

SCALING LIMITS VIA EXCURSION THEORY: INTERPLAY BETWEEN CRUMP-MODE-JAGERS BRANCHING PROCESSES AND PROCESSOR-SHARING QUEUES

BY AMAURY LAMBERT^{*}

UPMC Univ Paris 06

BY FLORIAN SIMATOS[†]

CWI and EURANDOM

BY BERT ZWART[‡]

CWI, EURANDOM, VU University Amsterdam and GeorgiaTech

We study the convergence of the $M/G/1$ Processor-Sharing queue length process in the heavy traffic regime, in the finite variance case. To do so, we combine results pertaining to Lévy processes, branching processes and queueing theory. These results yield the convergence of long excursions of the queue length processes, towards excursions obtained from those of some reflected Brownian motion with drift, after taking the image of their local time process by the Lamperti transformation. We also show, via excursion theoretic arguments, that this entails the convergence of the entire processes to some (other) reflected Brownian motion with drift. Along the way, we prove various invariance principles for homogeneous, binary Crump–Mode–Jagers processes. In the last section we discuss potential implications of the state space collapse property, well-known in the queueing literature, to branching processes.

CONTENTS

1	Introduction	2
2	Notation and preliminary results	5
3	CMJ branching processes, PS queues and Lévy processes	9
4	Heavy traffic of PS via excursion theory	13
5	Identification of the limit and general initial condition	19

^{*}The research of A. Lambert is funded by project MANEGE 09-BLAN-0215 of ANR (French national research agency).

[†]The research of F. Simatos and B. Zwart is sponsored by an NWO-VIDI grant.

AMS 2000 subject classifications: Primary 60F17, ; secondary 60J80, 60J55, 60K25

Keywords and phrases: scaling limit, excursion theory, Processor-Sharing queue, local time process of Lévy processes, Crump-Mode-Jagers branching processes

6	Discussion	23
A	Proof of Lemma 2.5	25
	References	27
	Author's addresses	29

1. Introduction. The standard machinery to show weak convergence of stochastic processes consists in proving tightness and characterizing accumulation points. Probably the most common technique to characterize accumulation points is to show that finite-dimensional distributions converge, but as Jacod and Shiryaev [15] point out, this is “very often [...] a very difficult (or simply impossible) task to accomplish”. In the present work, motivated by the Processor-Sharing (PS) queue length process, we develop new ideas to characterize limit points of a sequence of regenerative processes. The basic idea is to show that the convergence of suitably conditioned excursions implies the convergence of the full processes. Our starting point to control excursions is the Lamperti transformation that links excursions of the PS queue to Crump-Mode-Jagers (CMJ) branching processes. Further, control on CMJ processes comes from a recent result of Lambert [22] that relates them to Lévy processes via local times.

The Processor-Sharing Queue. The PS queue is a single-server queue in which the server splits its service capacity equally among all the users present. For instance, if the server has a service capacity c and if there are exactly $q \geq 1$ customers in the queue during the time interval $[t, t + h]$ (in particular there is no arrival or departure), then the residual service requirement of each customer is decreased by ch/q during this time interval while the total workload is decreased by ch .

Crump-Mode-Jagers Branching Processes. A CMJ process is a stochastic process counting the size of a population where individuals give birth to independent copies of themselves. It is defined through a pair (V, ξ) of possibly dependent random variables, where $V > 0$ is a real valued random variable and ξ is a point process on $(0, \infty)$ (in particular, its atoms have integer-valued weights). Each individual a of the branching process is given an independent copy (V_a, ξ_a) of (V, ξ) : if the individual a is born at time B_a , then at time $B_a \leq t \leq B_a + V_a$ she gives birth to $\xi_a(\{t - B_a\})$ i.i.d. copies of herself, V_a therefore being seen as her life length. A CMJ process is called binary and homogeneous when ξ is a Poisson process independent from V . Lambert [22] has shown that a binary, homogeneous CMJ process which is in addition critical or subcritical is the local time process of a suitable spectrally positive Lévy process.

The Lamperti Transformation. Connections between branching processes and queues have been known for a long time. Kendall [18] is usually referred to as one of the earliest publications in this area. Concerning the PS queue, Kitayev and Yashkov [20] have proved that a busy cycle of the PS queue length process becomes a CMJ process after a suitable time-change. This time-change transformation is the same one that links continuous-state space branching processes and Lévy processes and is called Lamperti transformation in the branching literature, see Lamperti [23] or Caballero et al. [7] for a more recent treatment. This connection between the PS queue and CMJ processes has been used to establish results on the stationary behavior of the PS queue, see for instance Grishchkin [11]. In this paper we make a deeper use of this connection, since we exploit it to study the entire trajectories of the processes.

The connections between CMJ processes, Lévy processes and PS queues lead to a natural proof of the weak convergence of CMJ processes. On the one hand, we can prove tightness of such processes by transferring, via the Lamperti transformation, a result in queueing theory on the departure process of queues with a symmetric service discipline. On the other hand, exploiting the fact that subcritical, binary and homogeneous CMJ processes are local time processes of suitable Lévy processes makes it possible to characterize accumulation points.

Using continuity properties of the Lamperti transformation, much in the spirit of those established by Helland [14], and the connection between CMJ processes and the PS queue, the convergence of suitably renormalized CMJ processes implies that excursions of the $(M/G/1)$ PS queue length processes converge. Thus convergence of excursions of the PS queue length process comes quite naturally by combining different results from queueing theory and the theory of Lévy and CMJ processes. Besides the original combination of these various results, the main methodological contribution of the present work is to show that from there, one can conclude that the whole PS queue length processes converge.

Weak convergence of CMJ processes. Binary, homogeneous CMJ processes considered in the present paper can be seen as branching processes and also as local time processes. Since Lamperti [24], we have a complete characterization of the possible asymptotic behaviors of branching processes in discrete time. Grimvall [10] improved Lamperti's results by proving tightness and hence weak convergence, see also Chapter 9 in Ethier and Kurtz [9] for another proof of Grimvall's results using time-change arguments. In

the continuous-time setting, the Markovian case has been studied by Helland [14]. Outside the Markovian case we are only aware of two papers by Sagitov [31, 30], that establish convergence of the finite-dimensional distributions for some particular CMJ that are not homogeneous.

As for convergence of local time processes, there is a wealth of literature studying the convergence of local time processes associated to random walks converging to Brownian motion. One of the earliest publication in this domain is Knight [21], see also Borodin [4, 5], Perkins [27] and references therein. The problem of finding sharp convergence rates has been the focus of intense activity, see for instance the introduction of Csörgő and Révész [8] for references. On the other hand, the question of the convergence of local time processes associated to compound Poisson processes (which is another natural way to approximate a Brownian motion) has been comparatively very little studied. In this context, Khoshnevisan [19] derived sharp convergence rates using embedding techniques requiring bounded fourth moments. In that respect, some of the results of the present paper seem to be new. Under a second moment assumption on lifetimes, Theorem 4.8 shows the weak convergence of suitably renormalized homogeneous, binary CMJ processes when started from one individual and conditioned by their total offspring. Theorem 5.4 states the weak convergence of these CMJ processes when started with large initial condition to the Feller diffusion. In this last setting, the choice of the initial condition turns out to be quite subtle, and is discussed in Section 6.

Heavy-Traffic of the PS Queue. The heavy traffic limit of the PS queue has been investigated by Gromoll [12]. Extending the state space collapse framework developed by Bramson [6] and Williams [33], he proved convergence of the measure-valued descriptors of the $G/G/1$ -PS queue, assuming that service requirements have bounded fourth moments, towards some measure-valued diffusion process. In the present paper we assume Poisson arrivals, i.e., we study the $M/G/1$ -PS queue, but we relax the moment assumption and prove convergence of the queue length process under a minimal second moment assumption. More specifically, Theorem 4.1 shows the convergence of these (suitably renormalized) queue length processes to the regenerative process whose excursions are obtained from those of some reflected Brownian motion with drift after taking the image of their local time process by the Lamperti transformation. Theorem 5.6 shows that this process actually is another reflected Brownian motion with drift.

We also believe that our method can be used to study the case of service requirements with infinite variance, where state space collapse cannot be

used since the workload and queue length processes have different orders of magnitude. To the very least, although so far there was no conjecture to this open problem, our method clearly suggests a candidate for the heavy traffic limit of the PS queue length process in the infinite variance case. Following the arguments in the previous paragraphs, this limit should be a regenerative process whose excursions away from 0 are obtained from those of some reflected, spectrally positive Lévy process by taking first their local time process and then applying Lamperti transformation.

Organization of the Paper. In Section 2 we introduce general notation and state preliminary results. In Section 3, we explain the connections between CMJ processes, PS queues and Lévy processes. We also introduce the processes studied throughout the rest of the paper. Section 4 is devoted to the proof of the main result of the paper, Theorem 4.1, which states the convergence of the PS queue length process towards a process that we define through its excursion measure. In Section 5 we extend this result by explicitly identifying the limiting process as being a(nother) reflected Brownian motion with drift and by considering a general initial condition. Finally, in Section 6 we make some comments about continuity properties of local time processes and possible implications of the state-space collapse property to branching processes.

Acknowledgements. We would like to thank the two anonymous referees who have contributed to significant improvements both in the presentation and content of the paper. We are especially grateful to one of the two referees who suggested the current proofs of Lemma 4.3 and Theorem 4.8 that lead to substantial simplifications.

