fact that $\pi \circ X$ under \mathbb{P}^n was regenerative at 0, in the natural sense that successive excursions away from 0 are i.i.d.. Under \mathbf{P} there is no first excursion away from 0 and so we need a more general notion of regeneration (we also don't know, at this point, that $\pi \circ X$ under \mathbf{P} is a Markov process). The goal of this step is to show that $\pi \circ X$ under \mathbf{P} is regenerative at 0 in the following sense (recall that θ_t and σ_t are the shift and stopping operators associated to $\pi \circ X$, see Section 2.1),:

- i) the zero set of $\pi \circ X$ has zero Lebesgue measure under \mathbf{P} ;
- ii) $\mathbf{P}(D'_0 = 0) = 1$, where $D'_0 = \inf\{t > 0 : \pi(X_t) = 0\}$;
- iii) for every $t \geq 0$ and every continuous, bounded functions f and g on \mathcal{E} :

$$(4.4) \quad \mathbf{E}[f(\sigma_{D_t})g(\theta_{D_t})] = \mathbf{E}[f(\sigma_{D_t})] \mathbf{E}[g(\pi \circ X)];$$

Note for (4.4) that D_t is \mathbf{P} -almost surely finite, since $D_t \leq \inf\{s \geq t : M(X_s) = 0\}$ by Lemma 4.2, and this upper bound is \mathbf{P} -almost surely finite by Lemma 4.1. It can be checked, following the proof of Theorem 22.11 in Kallenberg [16], that if $\pi \circ X$ under \mathbf{P}

satisfies the three properties i)–iii) above, then $\pi \circ X$ admits an excursion measure away from 0, denoted by \mathcal{N} , by which we mean, in accordance with the literature, that:

- (1) there is a continuous, nondecreasing process L increasing only on the zero set of $\pi \circ X$ (the local time);
- (2) the right-inverse L^{-1} of L is a subordinator and the excursion process sending t to the corresponding excursion if L^{-1} jumps at time t (and to a cemetery point otherwise) is a Poisson point process with intensity measure $\text{dmd}\mathcal{N}$ (m stands for the Lebesgue measure).

The rest of this step is devoted to showing that \mathbf{P} satisfies the properties i)–iii) above. We begin with a preliminary lemma.

Lemma 4.4. *For any $n \geq 1$ and $t \geq 0$, $\mathbb{E}^n(L_t^{n,\pi}) \leq Ct^{1/2}$.*

Proof. We have

$$\mathbb{E}^n(L_t^{n,\pi}) = n\mathbb{E}^n\left(\int_0^t \mathbb{1}_{\{\pi(X_u)=0\}} \text{d}u\right) = n\mathbb{E}\left(\int_0^t \mathbb{1}_{\{\pi(X_{n^2u})=0\}} \text{d}u\right) = \frac{1}{n}\mathbb{E}\left(\int_0^{n^2t} \mathbb{1}_{\{\pi(X_u)=0\}} \text{d}u\right).$$

Let $D^0 = 0$ and D^k for $k \geq 1$ be the endpoint of the k th excursion away from 0 of $\pi \circ X$, and let $K(y) = \sum_{k \geq 1} \mathbb{1}_{\{D^k \leq y\}}$ be the number of excursions finishing before y . In particular,

$$\int_0^{n^2t} \mathbb{1}_{\{\pi(X_u)=0\}} \text{d}u \leq \sum_{k=1}^{K(n^2t)+1} \int_{D^{k-1}}^{D^k} \mathbb{1}_{\{\pi(X_u)=0\}} \text{d}u.$$

The coupling shows that for each $k \geq 1$ we can write $D^k - D^{k-1} = E^k + V^k$, where E^k and V^k are independent and, under \mathbb{P} :

- $E^k = \int_{D^{k-1}}^{D^k} \mathbb{1}_{\{\pi(X_u)=0\}} \text{d}u$ is the time that the price process stays at 0 before the k th excursion starts. In particular, it follows an exponential distribution with parameter $\lambda\mathbb{P}(J=1)$;
- V^k is the time taken to explore the ambient tree \mathcal{T}^k , distributed according to \mathcal{T}_1 , corresponding to the k th excursion. According to the discussion at the end of Section 3, we can write $V^k = \mathcal{S}^k(\tau(\mathcal{T}^k))$ where \mathcal{S}^k is a random walk independent from \mathcal{T}^k and with step distribution an exponential random variable with parameter 2λ .

Note furthermore that, since $\pi \circ X$ under \mathbb{P} is regenerative at 0, the random variables $((E^k, \mathcal{S}^k, \mathcal{T}^k), k \geq 1)$ are i.i.d.. With this decomposition, we have

$$K(y) = \sum_{k \geq 1} \mathbb{1}_{\{D^k \leq y\}} = \sum_{k \geq 1} \mathbb{1}_{\{E^1 + \dots + E^k + V^1 + \dots + V^k \leq y\}} \leq \bar{K}(y)$$

with $\bar{K}(y) = \sum_{k \geq 1} \mathbb{1}_{\{V^1 + \dots + V^k \leq y\}}$, and so

$$\mathbb{E}^n(L_t^{n,\pi}) \leq \frac{1}{n}\mathbb{E}\left(\sum_{k=1}^{\bar{K}(n^2t)+1} E^k\right) = \frac{1}{\lambda\mathbb{P}(J=1)n}\mathbb{E}(\bar{K}(n^2t) + 1),$$

where the last inequality follows from the independence between $\bar{K}(n^2t)$ and the E^k 's. By definition of $\bar{K}(n^2t)$ we have

$$\mathbb{E}(\bar{K}(n^2t)) = \sum_{k \geq 1} \mathbb{P}(V^1 + \dots + V^k \leq n^2t) \leq \sum_{k \geq 1} \mathbb{P}\left(\max_{1 \leq i \leq k} V^i \leq n^2t\right)$$

and since the V^i 's are i.i.d. with common distribution $\mathcal{S}^1(\tau(\mathcal{T}^1))$, we end up with

$$\mathbb{E}(\tilde{K}(n^2 t)) \leq \sum_{k \geq 1} \{\mathbb{P}(V^1 \leq n^2 t)\}^k \leq \frac{1}{1 - \mathbb{P}(V^1 \leq n^2 t)} = \frac{1}{\mathbb{P}(\mathcal{S}^1(\tau(\mathcal{T}^1)) \geq n^2 t)}.$$

According to (A.7) we have $\mathbb{P}(\tau(\mathcal{T}_1) \geq u) \geq Cu^{-1/2}$ and so the lower bound in (3.2) implies that $\mathbb{P}(\mathcal{S}^1(\tau(\mathcal{T}^1)) \geq u)$ obeys to a similar lower bound, which completes the proof. \square

Lemma 4.5 (Proof of property i). *We have $\mathbf{P}(\forall t \geq 0 : \ell_t = 0) = 1$.*

Proof. We will use the compensation formula for the Poisson point process of excursions away from 0 of $\pi \circ X$ associated to the local time ℓ , see, e.g., Corollary IV.11 in Bertoin [3]. More precisely, under \mathbb{P}^n the first jump of ℓ occurs at rate $\lambda \mathbb{P}(J = 1)n^2$, which uniquely identifies the excursion measure of $\pi \circ X$ associated to ℓ as being equal to $\lambda \mathbb{P}(J = 1)n^2$ times the law of an excursion of $\pi \circ X$ away from 0, see for instance Proposition 0.2 in Bertoin [3]. In particular, for any measurable function $F : \mathcal{E} \rightarrow [0, \infty)$ and any $t \geq 0$ we have

$$\mathbb{E}^n \left(\sum_{0 \leq s \leq t} F(\beta_s) \right) = \lambda \mathbb{P}(J = 1) \times n \mathbb{E}^n(\ell_t) \times n \mathbb{E}_1^n[F(\sigma_{D_0})],$$

where $(\beta_s, s \geq 0)$ is the Poisson point process of excursions of $\pi \circ X$ away from 0 associated to ℓ . In particular, since $n\ell_t = L_t^{n,\pi}$ the previous lemma gives

$$(4.5) \quad \mathbb{E}^n \left(\sum_{0 \leq s \leq t} F(\beta_s) \right) \leq Ct^{1/2} \times n \mathbb{E}_1^n[F(\sigma_{D_0})], \quad n \geq 1, t \geq 0.$$

Let us now prove the result: actually, it is enough to prove that for any $\eta > 0$,

$$(4.6) \quad \limsup_{n \rightarrow +\infty} \mathbb{P}^n \left(\int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du \geq \eta \right) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Indeed, if this holds, then using the convergence of $\int \phi dX_t \Rightarrow^n \int \phi dX_t$ for continuous ϕ with a compact support, this implies

$$\mathbf{P} \left(\int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du \geq \eta \right) \xrightarrow{\varepsilon \rightarrow 0} 0$$

which yields $\mathbf{P}(\ell_t = 0) = 1$ for each fixed $t \geq 0$, and thus $\mathbf{P}(\forall t \geq 0 : \ell_t = 0) = 1$ by continuity of ℓ . So let us show (4.6): using Markov inequality, writing $\int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du = \ell_t + \int_0^t \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon\}} du$ and using $t \leq D_t$, we obtain

$$\mathbb{P}^n \left(\int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du \geq \eta \right) \leq \frac{1}{\eta} \mathbb{E}^n(\ell_t) + \frac{1}{\eta} \mathbb{E}^n \left(\int_0^{D_t} \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon\}} du \right).$$

Since $\ell_t = L_t^{n,\pi}/n$, the first term of the above upper bound is upper bounded by $Ct^{1/2}/(\eta n)$ by Lemma 4.4. As for the second term, using (4.5) with $F(\varepsilon) = \int_0^{\varepsilon} \mathbb{1}_{\{\varepsilon_u \leq \varepsilon\}} du$ for $\varepsilon \in \mathcal{E}$, we obtain

$$\mathbb{E}^n \left(\int_0^{D_t} \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon\}} du \right) \leq Ct^{1/2} n \mathbb{E}_1^n \left(\int_0^{D_0} \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du \right) = \frac{Ct^{1/2}}{n} \mathbb{E}_1 \left(\int_0^{D_0} \mathbb{1}_{\{\pi(X_u) \leq \varepsilon n\}} du \right).$$

In terms of the exploration of the ambient tree (equal in distribution to \mathcal{T}_1), the integral under the expectation corresponds to the time spent when the largest label of a

green node was $\leq \varepsilon n$. Since transitions occur at rate 2λ independently from everything else, we therefore have

$$\mathbb{E}_1 \left(\int_0^{D_0} \mathbb{1}_{\{\pi(X_u) \leq \varepsilon n\}} \mathrm{d}u \right) = \frac{1}{2\lambda} \mathbb{E} \left(\sum_{k=0}^{\tau(\mathcal{T}_1)-1} \mathbb{1}_{\{\psi(\gamma(\Phi_k(\mathcal{T}_1)), \mathcal{T}_1) \leq \varepsilon n\}} \right).$$

Further, we can write

$$\sum_{k=0}^{\tau(\mathcal{T}_1)-1} \mathbb{1}_{\{\psi(\gamma(\Phi_k(\mathcal{T}_1)), \mathcal{T}_1) \leq \varepsilon n\}} = \sum_{p=1}^{\varepsilon n} \sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}} \sum_{k \geq 0} \mathbb{1}_{\{\gamma(\Phi_k(\mathcal{T}_1)) = v\}}.$$

The sum $\sum_{k \geq 0} \mathbb{1}_{\{\gamma(\Phi_k(\mathcal{T}_1)) = v\}}$ counts the number of times the node v has been the price: this is actually equal to $1 + \mathcal{C}(v)$, with $\mathcal{C}(v)$ the number of children of v in \mathcal{T}_1 . The one accounts for the first time v becomes the price, and the additional $\mathcal{C}(v)$ accounts for the fact that each child of v makes v stay the price one more unit of time (either immediately, if the child has a smaller label, or later on if the child has a larger label). Thus

$$\sum_{k=0}^{\tau(\mathcal{T}_1)-1} \mathbb{1}_{\{\psi(\gamma(\Phi_k(\mathcal{T}_1)), \mathcal{T}_1) \leq \varepsilon n\}} = \sum_{p=1}^{\varepsilon n} \sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}} (1 + \mathcal{C}(v)).$$

Now this sum counts twice all nodes in $B(\mathcal{T}_1) \setminus \mathcal{K}(\mathcal{T}_1)$ with label in $\{1, \dots, \varepsilon n\}$; it also counts once the nodes in $\mathcal{K}(\mathcal{T}_1)$ as well as the nodes with label $\varepsilon n + 1$ whose parent has label εn . In particular,

$$(4.7) \quad \sum_{k=0}^{\tau(\mathcal{T}_1)-1} \mathbb{1}_{\{\psi(\gamma(\Phi_k(\mathcal{T}_1)), \mathcal{T}_1) \leq \varepsilon n\}} \leq 2 \sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq \varepsilon n + 1\}}$$

and so taking the mean and using (A.8) finally gives $\mathbb{E}_1 \left(\int_0^{D_0} \mathbb{1}_{\{\pi(X_u) \leq \varepsilon n\}} \mathrm{d}u \right) \leq C\varepsilon n$. Gathering the previous inequalities, we end up with

$$\mathbb{P}^n \left(\int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} \mathrm{d}u \geq \eta \right) \leq \frac{Ct^{1/2}}{\eta n} + \frac{C\varepsilon t^{1/2}}{\eta}$$

from which (4.6) follows by letting first $n \rightarrow +\infty$ and then $\varepsilon \rightarrow 0$. \square

Lemma 4.6 (Proof of properties ii) and iii)). *We have $\mathbf{P}(D'_0 = 0) = 1$ and for any $t \geq 0$ and f, g bounded, continuous functions on \mathcal{E} , the relation (4.4) holds.*

Proof. We first prove the property iii). First, assume that for every $\eta > 0$ it holds that

$$(4.8) \quad \limsup_{n \rightarrow +\infty} \mathbb{P}^n(D_t - D_{t,\varepsilon} \geq \eta) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

As explained in Section 2.2, this implies that $(\theta_{D_t}, \sigma_{D_t}) \Rightarrow^n (\theta_{D_t}, \sigma_{D_t})$. On the other hand, since $\pi \circ X$ under \mathbb{P}^n is regenerative at 0, (4.4) holds with \mathbb{E}^n instead of \mathbb{E} and so passing to the limit and using $(\theta_{D_t}, \sigma_{D_t}) \Rightarrow^n (\theta_{D_t}, \sigma_{D_t})$ we obtain the desired result. Thus we only have to prove (4.8), which we prove now.