2. Notation and preliminary results. Let D , resp. D_+ , be the set of càdlàg functions from $[0, \infty)$ to \mathbb{R} , resp. to $[0, \infty)$. For $f \in D$ and $m \geq 0$ let $\|f\|_m = \sup_{[0, m]} |f|$ and $\|f\|_\infty = \sup |f|$. We will endow D with the topology of uniform convergence on compact sets, i.e., we will write $f_n \rightarrow f$ for functions $f_n, f \in D$ if $\|f_n - f\|_m \rightarrow 0$ as $n \rightarrow \infty$ for every $m \geq 0$. The space D is more naturally endowed with the Skorohod J_1 topology, see for instance Billingsley [2], but since the Skorohod topology relativized to the space of continuous functions coincides with the topology of uniform convergence on compact sets there, this latter topology is enough for the purpose of the present paper, whenever considering sequences with continuous limit points.

If $f \in D$, we call *local time process* of f a Borel function $(L(a, t), a, t \geq 0)$

which satisfies

$$(1) \quad \int_0^t \phi(f(s))ds = \int_0^\infty L(a, t)\phi(a)da$$

for any $t \geq 0$ and any continuous function ϕ with compact support included in $[0, \infty)$. The local time process of an arbitrary function $f \in D$ may not exist, but when it does it is unique (up to an almost everywhere modification). In the sequel we will only consider local time processes associated to spectrally positive Lévy processes which either have infinite variation or negative drift. These local time processes are known to exist, see for instance Bertoin [1].

If $f \in D$ we define \underline{f} , the function f reflected above its past infimum, via

$$\underline{f}(t) = f(t) - \min \left(0, \inf_{0 \leq s \leq t} f(s) \right), \quad t \geq 1.$$

It is well-known that this transformation induces a continuous map, i.e., if $f_n, f \in D$ are such that $f_n \rightarrow f$, then $\underline{f_n} \rightarrow \underline{f}$.

For $f \in D$, let $\Delta f(t) = f(t) - f(t-)$ for $t > 0$, and $T_f = \inf\{t > 0 : f(t) = 0\}$ be the first time after time 0 at which f visits 0, with $T_f = \infty$ if f never visits 0 in $(0, \infty)$. We see it as a map $T : D \rightarrow [0, \infty]$ and we sometimes write $T(f)$ for T_f . In general this map is not continuous but we have the following result.

LEMMA 2.1. *If $f_n, f \in D$, $f_n \rightarrow f$ and f is continuous, then $T_f \leq \liminf_n T_{f_n}$.*

PROOF. Let $\tau = \liminf_n T_{f_n}$ and $(u(n))$ such that $T_{f_{u(n)}} \rightarrow \tau$. Since $f_n \rightarrow f$ and f is continuous, we obtain $f_{u(n)}(T_{f_{u(n)}}) \rightarrow f(\tau)$, hence $f(\tau) = 0$ which proves the result. \square

From now on \Rightarrow denotes weak convergence. When considering random vectors, we consider convergence in the product topology. The previous lemma has the following consequence.

COROLLARY 2.2. *If X_n, X are stochastic processes such that $X_n \Rightarrow X$, X is continuous and $T_{X_n} \Rightarrow T_X$, then $(X_n, T_{X_n}) \Rightarrow (X, T_X)$.*

PROOF. The sequence (X_n, T_{X_n}) is tight: let (X', T') be any accumulation point, so that X' is equal in distribution to X and T' to T_X . We show

that $T' = T_{X'}$, which will show that (X', T') is equal in distribution to (X, T_X) and will prove the result. Assume without loss of generality that $(X_n, T_{X_n}) \Rightarrow (X', T')$: the continuous mapping theorem and Lemma 2.1 imply that $T_{X'} \leq T'$. But since they are equal in distribution they must be equal almost surely, hence the result. \square

Stopping and shift operators. For $t \geq 0$ let σ_t and θ_t be the stopping and shift operators, respectively: for $f \in D$ and $t \geq 0$, $\sigma_t f = f(\cdot \wedge t)$ and $\theta_t f = f(\cdot + t)$. Note also for simplicity $\sigma = \sigma_T$ and $\theta = \theta_T$, i.e., $\sigma f = \sigma_{T(f)} f$ and $\theta f = \theta_{T(f)} f$. Formally, θ is only well-defined if $T(f)$ is finite, and in the rest of the paper we will only apply the map θ to such functions.

LEMMA 2.3. *If $f_n, f \in D$ and $t_n, t \geq 0$ are such that $f_n \rightarrow f$, f is continuous and $t_n \rightarrow t$, then $\theta_{t_n} f_n \rightarrow \theta_t f$ and $\sigma_{t_n} f_n \rightarrow \sigma_t f$.*

PROOF. Let w be the modulus of continuity of f , defined for $m, \varepsilon > 0$ by $w_m(\varepsilon) = \sup\{|f(t) - f(s)| : 0 \leq s, t \leq m \text{ and } |t - s| \leq \varepsilon\}$. Since f is continuous we have $w_m(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$, for any $m \geq 0$. Let $\bar{t} = \sup_{n \geq 1} t_n$: for any $0 \leq s \leq m$, we have

$$\begin{aligned} |\theta_{t_n} f_n(s) - \theta_t f(s)| &\leq |f_n(s + t_n) - f(s + t_n)| + |f(s + t_n) - f(s + t)| \\ &\leq \|f_n - f\|_{m+\bar{t}} + w_{m+\bar{t}}(|t_n - t|) \end{aligned}$$

and similarly, $|\sigma_{t_n} f_n(s) - \sigma_t f(s)| \leq \|f_n - f\|_m + w_m(|t_n - t|)$. These upper bounds are uniform in $s \leq m$, and since $f_n \rightarrow f$ and $t_n \rightarrow t$, letting $n \rightarrow +\infty$ gives the result. \square

In the sequel we say that a sequence (X_n) is C-tight if it is tight and any accumulation point is almost surely continuous. We will use several times that if (X_n) and (Y_n) are two C-tight sequences defined on the same probability space, then the sequence $(X_n + Y_n)$ is also C-tight, see for instance Corollary VI.3.33 in Jacod and Shiryaev [15].

COROLLARY 2.4. *If (X_n) is a C-tight sequence of processes and (κ_n) is a tight sequence of positive random variables, then $(\sigma_{\kappa_n} X_n)$ and $(\theta_{\kappa_n} X_n)$ are C-tight.*

PROOF. Let $(u(n))$ be a subsequence, we must find $(v(n))$ a subsequence of $(u(n))$ such that $(\sigma_{\kappa_{v(n)}} X_{v(n)})$ and $(\theta_{\kappa_{v(n)}} X_{v(n)})$ converge weakly to a continuous process. The sequence (X_n, κ_n) being tight, there exists $(v(n))$ a subsequence of $(u(n))$ such that $(X_{v(n)}, \kappa_{v(n)})$ converges weakly to some (X, κ) , with X a continuous process. Thus $\sigma_{\kappa_{v(n)}} X_n \Rightarrow \sigma_\kappa X$ and $\theta_{\kappa_{v(n)}} X_n \Rightarrow$

$\theta_\kappa X$ by Lemma 2.3 together with the continuous mapping theorem, hence the result. \square

Excursions. A function $e \in D_+$ will be called an excursion if $e(t) = 0$ for some $t > 0$ implies $e(u) = 0$ for all $u \geq t$. Observe that excursions are allowed to start at 0. Write \mathcal{E} for the set of excursions with finite length T_e . Let also $\mathcal{E}' \subset \mathcal{E}$ be the subset of excursions $e \in \mathcal{E}$ such that $\int(1/e)$ is finite, where from now on we write $\int_a^b f = \int_a^b f(t)dt$ and $\int f = \int_0^{T_f} f$, for $f \in D$. When dealing with excursions we will use the canonical notation for stochastic processes and write ϵ for the canonical map.

For $f \in D$ and $\varepsilon > 0$, let $e_\varepsilon(f)$ be the first excursion e of f away from 0 that satisfies $T_e > \varepsilon$, and let $g_\varepsilon(f) < d_\varepsilon(f)$ be its left and right endpoints. Note that there need not be such an excursion, but in the rest of the paper we will only apply the maps e_ε to functions f such that $e_\varepsilon(f)$ is well-defined for every $\varepsilon > 0$. Also, note that by definition, we have $e_\varepsilon = \sigma \circ \theta_{g_\varepsilon}$ in the sense that for any $f \in D$,

$$e_\varepsilon(f) = (\sigma \circ \theta_{g_\varepsilon(f)})(f) = (\sigma_{d_\varepsilon(f)-g_\varepsilon(f)} \circ \theta_{g_\varepsilon(f)})(f).$$

Lamperti transformation. We will call *Lamperti transformation* the map $\mathcal{L} : \mathcal{E} \rightarrow \mathcal{E}$ that to an excursion $f \in \mathcal{E}$ associates the excursion $h \in \mathcal{E}$ defined by $h(\int_0^t f) = f(t)$ for all $t \geq 0$. More specifically, if κ is the inverse of the strictly increasing, continuous function $t \mapsto \int_0^t f$ on $[0, \int f]$, then $\mathcal{L}(f) = f \circ \kappa$ on $[0, \int f]$ and 0 elsewhere. In particular, $(T \circ \mathcal{L})(f) = \int f$.

The inverse Lamperti transformation \mathcal{L}^{-1} also plays a crucial role. By definition, $\mathcal{L}^{-1}(f)$ is the solution in \mathcal{E} , when it exists and is unique, to the equation $h(t) = f(\int_0^t h)$, $t \geq 0$, where h is the unknown function. Existence and uniqueness to such equations are studied in Chapter 6 of Ethier and Kurtz [9]. Because we consider excursions which may start at 0, we cannot directly invoke Theorem 1.1 there, but an inspection of the proof reveals that it can be adapted to show that $\mathcal{L}^{-1}(f)$ is well-defined for $f \in \mathcal{E}'$. In this case, we have $\mathcal{L}^{-1}(f) = f \circ \pi$ on $[0, \int(1/f)]$, and 0 otherwise, with π the inverse of the strictly increasing, continuous function $t \mapsto \int_0^t (1/f)$ on $[0, \int(1/f)]$. In particular, $(T \circ \mathcal{L}^{-1})(f) = \int(1/f)$.

We will need the following results on \mathcal{L} and \mathcal{L}^{-1} , which are closely related to results by Helland [14] or Ethier and Kurtz [9, Chapter 6]. There are nonetheless significant differences and for completeness, we provide the proof of the following lemma in the appendix.

LEMMA 2.5. *Let X_n, X be random elements of \mathcal{E} such that the sequence (X_n) is C -tight and the sequence (T_{X_n}) is tight.*

If $X_n \Rightarrow X$ then $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$.

If $\mathbb{P}(X_n \in \mathcal{E}') = 1$, then the sequence $(\mathcal{L}^{-1}(X_n))$ is C -tight.

If $X_n \Rightarrow X$ and $\mathbb{P}(X_n \in \mathcal{E}') = \mathbb{P}(X \in \mathcal{E}') = 1$, then $\mathcal{L}^{-1}(X_n) \Rightarrow \mathcal{L}^{-1}(X)$.

3. CMJ branching processes, PS queues and Lévy processes.

Recall from the introduction that a Crump-Mode-Jagers (CMJ) process is a stochastic process with non-negative integer values counting the size of a population where individuals give birth to independent copies of themselves, and that Processor-Sharing (PS) is the service discipline where the server splits its service capacity equally among all users present in the queue at any time.

In the sequel, we will only consider homogeneous and binary CMJ processes, where individuals give birth to a single offspring at times of a Poisson process independent of their life length. With the notation of the introduction, ξ is a Poisson process independent of V . Similarly, we will only consider $M/G/1$ -PS queues, i.e., PS queues with Poisson arrivals and i.i.d. service requirements. Henceforth, CMJ will stand for homogeneous and binary CMJ, and PS for $M/G/1$ -PS.

In particular, thanks to the memoryless property of the exponential random variable, both a CMJ process and a PS queue can be described by a Markov process living in the state space $\mathcal{S} = \bigcup_{n \geq 0} (0, \infty)^n$, the set of finite sequences of positive real numbers. For a CMJ process the Markovian descriptor keeps track of the residual life lengths of the individuals alive; for a PS queue, it keeps track of the residual service requirements of the customers present in the queue. Although for the PS queue we will sometimes refer to this Markov process, we will actually only consider marginals of it, namely the workload process (corresponding to its total mass) and the queue length process (corresponding to the cardinal of its support). Thus although studying non-Markovian processes, we avoid the framework of measure valued processes.

Simple Facts about the PS Queue. Let q be the queue length process of a PS queue with unit service capacity, arrival rate λ and service distribution S . The workload process is the process keeping track of the total amount of work in the system, defined as the sum over all the customers of their residual service requirements. Since we assume Poisson arrivals, the workload process is a compensated compound Poisson process with drift -1 and Lévy measure $\lambda \mathbb{P}(S \in \cdot)$, reflected above its past infimum.

Set $\rho := \lambda \mathbb{E}(S)$ the load and assume $\rho < 1$ (subcritical case). Let S^* be the random variable with density $\mathbb{P}(S \geq \cdot)/\mathbb{E}(S)$ with respect to Lebesgue measure. It is sometimes called the forward recurrence time of S and has mean

$\mathbb{E}(S^*) = \mathbb{E}(S^2)/(2\mathbb{E}(S))$. The assumption $\rho < 1$ is equivalent to assuming that the Markov process describing the PS queue has a unique invariant distribution ν^* on \mathcal{S} . In that case, the invariant distribution is characterized by a geometrically distributed number of customers with parameter ρ , and i.i.d. residual service times with common distribution S^* , see, e.g., Robert [29, Proposition 7.13].

Connection between PS Queues, CMJ Processes and Lévy processes. In the following statement, q is the above PS queue and \mathbb{P}^χ is its law started at $\chi \in \mathcal{S}$. The following result is known since at least Kitayev and Yashkov [20], see also Chapter 7.3 in Robert [29].

THEOREM 3.1 (Connection between PS Queues and CMJ Processes). *Let $\chi = (\chi_i, 1 \leq i \leq k) \in \mathcal{S}$. The process $\mathcal{L}^{-1}(q)$ under \mathbb{P}^χ is a CMJ process starting with k ancestors, with birth rate λ and life length distribution S , except for the ancestors who have deterministic life lengths given by χ .*

Thus we can see σq as the time change of a CMJ process, since $\sigma q = \mathcal{L}(z)$ with $z = \mathcal{L}^{-1}(q)$ which by the above is a CMJ process. Further, since 0 is a regeneration point of q , every excursion of q away from 0 can be seen as the time change of a CMJ process started with one individual with life length distributed as S .

The following result can be found in Lambert [22]. The jumping contour process of a homogeneous, binary CMJ tree starting from one progenitor is the key object underlying this result. It is defined in Lambert [22], to which the reader is referred for more details.

In the following statement, x denotes a spectrally positive Lévy process starting from $\delta > 0$, with drift -1 , Lévy measure $\lambda\mathbb{P}(S \in \cdot)$ and local time process $(\ell(a, t), a, t \geq 0)$, as defined in (1). Note that in this case, $\ell(a, t)$ is also the number of times when x has taken the value a before time t .

THEOREM 3.2 (Connection between CMJ and Lévy Processes). *The process $(\ell(a, T_x), a \geq 0)$ is a CMJ process with birth rate λ and life length distribution S , started with one progenitor with life length δ .*

Scaling Near the Critical Point. For each integer $n \geq 1$, consider some $\lambda_n > 0$ and a positive random variable S_n , with forward recurrence time S_n^* . Let q_n denote the queue length process of a PS queue with arrival rate λ_n and service distribution S_n . Let $z_n = \mathcal{L}^{-1}(q_n)$, which according to Theorem 3.1 is a CMJ process with birth rate λ_n and life length distribution S_n . Last, let x_n be a compensated compound Poisson process with drift -1 and

Lévy measure $\lambda_n \mathbb{P}(S_n \in \cdot)$. Let $(\ell_n(a, t), a, t \geq 0)$ be its local time process, so that by Theorem 3.2, z_n is equal in distribution to $(\ell_n(a, T_{x_n}), a \geq 0)$. Also, x_n is equal in distribution to the workload process corresponding to q_n (with suitable initial conditions); in particular, the zero sets of q_n and x_n have the same distribution.

We restrict our attention to the subcritical case, namely, we assume that for each $n \geq 1$ the load $\rho_n := \lambda_n \mathbb{E}(S_n)$ satisfies $\rho_n < 1$. This assumption means that all hitting times of 0 by q_n , z_n and x_n with deterministic initial states (x_n starting in $(0, \infty)$) are almost surely finite and have finite expectations. We consider the following scaling near the critical point: in the sequel we assume that there exist finite and strictly positive real numbers λ, β and α such that

$$(2) \quad \lim_{n \rightarrow +\infty} \lambda_n = \lambda, \quad \lim_{n \rightarrow +\infty} \frac{\lambda_n}{2} \mathbb{E}(S_n^2) = \beta \quad \text{and} \quad \lim_{n \rightarrow +\infty} n(1 - \rho_n) = \alpha.$$

These three assumptions imply that $\mathbb{E}(S_n^*) \rightarrow \beta$. We are interested in the rescaled processes Q_n , Z_n and X_n defined by

$$Q_n(t) = \frac{q_n(n^2 t)}{n}, \quad Z_n(t) = \frac{z_n(nt)}{n} \quad \text{and} \quad X_n(t) = \frac{x_n(n^2 t)}{n}, \quad n \geq 1, t \geq 0.$$

Let also

$$L_n(a, t) = \frac{\ell_n(na, n^2 t)}{n}, \quad n \geq 1, a, t \geq 0.$$

Then L_n is the local time process of X_n , i.e., it satisfies

$$\int_0^t \phi(X_n(s)) ds = \int_0^\infty \phi(a) L_n(a, t) da.$$

By the Lévy-Khintchine formula, X_n is a Lévy process with Laplace exponent $\Psi_n(u) = nu - n^2 \lambda_n \mathbb{E}(1 - e^{-u S_n/n})$. In view of (2), standard arguments show that $\Psi_n(u) \rightarrow \alpha u + \beta u^2$ for any $u \geq 0$. As a consequence, see for instance Kallenberg [16], (X_n) converges in distribution to a drifted Brownian motion with drift $-\alpha$ and Gaussian coefficient 2β , which we write in the sequel X and whose local time process is denoted $(L(a, t), a, t \geq 0)$.