Let $A_t = t - G_t$ be the age of the excursion straddling t , and $G^u < D^u$ for $u > 0$ be the endpoints of the first excursion of $\pi \circ X$ with length $> u$, say e^u : then Theorem (5.9) in Gettoor [11] shows that for $u > 0$ and $v \in \mathcal{M}_F$ and conditionally on $\{A_t = u, X_{G_t}^n = v\}$ (recall the definition of X_t^n before Theorem 1.1), the excursion of $\pi \circ X$ straddling t is equal in distribution to e^u conditionally on $\{X_{G^u}^n = v\}$: in particular,

$$\mathbb{P}^n(D_t - D_{t,\varepsilon} \geq \eta) = \int \mathbb{P}^n(A_t \in \mathrm{d}u, X_{G_t}^n \in \mathrm{d}v) \mathbb{P}_v^n(D_u - D_{u,\varepsilon} \geq \eta \mid D_0 > u).$$

Further, under \mathbb{P}^n , $X_{G_t}^n$ is almost surely of the form $v = \zeta\delta_0 + \delta_1$ for some $\zeta \geq 0$. For such an initial condition, the number ζ of orders sitting at 0 does not influence the first excursion, which is distributed like the first excursion under \mathbb{P}_1 : thus

$$\mathbb{P}^n(D_t - D_{t,\varepsilon} \geq \eta) = \int \mathbb{P}^n(A_t \in du) \mathbb{P}_1^n(D_u - D_{u,\varepsilon} \geq \eta \mid D_0 > u)$$

and the goal is now to prove that

$$\limsup_{n \rightarrow +\infty} \sup_{u > 0} \mathbb{P}_1^n(D_u - D_{u,\varepsilon} > \eta \mid D_0 > u) \xrightarrow{\varepsilon \rightarrow 0} 0,$$

which will achieve the proof of (4.8). Rescaling, we obtain

$$\mathbb{P}_1^n(D_u - D_{u,\varepsilon} \geq \eta \mid D_0 > u) = \mathbb{P}_1(D_{un^2} - D_{un^2,\varepsilon n} \geq \eta n^2 \mid D_0 > un^2)$$

and to control this term we consider any $\varepsilon' > 0$ and write

$$(4.9) \quad \mathbb{P}_1^n(D_u - D_{u,\varepsilon} \geq \eta \mid D_0 > u) \leq \mathbb{P}_1(D_{un^2} - D_{un^2,\varepsilon n} \geq \eta n^2, S \leq (\varepsilon + \varepsilon')n \mid D_0 > un^2) \\ + \mathbb{P}_1(S \geq (\varepsilon + \varepsilon')n \mid D_0 > un^2)$$

where $S = \sup \pi \circ X$, where the supremum is taken over $[D_{un^2,\varepsilon n}, D_{un^2}]$.

High-level description. Let us explain in words how we are going to upper bound each term in the right-hand side of (4.9): this reasoning will also be used in the proof of Lemma 4.8. Let $\hat{\mathcal{T}}$ be the ambient tree corresponding to the first excursion of $\pi \circ X$ away from 0, so that D_0 is the sum of $\tau(\hat{\mathcal{T}})$ i.i.d. exponential random variables with parameter 2λ and the conditioning $D_0 > un^2$ therefore amounts, by (3.1), to $B(\hat{\mathcal{T}})$ having a large number of nodes.

When $S \leq (\varepsilon + \varepsilon')n$, then $D_{un^2} - D_{un^2,\varepsilon n}$ is smaller than the time spent exploring *all* the nodes in $\hat{\mathcal{T}}$ with label $\leq (\varepsilon + \varepsilon')n$. We have a good control on the number of such nodes (they are of the order of $(\varepsilon + \varepsilon')^2 n^2$) which thus translates into a good control on $D_{un^2} - D_{un^2,\varepsilon n}$ in this event.

On the other hand, to control the probability of S being large, i.e., $S > (\varepsilon + \varepsilon')n$, we observe that S is equal to the largest supremum of the excursions above εn that start between times $D_{un^2,\varepsilon n}$ and D_0 . By Lemma 1.2 these excursions are i.i.d. with common distribution the exploration of a tree distributed like \mathcal{T}_1 . In particular, we can control their supremum (which is equal in distribution to $\psi^*(\mathcal{T}_1)$), and to control their number, we use a crude upper bound by saying that there cannot be more excursions above level εn than there are nodes in $\hat{\mathcal{T}}$ with label $= \varepsilon n$. Again, we have a good control on these two quantities which, combined, will give us a sufficiently good control on the probability of S being large.

Let us now make these arguments rigorous. As just explained, $D_{un^2} - D_{un^2,\varepsilon n}$ is, in the event $\{S \leq (\varepsilon + \varepsilon')n, D_0 > un^2\}$, smaller than the time spent exploring the nodes of the ambient tree that have a label $\leq (\varepsilon + \varepsilon')n$. This means that if

$$N^{\leq} = \sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq (\varepsilon + \varepsilon')n\}}$$

is the number of such nodes, then

$$\mathbb{P}_1(D_{un^2} - D_{un^2,\varepsilon n} \geq \eta n^2, S \leq (\varepsilon + \varepsilon')n \mid D_0 > un^2) \leq \mathbb{P}(\mathcal{S}(2N^{\leq}) \geq \eta n^2 \mid \tau(\mathcal{T}_1) > un^2)$$

(the factor 2 in $2N^\leq$ comes from the same reason as the 2 in the right-hand side of (4.7)). Invoking (3.2), we get

$$\mathbb{P}_1(D_{un^2} - D_{un^2, \varepsilon n} \geq \eta n^2, S \leq (\varepsilon + \varepsilon')n \mid D_0 > un^2) \leq \mathbb{P}(N^\leq \geq \lambda \eta n^2 / 2 \mid \tau(\mathcal{T}_1) > un^2) + e^{-\bar{\mu} \eta n^2}$$

and so (A.10) finally gives

$$\mathbb{P}_1(D_{un^2} - D_{un^2, \varepsilon n} \geq \eta n^2, S \leq (\varepsilon + \varepsilon')n \mid D_0 > un^2) \leq \frac{C(\varepsilon + \varepsilon')^2}{\eta} + e^{-\bar{\mu} \eta n^2}.$$

We now control the second term in the right-hand side of (4.9). Let e^k be the k th excursion of $\pi \circ X$ above level $\varepsilon n - 1$ to start after time un^2 , and let N^\equiv be the number of excursions above level $\varepsilon n - 1$ that belong to the first excursion of $\pi \circ X$ away from 0: then as explained above, for any $\kappa_0 > 0$ we have

$$\mathbb{P}_1(S \geq (\varepsilon + \varepsilon')n \mid D_0 > un^2) \leq \mathbb{P}_1(N^\equiv \geq \kappa_0 n \mid D_0 > un^2) + \mathbb{P}_1\left(\sup_{1 \leq k \leq \kappa_0 n} \sup e^k \geq \varepsilon' n \mid D_0 > un^2\right).$$

Thanks to the coupling, we have

$$\mathbb{P}_1(N^\equiv \geq \kappa_0 n \mid D_0 > un^2) \leq \mathbb{P}\left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = \varepsilon n\}} \geq \kappa_0 n \mid \tau(\mathcal{T}_1) > un^2\right)$$

and so (A.11) gives $\mathbb{P}_1(N^\equiv \geq \kappa_0 n \mid D_0 > un^2) \leq C\varepsilon/\kappa_0$. On the other hand, since under $\mathbb{P}_1(\cdot \mid D_0 > un^2)$ the $(e^k, k \geq 1)$ are i.i.d., with common distribution the first excursion of $\pi \circ X$ under \mathbb{P}_1 (as a consequence of Lemma 1.2), we have thanks to the union bound

$$\mathbb{P}_1\left(\sup_{1 \leq k \leq \kappa_0 n} \sup e^k \geq \varepsilon' n \mid D_0 > un^2\right) \leq \kappa_0 n \mathbb{P}_1\left(\sup_{[0, D_0]} \pi \circ X \geq \varepsilon' n\right).$$

By the coupling,

$$\mathbb{P}_1\left(\sup_{[0, D_0]} \pi \circ X \geq \varepsilon' n\right) = \mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq \varepsilon' n) \leq \frac{C}{\varepsilon' n}$$

where the last inequality follows from Lemma A.3. Gathering the previous bounds, we see that

$$\mathbb{P}_1^n(D_u - D_{u, \varepsilon} \geq \eta \mid D_0 > u) \leq \frac{C(\varepsilon + \varepsilon')^2}{\eta} + \frac{C\varepsilon}{\kappa_0} + \frac{C\kappa_0}{\varepsilon'} + e^{-\bar{\mu} \eta n^2}.$$

Choosing $\kappa_0 = \varepsilon^{1/2}$ and $\varepsilon' = \varepsilon^{1/4}$, and letting first $n \rightarrow \infty$ and then $\varepsilon \rightarrow 0$ achieves to prove (4.8), and in particular property iii).

We now prove property ii), i.e., for any $\varepsilon > 0$ we must prove that $\mathbf{P}(D'_0 \geq \varepsilon) = 0$. For any $\eta > 0$ we have $\mathbf{P}(D'_0 \geq \varepsilon) \leq \mathbf{P}(D_\eta \geq \varepsilon - \eta)$. Because we have just proved that $D_\eta \Rightarrow^n D_\eta$, we have $\mathbb{P}^n(D_\eta \geq \varepsilon') \rightarrow \mathbf{P}(D_\eta \geq \varepsilon')$ for all ε' outside a countable set. Adapting the previous arguments, it is on the other hand not difficult to see that

$$\limsup_{n \rightarrow +\infty} \mathbb{P}^n(D_\eta \geq \varepsilon') \xrightarrow{\eta \rightarrow 0} 0$$

which concludes the proof of the lemma. \square

4.4. Fourth step : a regenerative property at the excursion level. Let in the sequel \mathcal{N} be an excursion measure of \mathbf{P} , whose existence has been proved in the previous step. Note that \mathcal{N} is only determined up to a multiplicative constant: in this step the value of this multiplicative constant is irrelevant (because we only consider \mathcal{N} upon some conditionings), and it will be fixed at the end of the next step.

The goal of this step is to show that \mathcal{N} satisfies the following regenerative property (R) studied in Weill [27]. In the sequel we use the canonical notation for excursions, and let $\epsilon = (\epsilon_t, t \geq 0)$ denote the canonical excursion and $\xi(a, u)$ for $a, u > 0$ denote the number of excursions of ϵ that reach level a and have height $> u$.

(R) For every $a, u > 0$ and $p \in \mathbb{N}$, under the probability measure $\mathcal{N}(\cdot \mid \sup \epsilon > a)$ and conditionally on the event $\{\xi(a, u) = p\}$, the p excursions of ϵ above level a with height greater than u are independent and distributed according to the probability measure $\mathcal{N}(\cdot \mid \sup \epsilon > u)$.

This property implies that \mathcal{N} is the law of the excursion height process of a spectrally positive Lévy process that does not drift to $+\infty$, see the next step for more details.

The rest of this step is therefore devoted to proving that \mathcal{N} satisfies the regenerative property (R). Fix until the rest of this step $a, u > 0$, $p \in \mathbb{N}$ and $(f_k, k = 1, \dots, p)$ continuous, bounded and non-negative functions on \mathcal{E} . Consider the first excursion of ϵ (or $\pi \circ X$) with exactly p excursions above a with height larger than u and let $(\hat{\epsilon}^k, k = 1, \dots, p)$ be these p excursions: in order to show that \mathcal{N} satisfies (R) we have to show that

$$\mathcal{N} \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \mid \xi(a, u) = p \right) = \prod_{k=1}^p \mathcal{N} (f_k(\epsilon) \mid \sup \epsilon > u).$$

To prove this we will prove that

$$(4.10) \quad \mathbb{E}^n \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right) \xrightarrow{n \rightarrow +\infty} \mathcal{N} \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \mid \xi(a, u) = p \right)$$

while at the same time

$$(4.11) \quad \mathbb{E}^n \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right) \xrightarrow{n \rightarrow +\infty} \prod_{k=1}^p \mathcal{N} (f_k(\epsilon) \mid \sup \epsilon > u).$$

Let in the rest of the proof $\hat{g}^k < \hat{d}^k$ be the endpoints of $\hat{\epsilon}^k$, ϵ^k be the k th excursion of $\pi \circ X$ above a with height $> u$, and $g^k < d^k$ be its endpoints. Note in particular that $(g^1, d^1) = (g_{a, a+u}, d_{a, a+u})$.

4.4.1. Proof of (4.10). Since \mathcal{N} is an excursion measure of $\pi \circ X$ under \mathbf{P} , the probability distribution $\mathcal{N}(\cdot \mid \xi(a, u) = p)$ is the law of the first excursion of $\pi \circ X$ under \mathbf{P} that has exactly p excursions above a with height $> u$, and in particular

$$\mathbb{E} \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right) = \mathcal{N} \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \mid \xi(a, u) = p \right).$$

Thus in order to prove (4.10) it is enough to show that

$$\mathbb{E}^n \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right) \xrightarrow{n \rightarrow +\infty} \mathbb{E} \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right),$$

i.e., that $(\hat{\epsilon}^k, k = 1, \dots, p) \Rightarrow^n (\hat{\epsilon}^k, k = 1, \dots, p)$. In view of Lemma 2.1 it is enough to show that the corresponding endpoints converge, i.e., that we have the convergence

$((\hat{g}^k, \hat{d}^k), k = 1, \dots, p) \Rightarrow ((\hat{g}^k, \hat{d}^k), k = 1, \dots, p)$. We first show in the following two lemmas that T_{a+u} and $(g_{a,a+u}, d_{a,a+u})$ converge, and explain after Lemma 4.8 why this implies the convergence of $((\hat{g}^k, \hat{d}^k), k = 1, \dots, p)$.