Notation for the Initial Condition. For any $\chi = (\chi_1, \dots, \chi_k) \in \mathcal{S}$, when q_n and z_n are started with $k \geq 1$ customers/individuals with residual service times/life lengths given by χ , the probability measure is denoted \mathbb{P}_n^χ . The law of the PS queue started empty will be denoted \mathbb{P}_n^\emptyset . We will use the following notation for random initial conditions.

When q_n and z_n are started with *one* individual with residual life length distributed as S_n , the law will simply be denoted \mathbb{P}_n . When there are initially ζ_n individuals with i.i.d. residual life lengths distributed as S_n^* , we will use the symbol $\mathbb{P}_n^{\zeta_n^*}$. When the initial condition is ν_n^* (a geometric number with parameter ρ_n of individuals with i.i.d. life lengths distributed as S_n^*), we will merely use the symbol \mathbb{P}_n^* . Note that Q_n under \mathbb{P}_n^* is a stationary process.

The probability measure for the Lévy processes is denoted \mathbf{P}_n^a when X_n itself is started at $a \in \mathbb{R}$. When X_n starts at a random initial value distributed as S_n/n we write \mathbf{P}_n . Finally, \mathbf{P}^a denotes the law of X started at a .

The scalings in time and space leading to Q_n , Z_n and X_n have been chosen in order to preserve the defining relationships between q_n , z_n and x_n .

LEMMA 3.3. *We have $Z_n = \mathcal{L}^{-1}(Q_n)$ and $\sigma Q_n = \mathcal{L}(Z_n)$, in particular $T_{Q_n} = \int Z_n$. Moreover, for any $\delta > 0$, Z_n under \mathbb{P}_n^χ with $\chi = (n\delta) \in \mathcal{S}$ is equal in distribution to $(L_n(a, T_{X_n}), a \geq 0)$ under \mathbf{P}_n^δ .*

Excursion measures. In the sequel, three distinct excursion measures will be considered. First, \mathcal{N} is the excursion measure of \underline{X} away from 0, where in this case, the excursion measure is normalized so that the local time of \underline{X} at 0 at time t is taken equal to $-\min(0, \inf_{0 \leq s \leq t} X(s))$. This normalization will always be considered for processes reflected above their past infimum.

Second, \mathcal{M} the push-forward of \mathcal{N} by the map $L(\cdot, T) = (L(a, T), a \geq 0)$, where from now on we will also denote by $(L(a, t), a, t \geq 0)$ the local time process of the canonical excursion ϵ . In other words, for any measurable function $f : \mathcal{E} \rightarrow [0, \infty)$, we have

$$\mathcal{M}(f) = \mathcal{N}(f \circ L(\cdot, T)).$$

Third, we define \mathcal{N}' as the push-forward of \mathcal{M} by \mathcal{L} , i.e., for any measurable function $f : \mathcal{E} \rightarrow [0, \infty)$, we have

$$\mathcal{N}'(f) = \mathcal{M}(f \circ \mathcal{L}).$$

Then the measures obtained by taking the push-forward of \mathcal{N} and \mathcal{N}' by T coincide, i.e., for any Borel set $A \subset [0, \infty)$ we have

$$(3) \quad \mathcal{N}(T \in A) = \mathcal{N}'(T \in A).$$

Indeed, we have by definition of \mathcal{N}' , \mathcal{M} and L :

$$\mathcal{N}'(T \in A) = \mathcal{M}\left(\int \epsilon \in A\right) = \mathcal{N}\left(\int_0^\infty L(a, T) da \in A\right) = \mathcal{N}(T \in A)$$

since $\int_0^\infty L(a, T) da = T$. As a side remark, note that it could be proved that $\mathcal{M}(1 \wedge T) = +\infty$ and so there is no regenerative process admitting \mathcal{M} as its excursion measure.

4. Heavy traffic of PS via excursion theory. The goal of this section is to prove forthcoming Theorem 4.1. Roughly speaking, it states that the sequence (Q_n) of PS queue-length processes started empty converge weakly to the regenerative process with excursion measure away from 0 equal to \mathcal{N}' . Recall that \mathcal{N}' is the push-forward of the excursion measure \mathcal{N} of \underline{X} by the successive application of $L(\cdot, T)$ (local time process at the first hitting time of 0) and \mathcal{L} (Lamperti transformation).

The push-forward \mathcal{M} of \mathcal{N} by the mere application of $L(\cdot, T)$ is not an excursion measure (in the above mentioned sense that $\mathcal{M}(1 \wedge T) = +\infty$), but we expect nonetheless that the distributions of the CMJ processes Z_n will converge in some sense to \mathcal{M} . This intuition is made precise in Theorem 4.8, where Z_n starts with one initial individual and is suitable conditioned, and Theorem 5.4, where Z_n starts from a large initial condition.

We also specify that \mathcal{N}' will be identified in Theorem 5.1 as the excursion measure away from 0 of $\beta^{-1}\underline{X}$, which is the reflected Brownian motion with drift $-\alpha/\beta$ and Gaussian coefficient $2/\beta$. This will ensure that the sequence (Q_n) actually converges weakly to this reflected process (see also Theorem 5.6 for general initial condition).

THEOREM 4.1. *Let Q_∞ be the process obtained by applying Itô's construction to the excursion measure \mathcal{N}' . Then the sequence (Q_n) under \mathbb{P}_n^\emptyset converges weakly to Q_∞ .*

To avoid any ambiguity, let us explain what we mean by Itô's construction, see Blumenthal [3] for instance. Let ∂ be some cemetery point and $e = (e_t, t \geq 0)$ be an $\mathcal{E} \cup \{\partial\}$ -valued Poisson point process with intensity measure \mathcal{N}' . Define

$$\tilde{L}(t) = \sum_{0 \leq s \leq t} T(e_s)$$

with the convention $T(\partial) = 0$. Since $\mathcal{N}'(1 \wedge T) < +\infty$ by (3), \tilde{L} is a subordinator with Lévy measure $\mathcal{N}'(T \in \cdot)$. Let \tilde{L}^{-1} be the right-continuous inverse of \tilde{L} , then the process Q_∞ is defined via the following formula:

$$Q_\infty(t) = e_{\tilde{L}^{-1}(t-)} \left(t - \tilde{L}(\tilde{L}^{-1}(t)-) \right) \mathbb{1}_{\Delta \tilde{L}(\tilde{L}^{-1}(t)) \neq 0}, \quad t \geq 0.$$

We first prove in Section 4.1 preliminary results on Lévy processes. Section 4.2 is devoted to tightness, Section 4.3 proves a result of independent interest on CMJ processes and Section 4.4 provides the proof of Theorem 4.1.

4.1. *Preliminary results on Lévy processes.* We will need the following results on Lévy processes.

LEMMA 4.2. *For any $a > 0$, the sequence (X_n, T_{X_n}) under \mathbf{P}_n^a converges weakly to (X, T_X) under \mathbf{P}^a .*

PROOF. Since $\mathbf{P}^a(\forall \varepsilon > 0, \inf_{[0, \varepsilon]} \theta X < 0) = 1$, the result follows directly from Proposition VI.2.11 in Jacod and Shiryaev [15]. \square

LEMMA 4.3. *For any $\varepsilon > 0$, the sequence $(g_\varepsilon(\underline{X}_n), d_\varepsilon(\underline{X}_n))$ under \mathbf{P}_n^0 converges weakly to $(g_\varepsilon(\underline{X}), d_\varepsilon(\underline{X}))$ under \mathbf{P}^0 .*

PROOF. Remember that Ψ_n is the Laplace exponent of X_n . Since X_n drifts to $-\infty$, Ψ_n is continuous and strictly increasing and we denote Φ_n its inverse. Let similarly Ψ be the Laplace exponent of X and Φ its inverse. Since $X_n \Rightarrow X$ it is not hard to show that $\Phi_n(u) \rightarrow \Phi(u)$ for every $u \geq 0$. Moreover, for $t \geq 0$ let

$$\gamma_n(t) = \inf \{s \geq 0 : X_n(s) = -t\} \quad \text{and} \quad \gamma(t) = \inf \{s \geq 0 : X(s) = -t\}.$$

Then it is well-known, see for instance Bertoin [1, Theorem VII.1], that γ_n and γ are subordinators with Laplace exponent Φ_n and Φ , respectively. Since $\Phi_n(u) \rightarrow \Phi(u)$ for every $u \geq 0$, standard arguments imply that $\gamma_n \Rightarrow \gamma$. Moreover, since γ_n and γ are the right-continuous inverses of the local time processes of \underline{X}_n and \underline{X} at 0, we have the identities $g_\varepsilon(\underline{X}_n) = \gamma_n(t_\varepsilon^1(\gamma_n)-)$ and $d_\varepsilon(\underline{X}_n) = \gamma_n(t_\varepsilon^1(\gamma_n))$ and similarly without the subscript n , where in the rest of the proof we define $t_\varepsilon^1(f) = \inf\{t \geq 0 : |\Delta f(t)| > \varepsilon\}$ for any $f \in D$ and $\varepsilon > 0$.

Proposition 2.7 in Jacod and Shiryaev [15] shows that if $f_n, f \in D$ are such that $f_n \rightarrow f$, $t_\varepsilon^1(f) < +\infty$ and $\varepsilon \notin \{|\Delta f(t)| : t \geq 0\}$, then $f_n(t_\varepsilon^1(f_n)-) \rightarrow f(t_\varepsilon^1(f)-)$ as well as $f_n(t_\varepsilon^1(f_n)) \rightarrow f(t_\varepsilon^1(f))$. The desired result therefore follows from an application of the continuous mapping theorem, together with the fact that $\mathbf{P}^0(t_\varepsilon^1(\gamma) < +\infty, \varepsilon \notin \{|\Delta \gamma(t)| : t \geq 0\}) = 1$ for every $\varepsilon > 0$. \square

LEMMA 4.4. *For any $\varepsilon > 0$, the sequence $(\sigma X_n, T_{X_n})$ considered under $\mathbf{P}_n(\cdot | T_{X_n} > \varepsilon)$ converges weakly to (ϵ, T_ϵ) under $\mathcal{N}(\cdot | T > \varepsilon)$.*

PROOF. From now on and unless otherwise specified, we implicitly consider X_n under \mathbf{P}_n^0 and X under \mathbf{P}^0 . Since by definition the process σX_n under $\mathbf{P}_n(\cdot | T_{X_n} > \varepsilon)$ is equal in distribution to $e_\varepsilon(\underline{X}_n)$ and $\mathcal{N}(\cdot | T > \varepsilon)$ is

the law of $e_\varepsilon(\underline{X})$, the result is equivalent to showing that $(e_\varepsilon, T \circ e_\varepsilon)(\underline{X}_n) \Rightarrow (e_\varepsilon, T \circ e_\varepsilon)(\underline{X})$.

For $f \in D$ let $\mathcal{J}(f) = (f, g_\varepsilon(f), d_\varepsilon(f))$. In view of Lemma 2.3 and the continuous mapping theorem, to prove that $(e_\varepsilon, T \circ e_\varepsilon)(\underline{X}_n) \Rightarrow (e_\varepsilon, T \circ e_\varepsilon)(\underline{X})$ it is enough to show that $\mathcal{J}(X_n) \Rightarrow \mathcal{J}(X)$. We have $\underline{X}_n \Rightarrow \underline{X}$, while Lemma 4.3 shows that $(g_\varepsilon, d_\varepsilon)(\underline{X}_n) \Rightarrow (g_\varepsilon, d_\varepsilon)(\underline{X})$. Hence the sequence $(\mathcal{J}(X_n))$ is tight and we only need to identify accumulation points. Let (X', g', d') be any accumulation point and assume without loss of generality using Skorohod's representation theorem that $\mathcal{J}(X_n) \rightarrow (X', g', d')$: then X' is equal in distribution to \underline{X} and (g', d') to $(g_\varepsilon, d_\varepsilon)(\underline{X})$, and we only have to show that $(g', d') = (g_\varepsilon, d_\varepsilon)(X')$.

Since \underline{X}_n and X_n have the same generator in $(0, \infty)$ and the functional T is \mathbf{P} -a.s. continuous (see Lemma 4.2), it can be proved that $T(\theta_{t_n}\underline{X}_n) \rightarrow T(\theta_t X')$ for any $t_n, t \geq 0$ such that $t_n \rightarrow t$.

Let $t < g'$: since $g_\varepsilon(\underline{X}_n) \rightarrow g'$ we have $t < g_\varepsilon(\underline{X}_n)$ for n large enough, and for those n it holds by definition of $g_\varepsilon(\underline{X}_n)$ that $T(\theta_t \underline{X}_n) \leq \varepsilon$. Since $T(\theta_t \underline{X}_n) \rightarrow T(\theta_t X')$ we obtain that $T(\theta_t X') \leq \varepsilon$. Since $t < g'$ is arbitrary, this proves that $g_\varepsilon(X') \geq g'$ and since they are equal in distribution they must be equal almost surely. Since $T(\theta_{g_\varepsilon(\underline{X}_n)} \underline{X}_n) = d_\varepsilon(\underline{X}_n) - g_\varepsilon(\underline{X}_n)$, letting $n \rightarrow +\infty$ shows that $T(\theta_{g'} X') = d' - g'$ and so $d' = g_\varepsilon(X') + T(\theta_{g_\varepsilon(X')} X') = d_\varepsilon(X')$. The proof is complete. \square

4.2. Tightness. Although tightness is usually a technical issue, it comes here from a simple queueing argument. Theorem 4.5 may look naive to an experienced reader, but to the best of our knowledge its implications in terms of tightness have never been used before; similar arguments could for instance have been used in Limic [25]. Since Processor-Sharing is a symmetric service discipline, the following result is a direct consequence of Theorems 3.10 and 3.6 in Kelly [17].

THEOREM 4.5. *The departure process of the queue length process q_n under \mathbb{P}_n^* is a Poisson process with parameter λ_n .*

COROLLARY 4.6. *The sequence of processes (Q_n) under \mathbb{P}_n^* is C-tight.*

PROOF. Writing a_n and d_n for the arrival and departure processes respectively, we can write $Q_n(t) = Q_n(0) + A_n(t) - D_n(t)$ for $t \geq 0$ with $A_n(t) = (a_n(n^2 t) - n^2 \lambda_n t)/n$ and $D_n(t) = (d_n(n^2 t) - n^2 \lambda_n t)/n$. By Theorem 4.5, a_n and d_n under \mathbb{P}_n^* are two Poisson processes with intensity λ_n . Since $\lambda_n \rightarrow \lambda$, both (A_n) and (D_n) converge in distribution to a Brownian motion and so are C-tight, hence so is the difference $(A_n - D_n)$. Since

$(Q_n(0))$ under \mathbb{P}_n^* converges to an exponential random variable, this shows that (Q_n) under \mathbb{P}_n^* is C-tight. \square

Corollary 4.6 encompasses all the tightness results we need. To be more specific, by considering Q_n under \mathbb{P}_n^* and shifting it at time T_{Q_n} we can get the tightness of (Q_n) under \mathbb{P}_n^\emptyset . Also, we can get the tightness of CMJ processes by suitably selecting excursions of Q_n and applying \mathcal{L}^{-1} . The elementary operations that we need to perform preserve C-tightness by Corollary 2.4 and Lemma 2.5, and so we get the following result.

COROLLARY 4.7. *The sequence (Q_n) under \mathbb{P}_n^\emptyset is C-tight and for any $\varepsilon > 0$, the sequence (Z_n) under $\mathbb{P}_n(\cdot \mid \int Z_n > \varepsilon)$ is C-tight.*

PROOF. By regeneration of Q_n , Q_n under \mathbb{P}_n^\emptyset is equal in distribution to θQ_n under \mathbb{P}_n^* . Since T_{Q_n} is equal in distribution to T_{X_n} , Lemma 4.2 shows that (T_{Q_n}) under \mathbb{P}_n^* is tight. Combining Corollaries 4.6 and 2.4 shows the C-tightness of (Q_n) under \mathbb{P}_n^\emptyset .

For Z_n , since $Z_n = \mathcal{L}^{-1}(Q_n)$ and $\int Z_n = T_{Q_n}$ by Lemma 3.3, one sees that Z_n under $\mathbb{P}_n(\cdot \mid \int Z_n > \varepsilon)$ is equal in distribution to $\mathcal{L}^{-1}(e_\varepsilon(Q_n))$ under \mathbb{P}_n^\emptyset . Since the zero set of Q_n is equal in distribution to the zero set of \underline{X}_n , $(g_\varepsilon, d_\varepsilon)(Q_n)$ under \mathbb{P}_n^\emptyset is equal in distribution to $(g_\varepsilon, d_\varepsilon)(\underline{X}_n)$ under \mathbf{P}_n^0 . Thus Lemma 4.3 implies that the sequence $(g_\varepsilon, d_\varepsilon)(Q_n)$ under \mathbb{P}_n^\emptyset is tight. Since by definition $e_\varepsilon(Q_n) = (\sigma_{d_\varepsilon(Q_n) - g_\varepsilon(Q_n)} \circ \theta_{g_\varepsilon(Q_n)})(Q_n)$, combining the results of Corollary 2.4 and Lemma 2.5 and using also that (Q_n) under \mathbb{P}_n^\emptyset is C-tight, we obtain the C-tightness of (Z_n) under $\mathbb{P}_n(\cdot \mid \int Z_n > \varepsilon)$. \square

4.3. Weak convergence of CMJ processes. The following result is of independent interest in the area of branching processes. Using similar techniques and ideas, the conditionings $\{\int Z_n > \varepsilon\}$ and $\{\int \epsilon > \varepsilon\}$ in the next statement could be replaced by $\{T_{Z_n} > \varepsilon\}$ and $\{T_\epsilon > \varepsilon\}$, respectively. Theorem 5.4 in the following section gives another result with a large initial condition.

THEOREM 4.8. *For any $\varepsilon > 0$, the sequence (Z_n) under $\mathbb{P}_n(\cdot \mid \int Z_n > \varepsilon)$ converges weakly to $\mathcal{M}(\cdot \mid \int \epsilon > \varepsilon)$.*

PROOF. In the remainder of the proof we implicitly consider Z_n under $\mathbb{P}_n(\cdot \mid \int Z_n > \varepsilon)$ and X_n under $\mathbf{P}_n(\cdot \mid T_{X_n} > \varepsilon)$ and we denote by L_n^0 the process $(L_n(a, T_{X_n}), a \geq 0)$. Lemma 3.3 shows that Z_n is equal in distribution to L_n^0 , so we only have to show that $L_n^0 \Rightarrow \mathcal{M}(\cdot \mid \int \epsilon > \varepsilon)$.