Lemma 4.7. *For any $b, \eta > 0$,*

$$\limsup_{n \rightarrow +\infty} \mathbb{P}^n (T_{\bar{b}} - T_{\underline{b}} \geq \eta) \xrightarrow{\bar{b} \downarrow b, \underline{b} \uparrow b} 0.$$

Proof. Consider any $0 < \underline{b}' < \underline{b}$ and let $\varepsilon = \bar{b} - \underline{b}$, $\varepsilon' = \underline{b} - \underline{b}'$ and $S = \sup \pi \circ X - \underline{b}'n$, where the supremum is taken over $[g_{\underline{b}'n, \underline{b}n}, d_{\underline{b}'n, \underline{b}n}]$. Rescaling, we obtain

$$\mathbb{P}^n (T_{\bar{b}} - T_{\underline{b}} > \eta) = \mathbb{P} (T_{\bar{b}n} - T_{\underline{b}n} > \eta n^2, S \geq (\varepsilon + \varepsilon')n) + \mathbb{P} (T_{\bar{b}n} - T_{\underline{b}n} > \eta n^2, S < (\varepsilon + \varepsilon')n).$$

When $S \geq (\varepsilon + \varepsilon')n$, then necessarily $g_{\underline{b}'n, \underline{b}n} \leq T_{\underline{b}n} \leq T_{\bar{b}n} \leq d_{\underline{b}'n, \underline{b}n}$ and so

$$\mathbb{P}^n (T_{\bar{b}} - T_{\underline{b}} > \eta) \leq \mathbb{P} (U_{\underline{b}'n, \underline{b}n} > \eta n^2) + \mathbb{P} (S < (\varepsilon + \varepsilon')n).$$

The coupling implies that $U_{\underline{b}'n, \underline{b}n}$ under \mathbb{P} is equal in distribution to $\mathcal{S}(\tau(\mathcal{T}_1) + 1)$ conditionally on $\psi^*(B(\mathcal{T}_1)) \geq \varepsilon'n$, and also that S under \mathbb{P} is equal in distribution to $\psi^*(B(\mathcal{T}_1))$ conditioned on $\psi^*(B(\mathcal{T}_1)) \geq \varepsilon'n$. Using $\tau(\mathcal{T}_1) + 1 \leq 3|\mathcal{T}_1|$ by (3.1) and using also (3.2), we therefore get

$$\begin{aligned} \mathbb{P}^n (T_{\bar{b}} - T_{\underline{b}} > \eta) &\leq \mathbb{P} (|\mathcal{T}_1| > \lambda \eta n^2 / 3 \mid \psi^*(B(\mathcal{T}_1)) \geq \varepsilon'n) + e^{-\bar{\mu} \eta n^2} \\ &\quad + \mathbb{P} (\psi^*(B(\mathcal{T}_1)) < (\varepsilon + \varepsilon')n \mid \psi^*(B(\mathcal{T}_1)) \geq \varepsilon'n). \end{aligned}$$

In view of (A.12) and (A.13), choosing $\varepsilon' = \varepsilon^{1/2}$ and letting first $n \rightarrow +\infty$ and then $\varepsilon \rightarrow 0$ gives the result. \square

Lemma 4.8. *For any $0 \leq a < b$ and any $\eta > 0$,*

$$(4.12) \quad \limsup_{n \rightarrow +\infty} \mathbb{P}^n (U_{\underline{a}, b} - U_{\bar{a}, b} \geq \eta) \xrightarrow{\bar{a} \downarrow a, \underline{a} \uparrow a} 0.$$

Proof. Since $U_{\underline{a}, b} - U_{\bar{a}, b} = (d_{\underline{a}, b} - d_{\bar{a}, b}) + (g_{\bar{a}, b} - g_{\underline{a}, b})$, we only have to prove that

$$\lim_{\bar{a} \downarrow a, \underline{a} \uparrow a} \limsup_{n \rightarrow +\infty} \mathbb{P}^n (d_{\underline{a}, b} - d_{\bar{a}, b} \geq \eta) = \lim_{\bar{a} \downarrow a, \underline{a} \uparrow a} \limsup_{n \rightarrow +\infty} \mathbb{P}^n (g_{\bar{a}, b} - g_{\underline{a}, b} \geq \eta) = 0.$$

The proofs for d and g are very similar to one another, and also very similar to the proof of Lemma 4.6. Let us first sketch the proof for d . First of all, we are interested in the excursion straddling T_{bn} and above $\underline{a}n$, so the ambient tree, say $\hat{\mathcal{T}}$, is distributed like \mathcal{T}_1 conditioned on $\psi^*(B(\mathcal{T}_1)) > (b - \underline{a})n$. Let $\varepsilon = \bar{a} - \underline{a}$, consider any $\varepsilon' > 0$ and define

$$N^{\leq} = \sum_{v \in \hat{\mathcal{T}}} \mathbb{1}_{\{\psi(v, \hat{\mathcal{T}}) \leq (\varepsilon + \varepsilon')n\}} \quad \text{and} \quad N^{\varepsilon} = \sum_{v \in \hat{\mathcal{T}}} \mathbb{1}_{\{\psi(v, \hat{\mathcal{T}}) = \varepsilon n\}}$$

Rescaling and introducing $S = \sup \pi \circ X - \underline{a}n$, where the supremum is taken over $[d_{\bar{a}n, bn}, d_{\underline{a}n, bn}]$, we obtain

$$(4.13) \quad \mathbb{P}^n (d_{\underline{a}, b} - d_{\bar{a}, b} \geq \eta) \leq \mathbb{P} (d_{\underline{a}n, bn} - d_{\bar{a}n, bn} \geq \eta n^2, S \leq (\varepsilon + \varepsilon')n) + \mathbb{P} (S > (\varepsilon + \varepsilon')n).$$

To control the right-hand side of the above upper bound we use a similar reasoning as in the proof of Lemma 4.6 (see the *High-level description* there). To control the first term of the above right-hand side, we observe that in the event $S \leq (\varepsilon + \varepsilon')n$, $d_{\underline{a}n, bn} - d_{\bar{a}n, bn}$

is upper bounded by the time spent exploring nodes with label $\leq (\varepsilon + \varepsilon')n$ in $\hat{\mathcal{F}}$, which leads to the bound

$$\mathbb{P}(d_{\underline{a}n, bn} - d_{\bar{a}n, bn} \geq \eta n^2, S \leq (\varepsilon + \varepsilon')n) \leq \mathbb{P}(N^{\leq} \geq \lambda \eta n^2 / 2 \mid \psi^*(B(\mathcal{T}_1)) > (b - \underline{a})n) + e^{-\bar{\mu} \eta n^2}.$$

To control the second term of the right-hand side of (4.13), we observe that (1) S is equal to the largest supremum of the excursions above $\bar{a}n$ that start after $d_{\bar{a}n, bn}$ and end before $d_{0, bn}$; (2) the number of such excursions is smaller than the number of nodes with label $= \varepsilon n$ in $\hat{\mathcal{F}}$; and (3) the excursions above $\bar{a}n$ and starting after time $d_{\bar{a}n, bn}$ are i.i.d. with common distribution the first excursion of $\pi \circ X$ under \mathbb{P}_1^n . This leads to the bound

$$\mathbb{P}(S > (\varepsilon + \varepsilon')n) \leq \mathbb{P}(N^{\leq} \geq \kappa_0 n \mid \psi^*(B(\mathcal{T}_1)) > (b - \underline{a})n) + \frac{C\kappa_0}{\varepsilon'}.$$

In view of (A.14) and (A.15) we get the desired result for d . For g we derive the exact same upper bound by considering $S = \sup \pi \circ X - \underline{a}n$, where the supremum is now taken over $[g_{\underline{a}n, bn}, g_{\bar{a}n, bn}]$. There is one additional minor difference, namely that excursions above $\bar{a}n$ that end before $g_{\bar{a}n, bn}$ are i.i.d., but with distribution the first excursion of $\pi \circ X$ above $\bar{a}n$ conditioned on having its height $< bn$. Since the probability of this event goes to one, this additional conditioning has no influence on the result. \square

We now explain why the two previous lemmas imply the convergence of the vector $((\hat{g}^k, \hat{d}^k), k = 1, \dots, p)$ (by which we mean that $((\hat{g}^k, \hat{d}^k), k = 1, \dots, p) \Rightarrow^n ((\hat{g}^k, \hat{d}^k), k = 1, \dots, p)$). First of all, the discussion in Section 2.2 shows that $\pi \circ X$ shifted at time T_{a+u} converges. Thus by Lemma 4.8, $d_{a, a+u}$, which is the hitting time of $(0, a+u]$ by the process $\pi \circ X$ shifted at time T_{a+u} , converges. Moreover, the arguments in the proof of Lemma 4.8 go through for $a = 0$, which shows that $D_{d_{a, a+u}}$ converges. Since $g_{a, a+u}$ is the hitting time of $(0, a+u]$ by the process $\pi \circ X$ shifted at time T_{a+u} and run backward in time, and the mapping that to a function associates the same function run backward in time is continuous, we obtain for the same reasons the convergence of $g_{a, a+u}$.

Recall that e^k is the k th excursion of $\pi \circ X$ above a with height $> u$, and $g^k < d^k$ are its endpoints. Let also $T^k = \inf\{t \geq g^k : \pi(X_t) \geq a+u\}$. The idea is now to iterate the above arguments by looking at the process $\pi \circ X$ shifted at time d^k . Let us look at $k = 1$, for which we have $(g^1, d^1, T^1) = (g_{a, a+u}, d_{a, a+u}, T_{a+u})$. Inspecting the proof of Lemma 4.7, we see that T^2 converges: indeed, all that matters in the proof of Lemma 4.7 is the local behavior around b , for which the initial state v , as long as $\pi(v)$ is far below b , is irrelevant (note that this is the case when shifting $\pi \circ X$ at time d^1 , since by definition $\pi(X_{d^1}) \leq a$).

Moreover, since the successive excursions above a are i.i.d. by Lemma 1.2, Lemma 4.8 implies, since T^2 converges, that d^2 , g^2 and D_{d^2} converge. Iterating, we obtain the convergence of d^k , g^k and D_{d^k} for every $k \geq 1$. Finally, it is not hard to see that these convergences hold jointly, i.e., $((g^k, d^k, D_{d^k}), k \geq 1) \Rightarrow^n ((g^k, d^k, D_{d^k}), k \geq 1)$. There are two different ways to see this: either use arguments as in end of the discussion in Section 2.2, or use the fact that the results of Lemmas 4.7 and 4.8 actually show more than just weak convergence, but actually that the limiting functions have no fixed point of discontinuity, and then use the continuous mapping theorem.

Having the joint convergence with the D_{d^k} 's makes it possible to know whether two successive excursions above a with height $> u$ belong to the same excursion away from 0. In particular, if $k^* \geq 0$ is the first index such that $D_{d^{k^*}} < D_{d^{k^*+1}} = \dots = D_{d^{k^*+p}} <$

$D_{d^{k^*+p+1}}$ (defining $d^0 = 0$), then $(\hat{g}^k, \hat{d}^k) = (g^{k^*+k}, d^{k^*+k})$ for $k = 1, \dots, p$. From the convergence of $((g^k, d^k, D_{d^k}), k \geq 1)$ we obtain the convergence of k^* , which therefore entails the convergence of $((\hat{g}^k, \hat{d}^k), k = 1, \dots, p)$ as desired. This finally achieves the proof of (4.10).

4.4.2. *Proof of (4.11).* Let $m^k = \sup \pi \circ X$, where the supremum is taken over the interval $[d^k, D_{d^k}]$: then $(\hat{\epsilon}^k, k = 1, \dots, p)$ is equal in distribution to $(\epsilon^k, k = 1, \dots, p)$ conditionally on $\{m^p < a + u < m^1, \dots, m^{p-1}\}$ (which is to be understood as $\{m^1 < a + u\}$ when $p = 1$). In particular,

$$\mathbb{E}^n \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right) = \frac{\mathbb{E}^n \left[\prod_{k=1}^p f_k(\epsilon^k); m^p < a + u < m^1, \dots, m^{p-1} \right]}{\mathbb{P}^n (m^p < a + u < m^1, \dots, m^{p-1})}.$$

For $\nu \in \mathcal{M}_F$ we define

$$A_q^n(\nu) = \mathbb{E}_\nu^n \left(\prod_{k=1}^q f_{k+p-q}(\epsilon^k); m^q < a + u < m^1, \dots, m^{q-1} \right) \text{ for } q = 1, \dots, p,$$

so that

$$\mathbb{E}^n \left(\prod_{k=1}^p f_k(\hat{\epsilon}^k) \right) = \frac{A_p^n(\mathbf{z})}{A_p^{n,1}(\mathbf{z})}$$

with $A_p^{n,1}$ defined similarly as A_p^n by taking all the f_k 's equal to the constant function which takes value one. Moreover, let us introduce for $\nu \in \mathcal{M}_F$

$$\begin{cases} B_q^n(\nu) = \mathbb{E}_\nu^n \left(\prod_{k=1}^q f_{k+p-q}(\epsilon^k); g^1 < D_0, m^q < a + u < m^1, \dots, m^{q-1} \right) & \text{for } q = 1, \dots, p-1, \\ B_0^n(\nu) = \mathbb{P}_\nu^n (D_0 < d^1) & \text{for } q = 0, \\ \Delta_q^n(\nu) = \mathbb{E}_\nu^n \left[f_{p-q}(\epsilon^1) \times \left(B_q^n(X_{d^1}^n) - B_q^n(X_{g^1-}^n) \right) \right] & \text{for } q = 0, \dots, p-1 \end{cases}$$

(recall that $\nu^n \in \mathcal{M}_F$ is the measure such that $\vartheta_n(\nu^n) = \nu$) and finally

$$\varphi_q^n = \mathbb{E}^n [f_{p-q+1}(\epsilon^1)] \text{ for } q = 1, \dots, p.$$

Note that Lemma 1.2 together with Lemmas 4.7, 4.8, 2.1 and the definition of \mathcal{N} imply that

$$(4.14) \quad \varphi_q^n \xrightarrow{n \rightarrow +\infty} \mathcal{N}(f_{p-q+1}(\epsilon) \mid \sup \epsilon > u) \text{ for } q = 1, \dots, p.$$