Lemma 4.7 shows that the sequence (L_n^0) is C-tight, so we only have to identify accumulation points. Let Z be any accumulation point and

assume without loss of generality that $L_n^0 \Rightarrow Z$. Lemma 4.4 shows that $(\sigma X_n, T_{X_n})$ converges weakly to (ϵ, T_ϵ) under $\mathcal{N}(\cdot | T > \varepsilon)$. Then the sequence $(\sigma X_n, T_{X_n}, L_n^0)$ is tight, let (e, τ, Z') be any accumulation point, so that (e, τ) is equal in distribution to (ϵ, T_ϵ) under $\mathcal{N}(\cdot | T > \varepsilon)$ and Z' to Z . Assume without loss of generality by Skorohod's representation theorem that $(\sigma X_n, T_{X_n}, L_n^0) \rightarrow (e, \tau, Z')$. By definition we have

$$\int_0^{T_{X_n}} \phi(\sigma X_n(t)) dt = \int_0^\infty \phi(a) L_n^0(a) da$$

for all continuous function ϕ with a compact support. Thus passing to the limit, the dominated convergence theorem (or uniform convergence arguments) shows that

$$\int_0^\tau \phi(e(t)) dt = \int_0^\infty \phi(a) Z'(a) da$$

which shows that Z' is the local time process of e up to time τ . The result is proved. \square

4.4. *Proof of Theorem 4.1.* We begin with a preliminary result.

LEMMA 4.9. *For any $\varepsilon > 0$, the sequence $(\sigma Q_n, T_{Q_n})$ considered under $\mathbb{P}_n(\cdot | T_{Q_n} > \varepsilon)$ converges weakly to (ϵ, T_ϵ) under $\mathcal{N}'(\cdot | T > \varepsilon)$.*

PROOF. Until the end of this step, we consider implicitly the process Q_n , and hence Z_n , under $\mathbb{P}_n(\cdot | T_{Q_n} > \varepsilon) = \mathbb{P}_n(\cdot | \int Z_n > \varepsilon)$. By Theorem 4.8, we know that $Z_n \Rightarrow \mathcal{M}(\cdot | \int \epsilon > \varepsilon)$. Moreover, Lemma 3.3 implies that T_{Z_n} is equal in distribution to $\|X_n\|_{T_{X_n}}$ under $\mathbf{P}_n(\cdot | T_{X_n} > \varepsilon)$ which converges, in view of Lemma 4.4 and using the continuous mapping theorem, to $\|\epsilon\|_\infty$ under $\mathcal{N}(\cdot | T > \varepsilon)$. In particular, the sequence (T_{Z_n}) is tight, so Lemma 2.5 implies that the sequence $(\mathcal{L}(Z_n))$ converges weakly to the push-forward of $\mathcal{M}(\cdot | \int \epsilon > \varepsilon)$ by \mathcal{L} . Since $\mathcal{L}(Z_n) = \sigma Q_n$ by Lemma 3.3 and the push-forward of $\mathcal{M}(\cdot | \int \epsilon > \varepsilon)$ by \mathcal{L} is by definition equal to $\mathcal{N}'(\cdot | T > \varepsilon)$ we obtain the convergence of the sequence (σQ_n) towards $\mathcal{N}'(\cdot | T > \varepsilon)$.

On the other hand, since the workload associated to Q_n is equal in distribution to \underline{X}_n , we obtain that T_{Q_n} is equal in distribution to T_{X_n} under $\mathbf{P}_n(\cdot | T_{X_n} > \varepsilon)$, hence (T_{Q_n}) converges weakly to T under $\mathcal{N}(\cdot | T > \varepsilon)$ in view of Lemma 4.4. Since T under $\mathcal{N}(\cdot | T > \varepsilon)$ is equal in distribution to T under $\mathcal{N}'(\cdot | T > \varepsilon)$ by (3) we obtain the convergence of (T_{Q_n}) towards T under $\mathcal{N}'(\cdot | T > \varepsilon)$. To conclude that the joint convergence holds we invoke Corollary 2.2. \square

We now prove Theorem 4.1. Since the sequence (Q_n) under \mathbb{P}_n^\emptyset is C-tight by Corollary 4.7, we only have to identify accumulation points. So let Q be any accumulation point and assume without loss of generality that $Q_n \Rightarrow Q$: we must prove that Q is equal in distribution to Q_∞ . In the rest of this section, for $\varepsilon > 0$ let $A_\varepsilon : D \rightarrow \mathcal{E} \times [0, \infty) \times [0, \infty)$ be the map given by $A_\varepsilon = (e_\varepsilon, g_\varepsilon, d_\varepsilon)$ and let $\Phi_\varepsilon : D \rightarrow D$ the map that truncates excursions with length smaller than ε , i.e., for $f \in D$ and $t \geq 0$ we put $\Phi_\varepsilon(f)(t) = f(t)$ if $f(t) \neq 0$ and the excursion e of f straddling t satisfies $T_e > \varepsilon$; otherwise we put $\Phi_\varepsilon(f)(t) = 0$. We prove that Q is equal in distribution to Q_∞ in two steps.

First step. Let $\varepsilon > 0$: we first prove that $(Q_n, A_\varepsilon(Q_n)) \Rightarrow (Q, A_\varepsilon(Q))$. First, note that $d_\varepsilon - g_\varepsilon = T \circ e_\varepsilon$ and so Lemma 4.9 implies, by definition of Q_∞ , that $(e_\varepsilon, d_\varepsilon - g_\varepsilon)(Q_n) \Rightarrow (e_\varepsilon, d_\varepsilon - g_\varepsilon)(Q_\infty)$. Moreover, $g_\varepsilon(Q_n)$ is equal in distribution to $g_\varepsilon(\underline{X}_n)$ under \mathbf{P}_n^0 and so Lemma 4.3 shows that $g_\varepsilon(Q_n) \Rightarrow g_\varepsilon(\underline{X})$. By (3) and the definition of Q_∞ , $g_\varepsilon(Q_\infty)$ and $g_\varepsilon(\underline{X})$ are equal in distribution and so $A_\varepsilon(Q_n) \Rightarrow A_\varepsilon(Q_\infty)$.

Let (Q', A') be any accumulation point of the tight sequence $(Q_n, A_\varepsilon(Q_n))$, so that Q' is equal in distribution to Q and A' to $A_\varepsilon(Q_\infty)$. Assume without loss of generality, using Skorohod's representation theorem, that the almost sure convergence $(Q_n, A_\varepsilon(Q_n)) \rightarrow (Q', A')$ holds: we show that $A' = A_\varepsilon(Q')$ which will prove that (Q', A') is equal in distribution to $(Q, A_\varepsilon(Q))$.

Note $A' = (e', g', d')$: the convergence $(Q_n, A_\varepsilon(Q_n)) \rightarrow (Q', A')$ implies in view of Lemma 2.3 and the definition of e_ε that $e' = (\sigma_{d'-g'} \circ \theta_{g'})(Q')$. Since e' is equal in distribution to $e_\varepsilon(Q_\infty)$, we see that e' is the excursion of Q' with endpoints $g' < d'$, and that it satisfies $T_{e'} > \varepsilon$. To show that $e' = e_\varepsilon(Q')$ it remains to show that this is the first such excursion. Since Q' is continuous it is enough to show that $\inf_{[a, a+\varepsilon]} Q' = 0$ for any $a < g'$. So let $a < g'$: since $g_\varepsilon(Q_n) \rightarrow g'$ we must have $a < g_\varepsilon(Q_n)$ for n large enough, and for those n , by definition of g_ε we must have $\inf_{[a, a+\varepsilon]} Q_n = 0$. Since $\inf_{[a, a+\varepsilon]} Q_n \rightarrow \inf_{[a, a+\varepsilon]} Q'$ by continuity, we obtain $\inf_{[a, a+\varepsilon]} Q' = 0$ which proves that $(Q_n, A_\varepsilon(Q_n)) \Rightarrow (Q, A_\varepsilon(Q))$. Note in particular that since we have argued that $A_\varepsilon(Q_n) \Rightarrow A_\varepsilon(Q_\infty)$ we also have $A_\varepsilon(Q) = A_\varepsilon(Q_\infty)$.