We now derive some relations between all these quantities. First of all,

$$(4.15) \quad B_q^n(\nu) = A_q^n(\nu) - B_0^n(\nu) A_q^n(\mathbf{z}) \text{ for } q = 1, \dots, p-1.$$

Indeed, for $q = 1, \dots, p-1$ we have

$$B_q^n(\nu) = A_q^n(\nu) - \mathbb{E}_\nu^n \left(\prod_{k=1}^q f_{k+p-q}(\epsilon^k); D_0 < g^1, m^q < a + u < m^1, \dots, m^{q-1} \right),$$

and since $\pi \circ X$ regenerates at D_0 , the second term of the above right-hand side is equal to

$$\mathbb{P}_\nu^n (D_0 < g^1) \mathbb{E}^n \left(\prod_{k=1}^q f_{k+p-q}(\epsilon^k); m^q < a + u < m^1, \dots, m^{q-1} \right).$$

Since the two events $\{D_0 < g^1\}$ and $\{D_0 < d^1\}$ coincide, we obtain (4.15). Second, for $\nu \in \mathcal{M}_F$ with $\pi(\nu) < an$, we have

$$(4.16) \quad A_q^n(\nu) = \varphi_q^n \times \mathbb{E}_\nu^n \left[B_{q-1}^n(X_{g^1-}^n) \right] + \Delta_{q-1}^n(\nu) \text{ for } q = 1, \dots, p.$$

Indeed, the strong Markov property at time d^1 gives for $q = 1, \dots, p$

$$A_q^n(\nu) = \mathbb{E}_\nu^n \left[f_{p-q+1}(\epsilon^1) \times B_{q-1}^n(X_{d^1}^n) \right] = \mathbb{E}_\nu^n \left[f_{p-q+1}(\epsilon^1) \times B_{q-1}^n(X_{g^1-}^n) \right] + \Delta_{q-1}^n(\nu).$$

Since $\pi(\nu) < an$, under \mathbb{P}_ν^n , ϵ^1 and X_{g^1-} are independent, and ϵ^1 is distributed according to ϵ^1 under \mathbb{P}^n , which gives (4.16). Combining (4.15) and (4.16), we end up with the following recursion for A_q^n :

$$(4.17) \quad A_q^n(\nu) = \varphi_q^n \times \left[\mathbb{E}_\nu^n \left(A_{q-1}^n(X_{g^1-}^n) \right) - A_{q-1}^n(\mathbf{z}) \mathbb{E}_\nu^n \left(B_0^n(X_{g^1-}^n) \right) \right] + \Delta_{q-1}^n(\nu), \quad q = 2, \dots, p,$$

with the boundary condition

$$(4.18) \quad A_1^n(\nu) = \varphi_1^n \times \mathbb{E}_\nu^n \left[B_0^n(X_{g^1-}^n) \right] + \Delta_0^n(\nu).$$

Since the functions f_k were arbitrary in deriving this recursion, we obtain a similar recursion for $A_q^{n,1}(\nu)$, but with all the terms φ_q^n replaced by one and the $\Delta_q^n(\nu)$'s replaced by $\Delta_q^{n,1}(\nu)$, defined similarly as $\Delta_q^n(\nu)$ but with the functions f_q equal to the constant function taking value one. Now consider \tilde{A}_q^n and $\tilde{A}_q^{n,1}$ that satisfy the same recursion (4.17)–(4.18), but with all the Δ_q^n equal to 0, i.e., for every $\nu \in \mathcal{M}_F$,

$$\tilde{A}_q^n(\nu) = \varphi_q^n \times \left[\mathbb{E}_\nu^n \left(\tilde{A}_{q-1}^n(X_{g^1-}^n) \right) - \tilde{A}_{q-1}^n(\mathbf{z}) \mathbb{E}_\nu^n \left(B_0^n(X_{g^1-}^n) \right) \right], \quad q = 2, \dots, p,$$

with the boundary condition

$$\tilde{A}_1^n(\nu) = \varphi_1^n \times \mathbb{E}_\nu^n \left[B_0^n(X_{g^1-}^n) \right],$$

and similarly for $\tilde{A}_q^{n,1}(\nu)$ with all the terms φ_q^n replaced by one. By induction one gets

$$\frac{\tilde{A}_p^n(\mathbf{z})}{\tilde{A}_p^{n,1}(\mathbf{z})} = \prod_{k=1}^p \varphi_k^n$$

and so (4.14) implies that

$$(4.19) \quad \frac{\tilde{A}_p^n(\mathbf{z})}{\tilde{A}_p^{n,1}(\mathbf{z})} \xrightarrow{n \rightarrow +\infty} \prod_{k=1}^p \mathcal{N}(f_k(\epsilon) \mid \sup \epsilon > u).$$

On the other hand, for ν with $\pi(\nu) < an$, we have $A_1^n(\nu) - \tilde{A}_1^n(\nu) = \Delta_0^n(\nu)$ while for $q = 2, \dots, p$,

$$\begin{aligned} |A_q^n(\nu) - \tilde{A}_q^n(\nu)| &\leq \varphi_q^n \mathbb{E}_\nu^n \left(\left| A_{q-1}^n(X_{g^1-}^n) - \tilde{A}_{q-1}^n(X_{g^1-}^n) \right| \right) \\ &\quad + \varphi_q^n \mathbb{E}_\nu^n \left(B_0^n(X_{g^1-}^n) \right) \left| A_{q-1}^n(\mathbf{z}) - \tilde{A}_{q-1}^n(\mathbf{z}) \right| + \left| \Delta_{q-1}^n(\nu) \right|. \end{aligned}$$

Thus if $\varepsilon^{(n)} = \max_{q=0, \dots, p-1} \varepsilon_q^{(n)}$ with

$$\varepsilon_q^{(n)} = \sup_{\nu \in \mathcal{M}_F: \pi(\nu) < an} \left| \Delta_q^n(\nu) \right| \text{ for } q = 0, \dots, p-1,$$

then by induction we obtain $|A_p^n(\mathbf{z}) - \tilde{A}_p^n(\mathbf{z})| \leq C\varepsilon^{(n)}$ for some finite constant C , and a similar upper bound holds for $|A_p^{n,1}(\mathbf{z}) - \tilde{A}_p^{n,1}(\mathbf{z})|$ (note that, to perform the induction, we use the fact that $\pi(X_{g^1-}^n) < an$ \mathbb{P}_ν^n -almost surely, for any ν with $\pi(\nu) < an$). In view of (4.19), the following lemma therefore achieves the proof of (4.11).

Lemma 4.9. $\varepsilon^{(n)} \rightarrow 0$ as $n \rightarrow +\infty$.

Proof. By definition we have for $q = 0, \dots, p-1$

$$|\Delta_q^n(\nu)| \leq C \mathbb{E}_\nu^n \left[\left| B_q^n(X_{g^1_-}) - B_q^n(X_{d^1}) \right| \right]$$

with $C = \max_{q=1, \dots, p} \sup f_q$. To control the difference appearing in this last expectation, we use the following observation: under \mathbb{P}_ν^n for any $\nu \in \mathcal{M}_F$ with $\pi(\nu) < an$, we can write $X_{d^1}^n = X_{g^1_-}^n + \Xi$, where $\Xi \in \mathcal{M}_F$ corresponds to the orders added below a during the first excursion of $\pi \circ X$ above a with height $> u$. In particular, the coupling implies that Ξ is independent from $X_{g^1_-}^n$, its law does not depend on ν and $M(\Xi)$ is equal in distribution to $|\mathcal{K}(\mathcal{T}_1)|$ conditioned on $\psi^*(B(\mathcal{T}_1)) > un$. In particular,

$$(4.20) \quad \varepsilon_q^{(n)} \leq C \mathbb{E}^n \left[\sup_{\nu \in \mathcal{M}_F: \pi(\nu) < an} \left| B_q^n(\nu) - B_q^n(\nu + \Xi) \right| \right].$$

In view of the definition of B_q^n , we thus need to show, roughly speaking, that the processes X under \mathbb{P}_ν^n and $\mathbb{P}_{\nu+\tilde{\nu}}^n$ are close, uniformly in ν . Note more precisely that all the events and random variables involved in the computation of B_q^n depend on X stopped at D_0 , and so we actually only need to show that the processes $(X_t, 0 \leq t \leq D_0)$ under \mathbb{P}_ν^n and $\mathbb{P}_{\nu+\tilde{\nu}}^n$ are close.

In order to do so we extend the coupling of Theorem 1.1. In essence, Theorem 1.1 states a coupling between X started from one single order at a and \mathcal{T}_a . This coupling can be readily extended by coupling X under \mathbb{P}_ν and a forest of trees $\mathcal{F}(\nu) = (\mathcal{T}_{(k)}, k = 1, \dots, M(\nu))$. The trees $\mathcal{T}_{(k)}$ are independent, and if $\nu = \sum_a \zeta_a \delta_a$, then exactly ζ_a of the trees $\mathcal{T}_{(k)}$ are distributed like \mathcal{T}_a .

This coupling between \mathbb{P}_ν and $\mathcal{F}(\nu)$ provides a coupling between \mathbb{P}_ν and $\mathbb{P}_{\nu+\tilde{\nu}}$: first, one considers the forest $\mathcal{F}(\nu)$ used to construct X under \mathbb{P}_ν . Then, one adds $M(\tilde{\nu})$ independent trees to this forest, say $(\tilde{\mathcal{T}}_{(k)}, k = 1, \dots, M(\tilde{\nu}))$, to get a larger forest $\tilde{\mathcal{F}} = \mathcal{F}(\nu) \cup \{\tilde{\mathcal{T}}_{(k)}, k = 1, \dots, M(\tilde{\nu})\}$. By exploring this forest with successive iterations of Φ , we get a new process \tilde{X} on the same probability space that X , which is a version of X under $\mathbb{P}_{\nu+\tilde{\nu}}$.

Note moreover that, as mentioned previously, we are only interested in X before time D_0 . In particular, we can truncate the trees $\mathcal{T}_{(k)}$ and $\tilde{\mathcal{T}}_{(k)}$ by removing all the nodes that have a label ≤ 0 . It is thus convenient to consider the operator $B_0: \mathbb{T} \rightarrow \mathbb{T}$ that removes all the nodes of a tree $T \in \mathbb{T}$ with label ≤ 0 .

If $\tilde{\varepsilon}^k, \tilde{g}^1, \tilde{m}^k$ and \tilde{D}_0 are the quantities associated to \tilde{X} in the same way that ε^k, g^1, m^k and D_0 are associated to X , then by definition of B_q^n we have

$$B_0^n(\nu) - B_0^n(\nu + \tilde{\nu}) = \mathbb{E}_\nu^n \left(\mathbb{1}_{\{D_0 < d^1\}} - \mathbb{1}_{\{\tilde{D}_0 < \tilde{d}^1\}} \right)$$

for $q = 0$, while for $q = 1, \dots, p-1$, $B_q^n(\nu) - B_q^n(\nu + \tilde{\nu})$ is equal to

$$\mathbb{E}_\nu^n \left(\prod_{k=1}^q f_k(\varepsilon^k) \mathbb{1}_{\{g^1 < D_0, m^q < a+u < m^1, \dots, m^{q-1}\}} - \prod_{k=1}^q f_k(\tilde{\varepsilon}^k) \mathbb{1}_{\{\tilde{g}^1 < \tilde{D}_0, \tilde{m}^q < a+u < \tilde{m}^1, \dots, \tilde{m}^{q-1}\}} \right).$$

Now the key observation is that in the event $\{\max_k \psi^*(B_0(\tilde{\mathcal{T}}_{(k)})) < (a+u)n\}$, the two random variables (the one defined in terms of X and the one defined in terms of \tilde{X}) in the previous expectation are equal. Indeed, in this event, the excursions above a with height $> u$ for X and \tilde{X} coincide. In particular, since the random variables under

consideration are bounded, we obtain

$$|B_q^n(v) - B_q^n(v + \tilde{v})| \leq C \mathbb{P}_v^n \left(\max_{k=1, \dots, M(\tilde{v})} \psi^*(B_0(\tilde{\mathcal{T}}_{(k)})) > (a+u)n \right).$$

Recall that the trees $\tilde{\mathcal{T}}_{(k)}$ are independent. Further, $\psi^*(B_0(\mathcal{T}_y)) \leq \psi^*(\mathcal{T}_y)$, and $\psi^*(\mathcal{T}_y)$ is (stochastically) increasing in y , so that using the union bounds we get for any $\tilde{v} \in \mathcal{M}_F$ with $\pi(\tilde{v}) < an$

$$\mathbb{P}_v^n \left(\max_{k=1, \dots, M(\tilde{v})} \psi^*(B_0(\tilde{\mathcal{T}}_{(k)})) > (a+u)n \right) \leq M(\tilde{v}) \mathbb{P}(\psi^*(\mathcal{T}_0) > un).$$

In view of (4.20) and the discussion preceding it, we therefore get

$$\varepsilon_q^{(n)} \leq C \mathbb{E}(\mathcal{N}(\mathcal{T}_1) \mid \psi^*(B(\mathcal{T}_1)) > un) \times \mathbb{P}(\psi^*(\mathcal{T}_0) > un).$$

The supremum over $n \geq 1$ of the expectation in the above right-hand side is finite by (A.16), and since $\mathbb{P}(\psi^*(\mathcal{T}_0) > un) \rightarrow 0$ as $n \rightarrow +\infty$, the result is proved. \square

4.5. Fifth step : $\pi \circ X$ under \mathbf{P} is a reflected Brownian motion. At this point, we know that \mathcal{N} is a σ -finite measure on \mathcal{E} that satisfies the following properties:

- I) $\mathcal{N}(\zeta = +\infty) = 0$ (since $\pi \circ X$ under \mathbf{P} is upper bounded by $M \circ X$ by Lemma 4.2, which by Lemma 4.1 is a Brownian motion with no drift reflected at 0);
- II) $\mathcal{N}(\varepsilon \text{ is not continuous}) = 0$ (since $\pi \circ X$ under \mathbf{P} is almost surely continuous).

In particular, \mathcal{N} induces a σ -finite measure Θ on the set of compact real trees via the usual coding of a compact real tree by a continuous excursion with finite length, see for instance Le Gall and Miermont [23, Section 3].

Further, let $y > 0$ and ε^1 be the first excursion of $\pi \circ X$ away from 0 with height $> y$. Then by Lemma 4.7 and 4.8, we have $\mathbb{P}^n(\sup \varepsilon^1 > x) \rightarrow \mathbf{P}(\sup \varepsilon^1 > x)$ for all x outside a countable set, where this latter quantity is equal to $\mathcal{N}(\sup \varepsilon > x \mid \sup \varepsilon > y)$ by definition of \mathcal{N} . On the other hand,

$$\mathbb{P}^n(\sup \varepsilon^1 > x) = \mathbb{P}^n(\sup \sigma_{D_0} > x \mid \sup \sigma_{D_0} > y) = \mathbb{P}(\psi^*(B(\mathcal{T}_1)) > xn \mid \psi^*(B(\mathcal{T}_1)) > yn)$$

which, for any $0 < y < x$, converges to y/x by Lemma A.3. Thus for all $x > y$ outside a countable set, we have

$$\mathcal{N}(\sup \varepsilon > x \mid \sup \varepsilon > y) = \frac{\mathcal{N}(\sup \varepsilon > x)}{\mathcal{N}(\sup \varepsilon > y)} = \frac{y}{x}$$

from which one deduces that $\mathcal{N}(\sup \varepsilon > x) = c/x$ for every $x > 0$, and for some finite constant $c > 0$ (this constant will be fixed shortly). Thus \mathcal{N} satisfies the following additional properties:

- III) $\mathcal{N}(\sup \varepsilon = 0) = 0$ (by definition of an excursion measure);
- IV) $0 < \mathcal{N}(\sup \varepsilon > x) < +\infty$ for every $x > 0$ (since $\mathcal{N}(\sup \varepsilon > x) = c/x$);
- V) $\mathcal{N}(\mathcal{E}) = +\infty$ (obtained by letting $x \downarrow 0$ in $\mathcal{N}(\sup \varepsilon > x) = c/x$);
- VI) \mathcal{N} satisfies the regenerative property (R) (by the previous step).

Properties III)–V) above immediately translate to Θ having infinite mass, $\Theta(\mathcal{H} = 0) = 0$ and $\Theta(\mathcal{H} > x) \in (0, \infty)$ (where \mathcal{H} denotes the height of the canonical tree \mathfrak{t}). Moreover, the last property VI) means exactly that Θ satisfies the property (R) of Weill [27]: indeed, excursions of ε above level a under \mathcal{N} correspond to the subtrees of \mathfrak{t} above level a under Θ . Finally, we see that the assumptions of Theorem 1.1 in Weill [27] are satisfied, which gives the existence of a spectrally positive Lévy process Y , with Laplace

exponent Ψ satisfying $\int^\infty (1/\Psi) < +\infty$, such that Θ is the (excursion) law of the Ψ -Lévy tree. In particular, \mathcal{N} is an excursion measure of the height process associated to Y .

We now fix the normalization constant as in Duquesne and Le Gall [8] (which amounts to choosing the constant c above), so that according to Corollary 1.4.2 in Duquesne and Le Gall [8] (remember that $\int^\infty (1/\Psi) < +\infty$) we have

$$\int_{\mathcal{N}(\sup \epsilon > x)}^\infty \frac{du}{\Psi(u)} = x, \quad x > 0,$$

which implies, since $\mathcal{N}(\sup \epsilon > x) = c/x$, that $\Psi(u) = u^2/c$. In other words, Y is equal in distribution to $(2/c)^{1/2}\tilde{W}$, with \tilde{W} a standard Brownian motion, and the height process associated to this Lévy process is equal in distribution to $(2c)^{1/2}W$ (to see this, consider for instance the CSBP Z associated to Y , which has branching mechanism Ψ and satisfies the SDE $dZ_t = (2Z_t/c)^{1/2}d\tilde{W}_t$, and use (20) and (21) in Pardoux and Wakolbinger [24]). Since $\pi \circ X$ under \mathbf{P} is equal in distribution to the height process of Y , we obtain that $\pi \circ X$ under \mathbf{P} is equal in distribution to $(2c)^{1/2}W$. The following lemma makes it possible to identify c and, more importantly, to conclude the proof of Theorem 1.1.

Lemma 4.10. *For any $\eta > 0$,*

$$\limsup_{n \rightarrow +\infty} \mathbb{P}^n \left(\left| L_t^{n,\pi} - \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du \right| \geq \eta \right) \xrightarrow{\varepsilon \downarrow 0} 0.$$

Proof. Writing $\int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du = \ell_t + \int_0^t \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon\}} du$ and using the triangular inequality, we first obtain

$$\begin{aligned} \mathbb{P}^n \left(\left| L_t^{n,\pi} - \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{\pi(X_u) \leq \varepsilon\}} du \right| \geq \eta \right) &= \mathbb{P}^n \left(L_t^{n,\pi} \geq \eta \varepsilon n / 2 \right) \\ &\quad + \mathbb{P}^n \left(\left| L_t^{n,\pi} - \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon\}} du \right| \geq \eta / 2 \right). \end{aligned}$$

The first term of the above upper bound goes to 0 by Lemma 4.4, and so we need to control the second term. Rescaling leads to

$$\begin{aligned} \mathbb{P}^n \left(\left| L_t^{n,\pi} - \frac{1}{\varepsilon} \int_0^t \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon\}} du \right| \geq \eta / 2 \right) \\ = \mathbb{P} \left(\left| \int_0^{n^2 t} \mathbb{1}_{\{\pi(X_u) = 0\}} du - \frac{1}{\varepsilon n} \int_0^{n^2 t} \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon n\}} du \right| \geq \eta n / 2 \right). \end{aligned}$$

Let as in the proof of Lemma 4.4 $K(y)$ be the number of excursions of $\pi \circ X$ away from 0 that end before time y , E^k be the time that $\pi \circ X$ stays at 0 before the k th excursion and $V^k(y)$ be the time spent exploring nodes with label $\leq y$ in the k th ambient tree: then if $\pi(X_{n^2 t}) = 0$, we have

$$\left| \int_0^{n^2 t} \mathbb{1}_{\{\pi(X_u) = 0\}} du - \frac{1}{\varepsilon n} \int_0^{n^2 t} \mathbb{1}_{\{0 < \pi(X_u) \leq \varepsilon n\}} du \right| \leq E^{K(n^2 t)} + \left| \sum_{k=1}^{K(n^2 t)} \left(E^k - \frac{1}{\varepsilon n} V^k(\varepsilon n) \right) \right|.$$

If $\pi(X_{n^2 t}) > 0$, then the residual term, instead of being $E^{K(n^2 t)}$, is the time spent exploring nodes with label $\leq \varepsilon n$ in the $K(n^2 t)$ th ambient tree. In each case, one can show that this residual term does not contribute in the regime $n \rightarrow +\infty$ and then $\varepsilon \rightarrow 0$ that

we are interested in, and so we only have to show that

$$\limsup_{n \rightarrow +\infty} \mathbb{P} \left(\left| \sum_{k=1}^{K(n^2 t)} \left(E^k - \frac{1}{\varepsilon n} V^k(\varepsilon n) \right) \right| \geq \eta n \right) \xrightarrow{\varepsilon \downarrow 0} 0.$$

For $y > 0$ introduce the following quantities: $m(y) = \mathbb{E}(E^1) - \mathbb{E}(V^1(y))/y$, $Y^k(y) = E^k - V^k(y)/y - m(y)$, $\sigma(y)^2 = \mathbb{E}(Y(y)^2)$, $\bar{Y}(y) = Y(y)/\sigma(y)$ and

$$\Sigma(n, \varepsilon) = \left| \frac{1}{K(n^2 t)^{1/2}} \sum_{k=1}^{K(n^2 t)} \bar{Y}^k(\varepsilon n) \right|.$$

Then the triangular inequality gives

$$\left| \sum_{k=1}^{K(n^2 t)} \left(E^k - \frac{1}{\varepsilon n} V^k(\varepsilon n) \right) \right| \leq \sigma(\varepsilon n) K(n^2 t)^{1/2} \Sigma(n, \varepsilon) + m(\varepsilon n) K(n^2 t)$$

and so

$$(4.21) \quad \mathbb{P} \left(\left| \sum_{k=1}^{K(n^2 t)} \left(E^k - \frac{1}{\varepsilon n} V^k(\varepsilon n) \right) \right| \geq \eta n \right) \leq \mathbb{P} \left(K(n^2 t) \geq \frac{\eta n}{2m(\varepsilon n)} \right) + \mathbb{P} \left(\Sigma(n, \varepsilon) \geq \frac{\eta n}{2\sigma(\varepsilon n) K(n^2 t)^{1/2}} \right).$$

Let $\bar{C} = \sup_{y \geq 0} (y^{-1/2} \mathbb{E}(K(y)))$, which has been showed in the proof of Lemma 4.4, to be finite. Using Markov's inequality, the first term of the above upper bound is thus upper bounded by

$$(4.22) \quad \mathbb{P} \left(K(n^2 t) \geq \frac{\eta n}{2m(\varepsilon n)} \right) \leq (2/\eta) \bar{C} t^{1/2} \times m(\varepsilon n).$$

By definition we have for $y > 0$

$$m(y) = \frac{1}{\lambda \mathbb{P}(J=1)} - \frac{1}{\lambda y} \mathbb{E} \left(\sum_{v \in B(\mathcal{J}_1)} \mathbb{1}_{\{\psi(v, \mathcal{J}_1) \leq y\}} \right)$$

and so (A.9) implies that $m(y) \rightarrow 0$. In view of (4.22), the first term in the right-hand side of (4.21) therefore vanishes as $n \rightarrow +\infty$. We now control the second term: let $C^* = \sup_{y \geq 0} (\sigma(y)/y^{1/2})$, which is proved to be finite in Section A.2, so that

$$\begin{aligned} \mathbb{P} \left(\Sigma(n, \varepsilon) \geq \frac{\eta n}{2\sigma(\varepsilon n) K(n^2 t)^{1/2}} \right) &\leq \mathbb{P} \left(\Sigma(n, \varepsilon) \geq \frac{\eta n^{1/2}}{2\varepsilon^{1/2} C^* K(n^2 t)^{1/2}} \right) \\ &\leq \mathbb{P} \left(\Sigma(n, \varepsilon) \geq \frac{\eta}{2\varepsilon^{1/2} C^* \bar{K}} \right) + \frac{\bar{C} t^{1/2}}{\bar{K}^2}. \end{aligned}$$

where the second inequality, valid for any \bar{K} , is obtained by considering the two events $\{n/K(n^2 t) \geq 1/\bar{K}^2\}$ and $\{n/K(n^2 t) \leq 1/\bar{K}^2\}$ and using Markov inequality in the second case. Since the $(\bar{Y}^k(\varepsilon n), k \geq 1)$ are i.i.d. centered random variables with unit variance, the central limit theorem gives

$$\limsup_{n \rightarrow +\infty} \mathbb{P} \left(\Sigma(n, \varepsilon) \geq \frac{\eta}{2\varepsilon^{1/2} C^* \bar{K}} \right) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Thus letting first $n \rightarrow +\infty$, then $\varepsilon \rightarrow 0$ and finally $\bar{K} \rightarrow +\infty$ achieves the proof. \square

4.6. **Last step.** At this point, we know that, under \mathbf{P} :

- (1) $\pi \circ X$ is equal in distribution to $(2c)^{1/2}W$ (by the fifth step);
- (2) $M \circ X$ is equal in distribution to $(2\lambda)^{1/2}W$ (by Lemmas 4.1 and 4.2);
- (3) $\mathcal{L}(\pi \circ X) = \mathcal{L}(\mathbb{E}(J)M \circ X)$ (by Lemmas 4.2 and 4.10);

These three properties have the following consequence.

Lemma 4.11. *Under \mathbf{P} , for every $t \geq 0$ we have $M(X_t) = \pi(X_t)/\mathbb{E}(J)$.*

Before proving this lemma, let us quickly conclude the proof of Theorem 1.1. Fix some $t, y \geq 0$: we have to prove (1.2). If $\pi(X_t) = 0$, then $M(X_t) = 0$ by Lemma 4.11 and (1.2) holds. Otherwise, assume first that $y < \pi(X_t)$ and let g be the left endpoint of the excursion of $\pi \circ X$ above y straddling t . Then according to Corollary 4.3, we have $X_t([0, y]) = X_g([0, y])$. On the other hand, we have $y = \pi(X_g)$ by definition of g and so $X_g([0, y]) = M(X_g)$ which is equal to $\pi(X_g)/\mathbb{E}(J) = y/\mathbb{E}(J)$ by Lemma 4.11. This proves that $X_t([0, y]) = y/\mathbb{E}(J)$ for $y < \pi(X_t)$, and since $X_t([0, y]) = M(X_t)$ for $y \geq \pi(X_t)$, which is equal to $\mathbb{E}(J)^{-1}\pi(X_t)$ by Lemma 4.11, this proves (1.2) and concludes the proof of Theorem 1.1.

Proof of Lemma 4.11. Thanks to (2.1) we can write $\pi = c\mathcal{L}(\pi) + \bar{\pi}$ and $M = \lambda\mathcal{L}(M) + \bar{M}$, where $(2c)^{-1/2}\bar{\pi}$ and $(2\lambda)^{-1/2}\bar{M}$ are two standard Brownian motions, and in the rest of the proof we write in order to ease the notation π and M for $\pi \circ X$ and $M \circ X$, respectively. Moreover, $\mathcal{L}(\pi)$ is on the one hand equal in distribution to $\mathcal{L}((c/\lambda)^{1/2}M)$ because π is equal in distribution to $(c/\lambda)^{1/2}M$, while on the other hand we have $\mathcal{L}(\pi) = \mathcal{L}(\mathbb{E}(J)M)$. This shows that $c = \lambda\mathbb{E}(J)^2$ and in particular, we have $M = (c/\mathbb{E}(J))\mathcal{L}(\pi) + \bar{M}$.