Second step. Since Q_n regenerates at 0, we have for any measurable functions $f, h, i : D \rightarrow [0, \infty)$

$$\begin{aligned} \mathbb{E}_n^\emptyset (f(\sigma_{g_\varepsilon(Q_n)} Q_n) h(e_\varepsilon(Q_n)) i(\theta_{d_\varepsilon(Q_n)} Q_n)) \\ = \mathbb{E}_n^\emptyset (f(\sigma_{g_\varepsilon(Q_n)} Q_n)) \mathbb{E}_n^\emptyset (h(e_\varepsilon(Q_n))) \mathbb{E}_n^\emptyset (i(Q_n)). \end{aligned}$$

Consider now f, h and i continuous and bounded and let $n \rightarrow +\infty$ in

both sides of the previous display. Since $(Q_n, A_\varepsilon(Q_n)) \Rightarrow (Q, A_\varepsilon(Q))$ by the first step, Lemma 2.3 together with the continuous mapping theorem gives

$$\mathbb{E} (f(\sigma_{g_\varepsilon(Q)}Q)h(e_\varepsilon(Q))i(\theta_{d_\varepsilon(Q)}Q)) = \mathbb{E} (f(\sigma_{g_\varepsilon(Q)}Q)) \mathbb{E} (h(e_\varepsilon(Q))) \mathbb{E} (i(Q)).$$

This implies that $\sigma_{g_\varepsilon(Q)}Q$, $e_\varepsilon(Q)$ and $\theta_{d_\varepsilon(Q)}Q$ are independent and that $\theta_{d_\varepsilon(Q)}Q$ is equal in distribution to Q . Since in addition $A_\varepsilon(Q)$ is equal in distribution to $A_\varepsilon(Q_\infty)$ by the previous step we obtain that $\Phi_\varepsilon(Q)$ and $\Phi_\varepsilon(Q_\infty)$ are equal in distribution. For any $f \in D$ and any $t \geq 0$, one easily sees that $\Phi_\varepsilon(f)(t) \rightarrow f(t)$ as $\varepsilon \rightarrow 0$. In particular, $\Phi_\varepsilon(Q)$ converges in the sense of finite-dimensional distributions to Q , and $\Phi_\varepsilon(Q_\infty)$ to Q_∞ , as $\varepsilon \rightarrow 0$. Hence Q and Q_∞ are equal in distribution which achieves the proof of Theorem 4.1.

5. Identification of the limit and general initial condition. Theorem 4.1 is the most important result of the paper, where the convergence of (Q_n) under \mathbb{P}_n^\emptyset is shown based on the convergence of its long excursions. The formulation of Theorem 4.1 reflects this approach, where the limiting process is defined through its excursion measure. In general, it is not clear whether a more explicit definition of Q_∞ can be given. For instance, in the infinite variance case we expect a similar statement to hold, where \mathcal{N} is the excursion measure of a reflected, spectrally positive Lévy process; in this case we do not know whether Q_∞ can be described in another way. However, here \mathcal{N}' turns out to be the excursion measure of $\beta^{-1}\underline{X}$, which is a reflected Brownian motion with drift $-\alpha/\beta$ and Gaussian coefficient $2/\beta$. This allows us to identify Q_∞ as $\beta^{-1}\underline{X}$ (see also forthcoming Theorem 5.6 for a general initial condition).

THEOREM 5.1. *\mathcal{N}' is the excursion measure of the process $\beta^{-1}\underline{X}$. In particular, Q_∞ is equal in distribution to $\beta^{-1}\underline{X}$ under \mathbf{P}^0 .*

PROOF. As a variation of the original Ray-Knight theorems [21, 28], it is known that \mathcal{M} , the push-forward of \mathcal{N} by the local time process, is the excursion measure of Feller diffusion, where we think here of the Feller diffusion as the solution Y to the stochastic differential equation

$$dY_t = -(\alpha/\beta) Y_t dt + \sqrt{(1/\beta) Y_t} dB_t,$$

with B the standard Brownian motion; see for instance Pardoux and Wakolbinger [26]. It remains to show that the push-forward of \mathcal{M} by \mathcal{L} gives \mathcal{N}'' , where \mathcal{N}'' stands for the excursion measure of $\beta^{-1}\underline{X}$ (recall that the local

time of a reflected process is chosen equal to the past infimum of the initial process).

For $\varepsilon > 0$ and $f \in D$, let $T^\varepsilon(f) = \inf\{t \geq 0 : f(t) \geq \varepsilon\}$: then $\theta_{T^\varepsilon(\epsilon)}(\epsilon)$, the canonical excursion shifted at time $T^\varepsilon(\epsilon)$, under $\mathcal{N}''(\cdot | T^\varepsilon < +\infty)$ is equal in distribution to $\sigma(\beta^{-1}X)$ under $\mathbf{P}^{\beta\varepsilon}$. The Lamperti representation theorem, see Lamperti [23], asserts that $\sigma(\beta^{-1}X)$ under $\mathbf{P}^{\beta\varepsilon}$ is the image by \mathcal{L} of the Feller diffusion started at ε , i.e., of $\theta_{T^\varepsilon(\epsilon)}(\epsilon)$ under $\mathcal{M}(\cdot | T^\varepsilon < +\infty)$. Hence the relation

$$\mathcal{N}''(f \circ \theta_{T^\varepsilon} | T^\varepsilon < +\infty) = \mathcal{M}(f \circ \theta_{T^\varepsilon \circ \mathcal{L}} \circ \mathcal{L} | T^\varepsilon \circ \mathcal{L} < +\infty)$$

holds for any non-negative, measurable function $f : \mathcal{E} \rightarrow [0, \infty)$. But by definition of \mathcal{N}' the right hand side is precisely $\mathcal{N}'(f \circ \theta_{T^\varepsilon} | T^\varepsilon < +\infty)$ and so

$$\mathcal{N}''(f \circ \theta_{T^\varepsilon} | T^\varepsilon < +\infty) = \mathcal{N}'(f \circ \theta_{T^\varepsilon} | T^\varepsilon < +\infty)$$

which can be rewritten as

$$\mathcal{N}''(f \circ \theta_{T^\varepsilon} \mathbb{1}_{\{T^\varepsilon < +\infty\}}) = \frac{\mathcal{N}''(T^\varepsilon < +\infty)}{\mathcal{N}'(T^\varepsilon < +\infty)} \mathcal{N}'(f \circ \theta_{T^\varepsilon} \mathbb{1}_{\{T^\varepsilon < +\infty\}}).$$

Applying this for $f = \mathbb{1}_{\{T^1 < +\infty\}}$ and $\varepsilon < 1$ we obtain

$$\frac{\mathcal{N}''(T^\varepsilon < +\infty)}{\mathcal{N}'(T^\varepsilon < +\infty)} = \frac{\mathcal{N}''(T^1 < +\infty)}{\mathcal{N}'(T^1 < +\infty)}$$

and so for $\varepsilon < 1$ we have

$$\mathcal{N}''(f \circ \theta_{T^\varepsilon} \mathbb{1}_{\{T^\varepsilon < +\infty\}}) = \frac{\mathcal{N}''(T^1 < +\infty)}{\mathcal{N}'(T^1 < +\infty)} \mathcal{N}'(f \circ \theta_{T^\varepsilon} \mathbb{1}_{\{T^\varepsilon < +\infty\}}).$$

Because $f \circ \theta_{T^\varepsilon} \mathbb{1}_{\{T^\varepsilon < +\infty\}}$ converges to f as $\varepsilon \rightarrow 0$, we get that \mathcal{N}' and \mathcal{N}'' are proportional. Moreover, since \mathcal{N}'' is the excursion measure of $\beta^{-1}\underline{X}$ and \mathcal{N} is that of \underline{X} , we have $\mathcal{N}''(T \in \cdot) = \mathcal{N}(T \in \cdot)$, and so (3) shows that the multiplicative constant must be equal to one. This proves the result. \square

In view of this result, it is consistent, and convenient, to redefine Q_∞ as $Q_\infty = \beta^{-1}\underline{X}$. In the rest of this section, $\zeta > 0$ is some positive real number, (ζ_n) is an integer-valued sequence such that $\zeta_n/n \rightarrow \zeta$ and we define

$$\tau_n = \inf\{t \geq 0 : L_n(0, t) > \zeta_n/n\} \text{ and } \tau = \inf\{t \geq 0 : L(0, t) = \zeta\}.$$

The goal of this section is to prove that the sequence (Q_n) under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $\beta^{-1}\underline{X}$ under $\mathbf{P}^{\zeta\beta}$. To do so, we first prove that $(\sigma Q_n, T_{Q_n})$

under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $(\sigma Q_\infty, T_{Q_\infty})$ under $\mathbf{P}^{\zeta\beta}$ through a series of steps similar to those performed in Section 4. In the sequel, we will use the fact that S_n^*/n is the distribution of $X_n(\eta_n)$ under $\mathbf{P}_n^0(\cdot | \eta_n < +\infty)$ where $\eta_n = \inf\{t \geq 0 : X_n(t) > 0\}$, see Theorem VII.17 in Bertoin [1]; we will informally call $X_n(\eta_n)$ the overshoot of X_n . The following result can be proved using standard arguments on Lévy processes, and so we omit the proof.

LEMMA 5.2. *The sequence (X_n, τ_n) under $\mathbf{P}_n^0(\cdot | \tau_n < +\infty)$ converges weakly to (X, τ) under $\mathbf{P}^0(\cdot | \tau < +\infty)$.*

LEMMA 5.3. *The sequence (Z_n) under $\mathbb{P}_n^{\zeta_n^*}$ is C-tight.*

PROOF. By Lemma 3.3, $Z_n = \mathcal{L}^{-1}(Q_n)$ and so the sequence (Z_n) under \mathbb{P}_n^* is C-tight, as can be seen by combining Corollary 4.6 and Lemma 2.5.