Fix some $a, t \geq 0$: then we can apply the optional sampling theorem (as in, e.g., Karatzas and Shreve [17, Problem 1.3.23a]) for the bounded stopping time $T_a \wedge t$ and the martingales \bar{M} and $\bar{\pi}$, and derive

$$(4.23) \quad \mathbf{E}(M_{T_a \wedge t}) = (c/\mathbb{E}(J))\mathbf{E}(\mathcal{L}(\pi)_{T_a \wedge t}) = (1/\mathbb{E}(J))\mathbf{E}(\pi_{T_a \wedge t}).$$

Since $(\bar{M}^2 - 2\lambda t, t \geq 0)$ and $(\bar{\pi}^2 - 2ct, t \geq 0)$ are also martingales, another application of the optional sampling theorem implies

$$(4.24) \quad \mathbf{E}(M_{T_a \wedge t}^2) = 2\lambda\mathbf{E}(T_a \wedge t) = \frac{2\lambda}{2c}\mathbf{E}(\pi_{T_a \wedge t}^2) = \frac{1}{\mathbb{E}(J)^2}\mathbf{E}(\pi_{T_a \wedge t}^2) \leq \frac{a^2}{\mathbb{E}(J)^2}.$$

The stopped process $(M_{T_a \wedge t}, t \geq 0)$ is therefore uniformly integrable, and letting $t \rightarrow +\infty$ in (4.23), we thus obtain $\mathbf{E}(M_{T_a}) = \mathbf{E}(\pi_{T_a})/\mathbb{E}(J) = a/\mathbb{E}(J)$. On the other hand, letting $t \rightarrow +\infty$ in (4.24) and using Fatou's lemma, we obtain $\mathbf{E}(M_{T_a}^2) \leq (a/\mathbb{E}(J))^2$ which implies that $M_{T_a} = a/\mathbb{E}(J)$.

A similar calculation to (4.24) shows that the stopped Brownian motions $(\bar{M}_{T_a \wedge t}, t \geq 0)$ and $(\bar{\pi}_{T_a \wedge t}, t \geq 0)$ are uniformly integrable. Then we can apply another version of the optional sampling theorem, such as in Karatzas and Shreve [17, Problem 1.3.19 and Theorem 1.3.22], and get

$$\mathbf{E}(M_{T_a} | \mathcal{F}_{t \wedge T_a}) = (c/\mathbb{E}(J))\mathbf{E}(\mathcal{L}(\pi)_{T_a} | \mathcal{F}_{t \wedge T_a}) + \bar{M}_{T_a \wedge t}.$$

Since we have proved that $M_{T_a} = a/\mathbb{E}(J)$, the last display leads to

$$a - c\mathbf{E}(\mathcal{L}(\pi)_{T_a} | \mathcal{F}_{t \wedge T_a}) = \mathbb{E}(J)\bar{M}_{T_a \wedge t}.$$

The exact same reasoning shows that the left hand side of the above display is also equal to $\bar{\pi}_{T_a \wedge t}$, and so $M_{t \wedge T_a} = \pi_{t \wedge T_a}/\mathbb{E}(J)$. Letting $a \rightarrow +\infty$ achieves the proof. \square

5. DISCUSSION

The main purpose of this paper was to exploit the (elegant, in our opinion) connection between the regenerative characterization of Lévy trees of Weill [27] and the present model of limit order book. The assumptions made on J in Theorem 1.1 correspond to the simplest interesting case where this connection can be exploited, but this result should hold under more general assumptions on J and λ . For instance, our arguments should readily extend to a triangular scheme where the rates at which orders are added to and removed from the book may be different, and the model's parameters depend on n in a suitable way. Typically, the results of Theorem 1.1 would still hold, with the limiting price process being a Brownian motion *with drift* reflected at 0.

A more delicate generalization consists in relaxing the assumption that $J \in \{-j^*, \dots, 1\}$. The proof of most results goes through in this more general case, but the main problem is that for a general random variable J , the successive excursions above level a are not i.i.d. anymore, which invalidates Lemma 1.2. However, the dependency between successive excursions lies in the overshoot of the price above a at the beginning of each excursion above a , and so upon suitable moment assumptions on J this dependency should be washed out in the limit.

Further, different boundary conditions can also be considered. In Lakner et al. [20] and Simatos [26] for instance, orders can be placed in the negative half-line. In Lakner et al. [20] there is the additional constraint that the number of orders cannot fall below some level, say εn . This is meant to model the presence of a *market maker*.

In the presence of such a market maker, Theorem 1.1 remains valid and the proofs go through. Indeed, imagine εn orders initially sit at 0. Since these orders can only be displaced when the price is at 0 and, while the price is at 0, the number of orders evolves according to a critical random walk, the price process needs to accumulate of the order of n^2 units of local time at 0 in order to go through this initial stack of orders. Lemma 4.1 shows that this takes of the order of n^4 units of time, and so on the time scale that we are interested in, this does not happen. Pushing this reasoning a bit further actually shows that Theorem 1.1 should remain valid as long as the minimal number of orders, say m_n , diverges to $+\infty$. Indeed, in this case after accumulating m_n^2 units of local time at 0 these orders will have only moved by a constant distance, and it would thus take nm_n^2 units of local time at 0, which take about $n^2 m_n^4 \gg n^2$ units of normal time to accumulate, to have them moved by a distance of the order of n .

On the other hand, when orders can be placed on the negative half-line and there is no market maker, then we conjecture that the price process should converge to a Brownian motion (without reflection), say \tilde{W} , and the measure-valued process should converge to the process having constant density $1/\mathbb{E}(J)$ with respect to Lebesgue measure restricted to $[I_t, \tilde{W}_t]$ with $I_t = -\inf_{[0,t]} \tilde{W}$. The key observation is indeed that if, in this “free” case, one reflects the measure-valued process by considering I^π , the past infimum of the price process, as the origin of space and collapsing all the orders below I^π at I^π , then one precisely gets the model studied here. Thus the only thing left to prove would be that I^π converges to the local time at 0 of the reflected price process.

Let us finally mention that we have focused here on the case $\mathbb{E}(J) > 0$. When $\mathbb{E}(J) < 0$, under moment minor assumptions on J the probability $\mathbb{P}(\psi^*(B(\mathcal{F}_1)) > u)$ decays *exponentially fast*, since for this to happen one needs the supremum of a random walk with a negative drift to be large (see for instance Theorem 2 in Addario-Berry and Broutin [1]).

This is in sharp contrast with the polynomial decay proved in Lemma A.3 when $\mathbb{E}(J) > 0$, and it implies, when $\mathbb{E}(J) < 0$, that $\pi \circ X$ under \mathbb{P}^n converges weakly to 0 (since one would need to see an exponential number of excursions before seeing a macroscopic one). Note that the case $\mathbb{E}(J) < 0$ with a different boundary condition (see discussion above) has been studied in Lakner et al. [20] via a completely different approach. The fact that $\pi \circ X$ under \mathbb{P}^n converges to 0 means, in terms of the free process studied in [20], that the limiting price process is increasing (see Proposition 9.12 there).

To conclude, we note that the case $\mathbb{E}(J) = 0$, which is in some sense the true critical case where both the offspring and displacement distributions of \mathcal{T}_1 are critical, remains open.

APPENDIX A. RESULTS ON $B(\mathcal{T}_1)$

We prove in this section the various results on $B(\mathcal{T}_1)$ that have been used in the proof of Theorem 1.1. These results may also be of independent interest, see for instance Durrett et al. [9] and Kesten [18] where closely related results are proved for \mathcal{T}_1 .

Note that we consider the case of a geometric offspring distribution, but except for the second estimate in (A.3) this does not play a role in the proofs. Similarly, except for some explicit expressions such as (A.4), the following results would hold for any random variable J with $\mathbb{E}(J) > 0$ and a suitable moment assumption, e.g., $\mathbb{E}(e^{\rho J}) < +\infty$ for some $\rho > 0$.

A.1. Preliminary results. Let in the sequel $Z_m = \sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{|v|=m\}}$ for $m \geq 1$ be the number of nodes at depth m in \mathcal{T}_1 , so that $(Z_m, m \geq 1)$ is a Galton Watson branching process with geometric offspring distribution with parameter $1/2$, $h(\mathcal{T}_1)$ is its extinction time and $|\mathcal{T}_1|$ is its total progeny. By induction one easily proves that

$$(A.1) \quad \mathbb{E}[(Z_m - 1)^2] = 2m, \quad m \geq 1$$

(the factor 2 corresponds to the variance of the offspring distribution). Because of the memoryless property of the geometric distribution, the contour function of Z_m is a simple random walk, say $(\mathcal{C}(k), k \geq 0)$, started at 1 and stopped at its time of first visit to 0 (see for instance Proposition 2.6 in Le Gall and Miermont [23]). In terms of \mathcal{C} , $2Z_m$ is its local time at m , $h(\mathcal{T}_1)$ is its supremum and $2|\mathcal{T}_1|$ is the time of its first visit to 0. The following results can therefore be derived using standard results from random walk theory: first, as $u \rightarrow +\infty$ we have

$$(A.2) \quad \mathbb{P}(|\mathcal{T}_1| \geq u) \underset{u \rightarrow +\infty}{\sim} C_S / u^{1/2} \quad \text{and} \quad P(h(\mathcal{T}_1) \geq u) \underset{u \rightarrow +\infty}{\sim} 1/u,$$

where $C_S > 0$ is some finite constant, and second, there exists a finite constant $C > 0$ such that for every $u, m \geq 1$,

$$(A.3) \quad \mathbb{E}(Z_m \mid |\mathcal{T}_1| > u) \leq Cm \quad \text{and} \quad \mathbb{E}(Z_m \mid h(\mathcal{T}_1) > u) \leq Cm.$$

Note that (A.2) and the bound on $\mathbb{E}(Z_m \mid |\mathcal{T}_1| > u)$ are known to hold for any critical Galton Watson branching process with finite variance (see, e.g., Theorem 23 in Aldous [2] for (A.2) and Theorem 1.13 in Janson [15] for the bound on $\mathbb{E}(Z_m \mid |\mathcal{T}_1| > u)$). Also, most of the times upper and lower bounds on $\mathbb{P}(|\mathcal{T}_1| \geq u)$ and $\mathbb{P}(h(\mathcal{T}_1) \geq u)$ will be enough, and we will for instance often write

$$1/(Cu^{1/2}) \leq \mathbb{P}(|\mathcal{T}_1| \geq u) \leq C/u^{1/2} \quad \text{and} \quad 1/(Cu) \leq P(h(\mathcal{T}_1) \geq u) \leq C/u.$$

Let in the rest of this section $S = (S_m, m \geq 0)$ be a random walk started at 0 and with step distribution J , independent from \mathcal{T}_1 , and $\underline{S}_m = \min_{0 \leq k \leq m} S_k$.

Lemma A.1. *We have $\mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|^2) < +\infty$ and*

$$(A.4) \quad \mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|) = 1 - \frac{\mathbb{E}(J)}{\mathbb{P}(J=1)}.$$

Proof. To compute the mean number of killed nodes, we write

$$|\mathcal{K}(\mathcal{T}_1)| = \sum_{m \geq 1} \sum_{v \in \mathcal{T}_1: |v|=m} f(v) \text{ with } f(v) = \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq 0, \psi(v_1, \mathcal{T}_1), \dots, \psi(v_{m-1}, \mathcal{T}_1) \geq 1\}}.$$

Thus, taking the mean and using the fact that labels are independent from the genealogical structure, we obtain

$$\mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|) = \sum_{m \geq 1} \mathbb{P}(S_m < 0, \underline{S}_{m-1} \geq 0) \mathbb{E}(Z_m).$$

Since the genealogical structure Z of \mathcal{T}_1 is a critical Galton Watson process, we have $\mathbb{E}(Z_m) = 1$ which gives $\mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|) = \mathbb{P}(\underline{S}_\infty < 0)$. Since $J \in \{-j^*, -j^* + 1, \dots, 0, 1\}$, S is, in the terminology of Brown et al. [6], a skip-free (to the right) random walk with positive drift. In particular, Corollary 1 in this reference implies that $\mathbb{P}(\underline{S}_\infty \geq 0) = \mathbb{E}(J)/\mathbb{P}(J=1)$ which gives (A.4).

As for the second moment, we define $v \wedge v'$ for $v, v' \in \mathcal{T}_1$ as the most recent common ancestor of v and v' , and write $|\mathcal{K}(\mathcal{T}_1)|^2 = |\mathcal{K}(\mathcal{T}_1)| + \Sigma$, so that we only have to prove that $\mathbb{E}(\Sigma) < +\infty$, with

$$\Sigma = \sum_{\substack{M \geq 1 \\ m, m' \geq M}} \sum_{V: |V|=M} \sum_{\substack{v: |v|=m \\ v': |v'|=m'}} f(v)f(v') \mathbb{1}_{\{v \wedge v' = V, v \neq v'\}}.$$

Let $M \geq 1$, $m, m' \geq M$ and $V, v, v' \in \mathcal{T}_1$ with $|V| = M$, $|v| = m$, $|v'| = m'$, $v \wedge v' = V$ and $v \neq v'$. If v is an ancestor of v' (or the other way around), then $f(v)f(v') = 0$. Otherwise, $m, m' > M$ and the paths from the root to v and v' coincide on the first M steps and are independent afterwards, on the $m - M$ and $m' - M$ remaining steps, respectively. Thus in this case, if S' is an independent copy of S we have

$$\begin{aligned} \mathbb{E}(f(v)f(v') | Z) &= \mathbb{P}(\underline{S}_{m-1} \geq 0, S_m < 0, S_M + S'_1, \dots, S_M + S'_{m'-M-1} \geq 0, S_M + S'_{m'-M} < 0) \\ &= \mathbb{E}[g(m-M, S_M)g(m'-M, S_M); \underline{S}_M \geq 0] \end{aligned}$$

where $g(i, s) = \mathbb{P}(\underline{S}_{i-1} \geq -s, S_i < -s)$ for $i \geq 1$ and $s \in \mathbb{N}$. Defining $g(0, s) = 0$, we therefore have

$$\mathbb{E}(\Sigma) \leq \sum_{\substack{M \geq 1 \\ m, m' \geq M}} \mathbb{E}(g(m-M, S_M)g(m'-M, S_M)) \mathbb{E} \left(\sum_{V: |V|=M} \sum_{\substack{v: |v|=m \\ v': |v'|=m'}} \mathbb{1}_{\{v \wedge v' = V\}} \right).$$

Since by the branching property, the subtrees rooted at nodes at depth M in the tree are i.i.d., independent from the number Z_M of nodes at depth M , and since further $\mathbb{E}(Z_M) = 1$, we have

$$\mathbb{E} \left(\sum_{V: |V|=M} \sum_{\substack{v: |v|=m \\ v': |v'|=m'}} \mathbb{1}_{\{v \wedge v' = V\}} \right) = \mathbb{E} \left(\sum_{\substack{v: |v|=m-M \\ v': |v'|=m'-M}} \mathbb{1}_{\{v \wedge v' = \emptyset\}} \right).$$

To count the number of nodes at depths $m-M$ and $m'-M$ with most recent common ancestor the root, we can pick two distinct children of the root and then count the

number of nodes at depth $m-M-1$ and $m'-M-1$ in each subtree, so that

$$\sum_{\substack{v:|v|=m-M \\ v':|v'|=m'-M}} \mathbb{1}_{\{v \wedge v' = \emptyset\}} = \sum_{u, u':|u|=|u'|=1} Z(u, m-M-1)Z(u', m'-M-1)$$

where $Z(w, i)$ is the number of nodes at depth i in the subtree of \mathcal{T} rooted at $w \in \mathcal{T}$. Thus taking the mean and noting that the number of distinct pairs of children of the root is equal in distribution to $Z_1(Z_1 - 1)$ which has mean 2, we obtain

$$\mathbb{E} \left(\sum_{V:|V|=M} \sum_{\substack{v:|v|=m \\ v':|v'|=m'}} \mathbb{1}_{\{v \wedge v' = V, v \neq v'\}} \right) = 2\mathbb{E}(Z_{m-M-1}Z_{m'-M-1}) = 2\mathbb{E}(Z_{m^*})$$

with $m^* = \min(m-M, m'-M) + 1$. Using that $g(0, s) = 0$, upper bounding m^* by $3((m-M) + (m'-M))$ when $m, m' > M$, changing variables in the sum and using (A.1), we get

$$\mathbb{E}(\Sigma) \leq 3 \sum_{M, m, m' \geq 1} \mathbb{E}(g(m, S_M)g(m', S_M))(m + m').$$

We have by definition $g(m, s) = \mathbb{P}(S_{m-1} \geq -s, S_m < -s)$, so that for any $\kappa > 0$,

$$(A.5) \quad g(m, s) \leq \mathbb{P}(S_m < -s) \leq e^{-\kappa s} [\mathbb{E}(e^{-\kappa J})]^m \leq e^{-\kappa s} [\mathbb{E}(e^{-2\kappa J})]^{m/2}.$$

Since $\mathbb{E}(J) > 0$, we can choose $\kappa > 0$ such that $\beta = \mathbb{E}(e^{-2\kappa J}) < 1$, and so we get the bound

$$\mathbb{E}(\Sigma) \leq 3 \sum_{M, m, m' \geq 1} \mathbb{E} \left(e^{-2\kappa S_M} \beta^{(m+m')/2} \right) (m + m') = 3 \sum_{M \geq 1} \beta^M \sum_{m \geq 1} m^2 \beta^m.$$

Since $\beta < 1$, these two sums are finite, which achieves to prove that $|\mathcal{K}(\mathcal{T}_1)|$ has a finite second moment. \square

Lemma A.2. *As $u \rightarrow +\infty$, we have $u\mathbb{P}(h(B(\mathcal{T}_1)) \geq u) \rightarrow \mathbb{E}(J)/\mathbb{P}(J = 1)$.*

Proof. Define for simplicity $\kappa = |\mathcal{K}(\mathcal{T}_1)|$ and let $(v_k^B, k = 1, \dots, \kappa)$ be the κ killed nodes in \mathcal{T}_1 , and $(\mathcal{T}^{(k)}, k = 1, \dots, \kappa)$ be the subtrees attached to them.

$$h(\mathcal{T}_1) = \max(h(B(\mathcal{T}_1)), |v_1^B| + h(\mathcal{T}^{(1)}) - 1, \dots, |v_\kappa^B| + h(\mathcal{T}^{(\kappa)}) - 1)$$

so that

$$\mathbb{P}(h(\mathcal{T}_1) \geq u) = \mathbb{P} \left(h(B(\mathcal{T}_1)) < u \text{ and } h(\mathcal{T}^{(k)}) \geq u + 1 - |v_k^B| \text{ for some } k \in \{1, \dots, \kappa\} \right) + \mathbb{P}(h(B(\mathcal{T}_1)) \geq u).$$

Next, we observe that conditionally on $B(\mathcal{T}_1)$, the $(h(\mathcal{T}^{(k)}), k = 1, \dots, \kappa)$ are i.i.d. with common distribution $h(\mathcal{T}_1)$. Thus defining $H(u) = \mathbb{P}(h(\mathcal{T}_1) \geq u)$, we obtain

$$\mathbb{P} \left(h(\mathcal{T}^{(k)}) \geq u + 1 - |v_k^B| \text{ for some } k \in \{1, \dots, \kappa\} \mid B(\mathcal{T}_1) \right) = 1 - \prod_{k=1}^{\kappa} (1 - H(u + 1 - |v_k^B|))$$

and consequently,

$$H(u) = \mathbb{E}(Y(u); h(B(\mathcal{T}_1)) < u) + \mathbb{P}(h(B(\mathcal{T}_1)) \geq u)$$

with

$$Y(u) = 1 - \prod_{k=1}^{\kappa} (1 - H(u + 1 - |v_k^B|)).$$

It follows from (A.2) that the random variable $uY(u)\mathbb{1}_{\{h(B(\mathcal{T}_1)) < u\}}$ converges almost surely as $u \rightarrow +\infty$ to κ . If we had uniform integrability, then we would obtain

$$u\mathbb{P}(h(B(\mathcal{T}_1)) \geq u) = uH(u) - \mathbb{E}(uY(u); h(B(\mathcal{T}_1)) < u) \xrightarrow{u \rightarrow +\infty} 1 - \mathbb{E}(\kappa)$$

which would prove the result by (A.4). Thus it remains to show that the family of random variables $(uY(u)\mathbb{1}_{\{h(B(\mathcal{T}_1)) < u\}}, u \geq 0)$ is uniformly integrable: it is enough to show that $\sup_{u \geq 1} \mathbb{E}(u^2 Y(u)^2)$ is finite. Let $V^B = \max_{k=1, \dots, \kappa} |v_k^B|$: since $Y(u) \leq 1$ and Y is increasing in each $|v_k^B|$, we have

$$\mathbb{E}(Y(u)^2) \leq \mathbb{P}(V^B \geq u/2) + \mathbb{E}\left[\left(1 - (1 - H(u/2 + 1))^\kappa\right)^2\right].$$

In the event $V^B \geq u/2$, we have $N \geq 1$ where N is the number of nodes $v \in \mathcal{T}_1$ that satisfy $|v| \geq u/2$ and $\psi(v, \mathcal{T}_1) \leq 0$. Using Markov inequality, we therefore get

$$\mathbb{P}(V^B \geq u/2) \leq \mathbb{E}\left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{|v| \geq u/2, \psi(v, \mathcal{T}_1) \leq 0\}}\right) = \sum_{m \geq u/2} \mathbb{P}(S_m \leq 0).$$

Using $1 - (1 - x)^y \leq xy$ for $y \geq 0$, we get on the other hand

$$\mathbb{E}\left[\left(1 - (1 - H(u/2 + 1))^\kappa\right)^2\right] \leq H(u/2 + 1)^2 \mathbb{E}(\kappa^2)$$

so that finally,

$$u^2 \mathbb{E}(Y(u)^2) \leq u^2 \sum_{m \geq u/2} \mathbb{P}(S_m \leq 0) + (u\mathbb{P}(h(\mathcal{T}_1) \geq u/2 + 1))^2 \mathbb{E}(\kappa^2).$$

Since the probability $\mathbb{P}(S_m \leq 0)$ decays exponentially fast as $m \rightarrow +\infty$ by (A.5), the first term of the above upper bound is bounded in u . The second term being also bounded in u by (A.2) and Lemma A.1, the proof is complete. \square

Lemma A.3. *As $u \rightarrow +\infty$, we have $u\mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u) \rightarrow (\mathbb{E}(J))^2 / \mathbb{P}(J = 1)$.*

Proof. Let

$$\bar{\psi} = \sup_{v \in \mathcal{T}_1} |v|^{-2/3} |\psi(v, \mathcal{T}_1) - |v|\mathbb{E}(J)|,$$

so that for any $\varrho > 0$,

$$(A.6) \quad \mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u, \bar{\psi} \leq u^\varrho) \leq \mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u) \leq \mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u, \bar{\psi} \leq u^\varrho) + u^{-12\varrho} \mathbb{E}(\bar{\psi}^{-12}).$$

We show that $\mathbb{E}(\bar{\psi}^{-12})$ is finite. By upper bounding the supremum by the sum, we get

$$\mathbb{E}(\bar{\psi}^{-12}) \leq \sum_{m \geq 1} \frac{1}{m^8} \mathbb{E}\left(\sum_{v \in \mathcal{T}_1, |v|=m} |\psi(v, \mathcal{T}_1) - m\mathbb{E}(J)|^{12}\right) = \sum_{m \geq 1} \frac{1}{m^8} \mathbb{E}[|Y_1 + \dots + Y_m|^{12}]$$

where (Y_i) are i.i.d. centered random variables with distribution $J - \mathbb{E}(J)$ and where, in order to derive the last equality, we used the independence in \mathcal{T}_1 between the genealogical structure and the labels. The central limit theorem implies that $|Y_1 + \dots + Y_m|/m^{1/2}$ converges weakly, and since the Y_k 's are bounded, all the moments of this random variable are bounded uniformly in m , so that by uniform integrability we can write $\mathbb{E}[|Y_1 + \dots + Y_m|^{12}] \leq Cm^6$ for all $m \geq 1$ and some finite constant C , independent from m . This gives $\mathbb{E}(\bar{\psi}^{-12}) \leq C \sum_{m \geq 1} m^{-2}$ which is finite.

We now derive an upper bound on the term $\mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u, \bar{\psi} \leq u^\varrho)$ in (A.6). In the event $\{\psi^*(B(\mathcal{T}_1)) \geq u\}$, there exists $v^* \in B(\mathcal{T}_1)$ such that $u \leq \psi(v^*, \mathcal{T}_1) \leq |v^*|$. Moreover,

by definition of $\bar{\psi}$ we have $\psi(v^*, \mathcal{T}_1) \leq |v^*| \mathbb{E}(J) + |v^*|^{2/3} \bar{\psi}$, and so when both events $\{\psi^*(B(\mathcal{T}_1)) \geq u\}$ and $\{\bar{\psi} \leq u^\rho\}$ hold, there exists $v^* \in B(\mathcal{T}_1)$ such that

$$u \leq |v^*| \mathbb{E}(J) + |v^*|^{2/3} u^\rho \leq |v^*| \mathbb{E}(J) + |v^*| e^{+2/3},$$

which can be rewritten as $u \leq \phi(|v^*| \mathbb{E}(J))$ where $\phi(x) = x + (x/\mathbb{E}(J)) e^{+2/3}$ for $x \geq 0$. If ϕ^{-1} stands for its inverse, we therefore have

$$\mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u, \bar{\psi} \leq u^\rho) \leq \mathbb{P}(h(B(\mathcal{T}_1)) \geq \phi^{-1}(u)/\mathbb{E}(J))$$

so that plugging these inequalities in (A.6), we obtain

$$u \mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u) \leq \frac{C}{u^{12\rho-1}} + u \mathbb{P}(h(B(\mathcal{T}_1)) \geq \phi^{-1}(u)/\mathbb{E}(J)).$$

For $\rho > 1/12$ the first term of the above upper bound vanishes, while for $\rho < 1/3$ we have $\phi^{-1}(u) \sim u$ as $u \rightarrow +\infty$ and so the second one goes to $(\mathbb{E}(J))^2 / \mathbb{P}(J=1)$ by Lemma A.2. Thus choosing $1/12 < \rho < 1/3$ we obtain

$$\limsup_{u \rightarrow +\infty} u \mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq u) \leq \frac{(\mathbb{E}(J))^2}{\mathbb{P}(J=1)}.$$

Starting from the lower bound in (A.6) a corresponding lower bound can be proved using the same arguments, which completes the proof. \square

A.2. Various results. We now provide the proof of the various results that have been used in the proof of Theorem 1.1.

Result needed in the proof of Lemma 4.4. To complete the proof of Lemma 4.4, we need to show that there exists a finite constant $C > 0$ such that for every $u > 0$,

$$(A.7) \quad \mathbb{P}(\tau(\mathcal{T}_1) \geq u) \geq C u^{-1/2}.$$

Indeed, we have

$$\mathbb{P}(\tau(\mathcal{T}_1) \geq u) \geq \mathbb{P}(|\mathcal{T}_1| \geq u, \mathcal{K}(\mathcal{T}_1) = \emptyset) = \mathbb{P}(|\mathcal{T}_1| \geq u) - \mathbb{P}(|\mathcal{T}_1| \geq u, |\mathcal{K}(\mathcal{T}_1)| \geq 1).$$

We have $\mathbb{P}(|\mathcal{T}_1| \geq u) \geq C u^{-1/2}$ by (A.2), while

$$\mathbb{P}(|\mathcal{T}_1| \geq u, |\mathcal{K}(\mathcal{T}_1)| \geq 1) \leq \sqrt{\mathbb{P}(|\mathcal{T}_1| \geq u) \mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|)} \leq C u^{-1/4},$$

where the first inequality comes from using first Cauchy-Schwarz inequality and then Markov inequality, and the second inequality comes from (A.2) and the fact that $\mathbb{E}(|\mathcal{K}(\mathcal{T}_1)|)$ is finite by Lemma A.1. We thus get

$$\mathbb{P}(\tau(\mathcal{T}_1) \geq u) \geq C(u^{-1/2} - u^{-1/4}) \geq C u^{-1/2}$$

which concludes the proof. \square

Result needed in the proof of Lemmas 4.5 and 4.10. To complete the proof of Lemma 4.5, we need to show that there exists a finite constant $C > 0$ such that for every $p \geq 0$,

$$(A.8) \quad \mathbb{E} \left(\sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq p\}} \right) \leq C p,$$

while in the proof of Lemma 4.10 we need to prove that

$$(A.9) \quad \frac{1}{y} \mathbb{E} \left(\sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq y\}} \right) \xrightarrow{y \rightarrow +\infty} \frac{1}{\mathbb{P}(J=1)}.$$

Since (A.9) implies (A.8) we prove (A.9). We have

$$\mathbb{E} \left(\sum_{v \in B(\mathcal{T}_1)} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq y\}} \right) = \sum_{m \geq 1} \mathbb{P}(S_m \leq y, \underline{S}_m \geq 0).$$

As $m \rightarrow +\infty$, S_m/m conditionally on $\{\underline{S}_m \geq 0\}$ converges to $\mathbb{E}(J)$. One can therefore show that

$$\sum_{m \geq 1} \mathbb{P}(S_m \leq y, \underline{S}_m \geq 0) = \sum_{m=1}^{y/\mathbb{E}(J)} \mathbb{P}(\underline{S}_m \geq 0) + o(y).$$

Since $\mathbb{P}(\underline{S}_m \geq 0) \rightarrow \mathbb{P}(\underline{S}_\infty \geq 0) = \mathbb{E}(J)/\mathbb{P}(J=1)$, this gives

$$\sum_{m \geq 1} \mathbb{P}(S_m \leq y, \underline{S}_m \geq 0) = \frac{y}{\mathbb{E}(J)} \mathbb{P}(\underline{S}_\infty \geq 0) + o(y) = \frac{y}{\mathbb{P}(J=1)} + o(y)$$

which proves (A.9). \square

Results needed in the proof of Lemma 4.6. To complete the proof of Lemma 4.6, we must show that there exists a finite constant $C > 0$ such that for every $p \geq 1$ and every $\kappa, u > 0$

$$(A.10) \quad \mathbb{P} \left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq p\}} \geq \kappa \mid \tau(\mathcal{T}_1) > u \right) \leq \frac{Cp^2}{\kappa}$$

and

$$(A.11) \quad \mathbb{P} \left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}} \geq \kappa \mid \tau(\mathcal{T}_1) > u \right) \leq \frac{Cp}{\kappa}.$$

Note that (A.11) implies (A.10) by summation over p , so we only need prove (A.11). Let $N_p = \sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}}$: we have to control $\mathbb{P}(N_p \geq \kappa \mid \tau(\mathcal{T}_1) \geq u)$, and we start by writing

$$\mathbb{P}(N_p \geq \kappa \mid \tau(\mathcal{T}_1) > u) = \frac{\mathbb{P}(N_p \geq \kappa, \tau(\mathcal{T}_1) > u)}{\mathbb{P}(\tau(\mathcal{T}_1) > u)} \leq \frac{C}{\kappa} \mathbb{E}(N_p \mid |\mathcal{T}_1| > u/2)$$

using (A.7), $\tau(\mathcal{T}_1) \leq 2|\mathcal{T}_1|$ by (3.1), Markov inequality and (A.2) to derive the inequality. Conditioning on the genealogical structure leads as before to

$$\mathbb{E}(N_p \mid |\mathcal{T}_1| > u/2) = \sum_{m \geq 1} \mathbb{E}(Z_m \mid |\mathcal{T}_1| > u/2) \mathbb{P}(S_m = p)$$

and combining the two previous displays with (A.3), we end up with

$$\mathbb{P}(N \geq \kappa \mid \tau(\mathcal{T}_1) > u) \leq \frac{C}{\kappa} \sum_{m \geq 1} m \mathbb{P}(S_m = p) = \frac{C}{\kappa} \mathbb{E} \left(\sum_{m \geq 1} m \mathbb{1}_{\{S_m = p\}} \right).$$

Since S takes a geometric number of times the value p at times around $p/\mathbb{E}(J)$, the term $\mathbb{E}(\sum_{m \geq 1} m \mathbb{1}_{\{S_m = p\}})$ is of the order of p when p grows large, which concludes the proof. \square

Results needed in the proof of Lemma 4.7. To complete the proof of Lemma 4.7 we need to prove the two following results:

$$(A.12) \quad \limsup_{n \rightarrow +\infty} \mathbb{P}(\psi^*(B(\mathcal{T}_1)) < (\varepsilon^2 + \varepsilon)n \mid \psi^*(B(\mathcal{T}_1)) \geq \varepsilon n) \xrightarrow{\varepsilon \rightarrow 0} 0$$

and

$$(A.13) \quad \limsup_{n \rightarrow +\infty} \mathbb{P}(|\mathcal{T}_1| > \varepsilon n^2 \mid \psi^*(B(\mathcal{T}_1)) \geq \varepsilon n) \xrightarrow{\varepsilon \rightarrow 0} 0.$$

Note that (A.12) follows immediately from Lemma A.3. As for (A.13), we have

$$\mathbb{P}(|\mathcal{T}_1| > \varepsilon n^2 \mid \psi^*(B(\mathcal{T}_1)) \geq \varepsilon n) = \frac{\mathbb{P}(|\mathcal{T}_1| > \varepsilon n^2, \psi^*(B(\mathcal{T}_1)) \geq \varepsilon n)}{\mathbb{P}(\psi^*(B(\mathcal{T}_1)) \geq \varepsilon n)} \leq C\varepsilon n \mathbb{P}(|\mathcal{T}_1| > \varepsilon n^2)$$

where the last inequality results from Lemma A.3. Invoking (A.2) thus gives (A.13). \square

Results needed in the proof of Lemma 4.8. To complete the proof of Lemma 4.8, we must show that there exists a finite constant $C > 0$ such that for every $u, \kappa > 0$ and $p \geq 1$,

$$(A.14) \quad \mathbb{P}\left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) \leq p\}} \geq \kappa \mid \psi^*(B(\mathcal{T}_1)) > u\right) \leq \frac{Cp^2}{\kappa}$$

and

$$(A.15) \quad \mathbb{P}\left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}} \geq \kappa \mid \psi^*(B(\mathcal{T}_1)) > u\right) \leq \frac{Cp}{\kappa}.$$

As for (A.10) and (A.11) we only need prove (A.15): combining $\psi^*(B(\mathcal{T}_1)) \leq h(\mathcal{T}_1)$, Lemma A.3 and (A.2), we obtain

$$\mathbb{P}\left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}} \geq \kappa \mid \psi^*(B(\mathcal{T}_1)) > u\right) \leq C \mathbb{P}\left(\sum_{v \in \mathcal{T}_1} \mathbb{1}_{\{\psi(v, \mathcal{T}_1) = p\}} \geq \kappa \mid h(\mathcal{T}_1) > u\right).$$

From there, (A.15) can be proved by repeating verbatim the proof of (A.11) with the following caveat: one needs to replace the conditioning on $|\mathcal{T}_1|$ by a conditioning on $h(\mathcal{T}_1)$, and thus to use the second bound in (A.3) instead of the first one. \square

Result needed in the proof of Lemma 4.9. To complete the proof of Lemma 4.9, we need to prove that

$$(A.16) \quad \sup_{u \geq 0} \mathbb{E}(|\mathcal{K}(\mathcal{T}_1)| \mid \psi^*(B(\mathcal{T}_1)) > u) < +\infty.$$

The same arguments as in the previous proof apply and show that

$$\mathbb{E}(|\mathcal{K}(\mathcal{T}_1)| \mid \psi^*(B(\mathcal{T}_1)) > u) \leq C \mathbb{E}(|\mathcal{K}(\mathcal{T}_1)| \mid h(\mathcal{T}_1) > u).$$

Conditioning on the genealogical structure $(Z_m, m \geq 1)$, we get

$$\mathbb{E}(|\mathcal{K}(\mathcal{T}_1)| \mid h(\mathcal{T}_1) > u) = \sum_{m \geq 1} \mathbb{E}(Z_m \mid h(\mathcal{T}_1) > u) \mathbb{P}(S_m < 0, \underline{S}_{m-1} \geq 0) \leq C \sum_{m \geq 1} m \mathbb{P}(S_m < 0)$$

using (A.3) for the last inequality. Since $\mathbb{P}(S_m < 0)$ decays exponentially fast in m by (A.5), the sum $\sum_{m \geq 1} m \mathbb{P}(S_m < 0)$ is finite, which gives (A.16). \square

Result needed in the proof of Lemma 4.10. To complete the proof of Lemma 4.10 we need to prove that the constant C^* defined there is finite. Let $N(y)$ be the number of nodes in $B(\mathcal{T}_1)$ with label $\leq y$: then going back to the definition of C^* , we see that we have to prove that $\sup_y (\text{Var}(\mathcal{S}(N(y))) / y^3)$ is finite, where $\text{Var}(Y)$ denotes the variance of a real valued random variable Y . Thanks to (3.2), we only have to show that $\sup_y (\text{Var}(N(y)) / y^3)$ is finite. Further, proceeding as in the rest of this section, we can show that $N(y)$ behaves like the number of nodes in \mathcal{T}_1 at depth $\leq y/\mathbb{E}(J)$, and in particular $\text{Var}(N(y))$ is of the order of $\text{Var}(Z_1 + \dots + Z_{y/\mathbb{E}(J)})$. Thus in order to prove that $C^* < +\infty$, we only have to prove that $\text{Var}(Z_1 + \dots + Z_n)$ grows at most like n^3 . Let

$$v_n = \mathbb{E}\left((Z_{n+1} - 1) \sum_{k=1}^n (Z_k - 1)\right).$$

Then conditioning on $(Z_k, k \geq n)$, we obtain

$$v_n = \mathbb{E} \left((Z_n - 1) \sum_{k=1}^n (Z_k - 1) \right)$$

and so (A.1) gives $v_n = 2n + v_{n-1}$. In particular, v_n grows quadratically. On the other hand, we have

$$\text{Var}(Z_1 + \dots + Z_{n+1}) = 2(n+1) + 2v_n + \text{Var}(Z_1 + \dots + Z_n),$$

and since v_n grows quadratically in n , this implies that $\text{Var}(Z_1 + \dots + Z_n)$ grows like n^3 , which proves the result. \square

REFERENCES

- [1] L. Addario-Berry and N. Broutin. Total progeny in killed branching random walk. *Probab. Theory Related Fields*, 151(1-2):265–295, 2011.
- [2] David Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.
- [3] Jean Bertoin. *Lévy processes*, volume 121 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1996.
- [4] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.
- [5] Jose Blanchet and Xinyun Chen. Continuous-time modeling of bid-ask spread and price dynamics in limit order books. arXiv:1310.1103.
- [6] Mark Brown, Erol A. Peköz, and Sheldon M. Ross. Some results for skip-free random walk. *Probab. Engrg. Inform. Sci.*, 24(4):491–507, 2010.
- [7] Rama Cont, Sasha Stoikov, and Rishi Talreja. A stochastic model for order book dynamics. *Oper. Res.*, 58(3):549–563, 2010.
- [8] Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.
- [9] Rick Durrett, Harry Kesten, and Ed Waymire. On weighted heights of random trees. *J. Theoret. Probab.*, 4(1):223–237, 1991.
- [10] A. Garèche, G. Disdier, J. Kockelkoren, and J.-P. Bouchaud. Fokker-planck description for the queue dynamics of large tick stocks. *Phys. Rev. E*, 88:032809, Sep 2013.
- [11] R. K. Gettoor. Excursions of a Markov process. *Ann. Probab.*, 7(2):244–266, 1979.
- [12] Matin David Gould, Mason Alexander Porter, Stacy Williams, Mark McDonald, Daniel Fenn, and Sam Howison. Limit order books. arXiv:1012.0349.
- [13] Ulrich Horst and Michael Paulsen. A law of large numbers for limit order books. Preprint available on the authors' webpage.
- [14] Jean Jacod and Albert N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.
- [15] Svante Janson. Random cutting and records in deterministic and random trees. *Random Structures Algorithms*, 29(2):139–179, 2006.
- [16] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [17] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [18] Harry Kesten. A limit theorem for weighted branching process trees. In *The Dynkin Festschrift*, volume 34 of *Progr. Probab.*, pages 153–166. Birkhäuser Boston, Boston, MA, 1994.
- [19] Andrei Kirilenko, Richard B. Sowers, and Xiangqian Meng. A multiscale model of high-frequency trading. *Algorithmic Finance*, 2(1):59–98, 01 2013.
- [20] Peter Lakner, Josh Reed, and Sasha Stoikov. High frequency asymptotics for the limit order book. Preprint available on the authors' webpage.
- [21] Amaury Lambert and Florian Simatos. Asymptotic behavior of local times of compound Poisson processes with drift in the infinite variance case. To appear in *Journal of Theoretical Probability*.
- [22] Amaury Lambert, Florian Simatos, and Bert Zwart. Scaling limits via excursion theory: Interplay between Crump-Mode-Jagers branching processes and Processor-Sharing queues. *Ann. Appl. Probab.*, 23(6):2357–2381, 2013.

- [23] Jean-François Le Gall and Grégory Miermont. Scaling limits of random trees and planar maps. In *Probability and statistical physics in two and more dimensions*, volume 15 of *Clay Math. Proc.*, pages 155–211. Amer. Math. Soc., Providence, RI, 2012.
- [24] Etienne Pardoux and Anton Wakolbinger. From exploration paths to mass excursions – variations on a theme of Ray and Knight. In J. Blath, P. Imkeller, and S. Roelly, editors, *Surveys in Stochastic Processes*, EMS Series of Congress Reports, pages 87–106, 2011.
- [25] Sylvie Roelly-Coppoletta. A criterion of convergence of measure-valued processes: application to measure branching processes. *Stochastics*, 17(1-2):43–65, 1986.
- [26] Florian Simatos. Coupling limit order books and branching random walks. To appear in *Journal of Applied Probability*.
- [27] Mathilde Weill. Regenerative real trees. *Ann. Probab.*, 35(6):2091–2121, 2007.
- [28] Elena Yudovina. A simple model of a limit order book. arXiv:1205.7017.