Let Z'_n be a process defined on the same probability space as Z_n , independent of Z_n and with the same law as Z_n under \mathbb{P}_n^* . Now let Z''_n be the (rescaled) CMJ process defined as $Z''_n := Z'_n + Z_n$. Because of the lack-of-memory property of the geometric random variable, Z''_n under $\mathbb{P}_n^{\zeta_n^*}$ has the same law as Z_n under $\mathbb{P}_n^*(\cdot | Z_n(0) \geq \zeta_n/n)$.

Since (Z_n) under \mathbb{P}_n^* is C-tight and $(Z_n(0))$ under \mathbb{P}_n^* converges weakly to an exponential random variable, it is easy to show that (Z_n) under $\mathbb{P}_n^*(\cdot | Z_n(0) \geq \zeta_n/n)$ is C-tight. In particular, the two sequences (Z''_n) and (Z'_n) under $\mathbb{P}_n^{\zeta_n^*}$ are C-tight. Since the difference of two C-tight sequences is also C-tight we obtain the C-tightness of (Z_n) under $\mathbb{P}_n^{\zeta_n^*}$ which was to be proved. \square

THEOREM 5.4. *The sequence (Z_n) considered under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $(L(a, \tau), a \geq 0)$ under $\mathbf{P}^0(\cdot | \tau < +\infty)$. In particular, the sequence (Z_n) under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $\mathcal{L}^{-1}(Q_\infty)$ under $\mathbf{P}^{\beta\zeta}$, which is the solution Y to the stochastic differential equation*

$$dY_t = -(\alpha/\beta) Y_t dt + \sqrt{(1/\beta) Y_t} dB_t,$$

with initial condition $Y_0 = \zeta$.

PROOF. From Lemma 3.3 and the branching property, one gets that Z_n under $\mathbb{P}_n^{\zeta_n^*}$ is equal in distribution to $(L_n(a, \tau_n), a \geq 0)$ considered under $\mathbf{P}_n^0(\cdot | \tau_n < +\infty)$: then the proof follows analogously as for Theorem 4.8 using Lemma 5.2 instead of Lemma 4.4. The identification of the limit as being $\mathcal{L}(Q_\infty)$ comes as in the proof of Theorem 5.1 from a combination of Ray-Knight theorem together with the Lamperti representation theorem. \square

LEMMA 5.5. *The sequence $(\sigma Q_n, T_{Q_n})$ under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $(\sigma Q_\infty, T_{Q_\infty})$ under $\mathbf{P}^{\beta\zeta}$.*

PROOF. Unless otherwise specified we consider implicitly Z_n and Q_n under $\mathbb{P}_n^{\zeta_n^*}$ and Q_∞ under $\mathbf{P}^{\beta\zeta}$. Thanks to Corollary 2.2, we only have to show that $\sigma Q_n \Rightarrow Q_\infty$ and $T_{Q_n} \Rightarrow T_{Q_\infty}$.

By Theorem 5.4, we know that $Z_n \Rightarrow \mathcal{L}(Q_\infty)$. By the branching property and the fact that the overshoot of X_n is distributed like S_n^*/n , T_{Z_n} is equal in distribution to $\|X_n\|_{\tau_n}$ under $\mathbf{P}_n^0(\cdot | \tau_n < +\infty)$. In view of Lemma 5.2 combined with the continuous mapping theorem we get weak convergence, and in particular tightness, of (T_{Z_n}) . Thus Lemma 2.5 implies that $\mathcal{L}(Z_n) \Rightarrow \mathcal{L}(\mathcal{L}^{-1}(Q_\infty))$. Since $\mathcal{L}(Z_n) = \sigma Q_n$ and $\mathcal{L}(\mathcal{L}^{-1}(Q_\infty)) = \sigma Q_\infty$ this proves that $\sigma Q_n \Rightarrow \sigma Q_\infty$.

Since the workload process has the same law as \underline{X}_n , T_{Q_n} is equal in distribution to T_{X_n} with the initial condition $X_n(0) = A_n$, where A_n is equal to the sum of ζ_n independent copies of S_n^*/n . In particular, since $\mathbb{E}(S_n^*) \rightarrow \beta$ by (2), the strong law of large number implies that $A_n \rightarrow \zeta\beta$. Thus Lemma 5.2 implies that $T_{Q_n} \Rightarrow T_{Q_\infty}$ which concludes the proof. \square

PROPOSITION 5.6. *The sequence (Q_n) under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $\beta^{-1}\underline{X}$ under $\mathbf{P}^{\zeta\beta}$.*

PROOF. Let $\mathcal{C} : D \times [0, \infty) \times D \rightarrow D$ be the concatenation map defined for $f, h \in D$ and $s, t \geq 0$ by

$$\mathcal{C}(f, t, h)(s) = \begin{cases} f(s) & \text{if } s < t, \\ h(t - s) & \text{if } s \geq t. \end{cases}$$

Imagine for a moment that we knew that \mathcal{C} was continuous in the following sense: if $f_n, h_n, f, h \in D$ and $t_n, t > 0$ are such that $f_n \rightarrow f$, $h_n \rightarrow h$, f and h are continuous with $f(t) = h(t)$ and $t_n \rightarrow t$, then $\mathcal{C}(f_n, t_n, h_n) \rightarrow \mathcal{C}(f, t, h)$. Then the result would follow from this result and the continuous mapping theorem, since $Q_n = \mathcal{C}(\sigma Q_n, T_{Q_n}, \theta Q_n)$ and $(\sigma Q_n, T_{Q_n}, \theta Q_n)$ under $\mathbb{P}_n^{\zeta_n^*}$ converges weakly to $(\sigma Q_\infty, T_{Q_\infty}, \theta Q_\infty)$ under $\mathbf{P}^{\beta\zeta}$ by Lemmas 5.5 and Theorem 4.1 (using that $(\sigma Q_n, T_{Q_n})$ and θQ_n are independent).

Hence we only have to prove continuity of \mathcal{C} . Let (μ_n) be any sequence of functions such that $\sup_{t \geq 0} |\mu_n(t) - t| \rightarrow 0$ and such that for each $n \geq 1$, μ_n is continuous, strictly increasing and satisfies $\mu_n(0) = 0$ and $\mu_n(t) = t_n$. Then for any $s \geq 0$ one has

$$|\mathcal{C}(f_n, t_n, h_n)(\mu_n(s)) - \mathcal{C}(f, t, h)(s)| = \begin{cases} |f_n(\mu_n(s)) - f(s)| & \text{if } s < t, \\ |h_n(\mu_n(s) - t_n) - h(s - t)| & \text{if } s \geq t. \end{cases}$$

Since $f_n \rightarrow f$, $h_n \rightarrow h$, f and h are continuous and $\sup_{t \geq 0} |\mu_n(t) - t| \rightarrow 0$ this implies that $\mathcal{C}(f_n, t_n, h_n) \circ \mu_n \rightarrow \mathcal{C}(f, t, h)$. This means precisely that $(\mathcal{C}(f_n, t_n, h_n))$ converges in the Skorohod J_1 topology to $\mathcal{C}(f, t, h)$ and since $\mathcal{C}(f, t, h)$ is continuous by choice of f and h this means that $\mathcal{C}(f_n, t_n, h_n) \rightarrow \mathcal{C}(f, t, h)$. The result is proved. \square

6. Discussion. From a branching perspective, the result of Theorem 5.4 is quite surprising: for the sequence (Z_n) to converge, one would naively think that the initial individuals should start with the “normal” life length distribution S_n instead of its forward recurrence time S_n^* . This subtlety does not seem to appear in previous works on scaling limits of continuous-time branching processes. Although surprising from a branching perspective, this phenomenon is well-known in the folklore of queueing theory. In the rest of this discussion, fix an integer sequence (ζ_n) such that $\zeta_n/n \rightarrow \zeta > 0$.

Discontinuity of Local Times. Given some random variable V_n , let $Y_n^{V_n}$ be the process obtained as follows:

- $Y_n^{V_n}(0)$ is distributed like V_n/n ;
- $Y_n^{V_n}$ has the same generator as X_n in $(0, \infty)$;
- when $Y_n^{V_n}$ hits 0, it stays there for an exponential duration with parameter $n\lambda_n$ and then jumps according to V_n/n ;
- $Y_n^{V_n}$ is stopped at the time of its ζ_n th visit to 0.

Note also $L_n^{V_n}$ the local time process of $Y_n^{V_n}$: the branching property together with Lemma 3.3 show that $L_n^{V_n}$ is equal in distribution to Z_n started with ζ_n individuals with i.i.d. life lengths with common distribution V_n .

Under the conditions that we imposed, it can be proved that $Y_n^{V_n} \Rightarrow \underline{X}$ for both $V_n = S_n$ and $V_n = S_n^*$. Nonetheless, their local time processes $(L_n^{S_n})$ and $(L_n^{S_n^*})$ have different asymptotic behavior. On the one hand, $(L_n^{S_n^*})$ converges in view of Theorem 5.4 to the Feller diffusion with drift $-\alpha/\beta$ and Gaussian coefficient $2/\beta$, started at ζ . On the other hand, it can be proved that $(L_n^{S_n})$ converges in the sense of finite-dimensional distributions to a discontinuous process with value ζ at time 0, but distributed for nonzero times as the same Feller diffusion started at $\zeta/(\beta\lambda)$. In particular, the sequence $(L_n^{S_n^*})$ cannot be tight, although for each $\varepsilon > 0$ the sequence $(\theta_\varepsilon L_n^{S_n^*})$ is tight. We now provide an interpretation of this phenomenon in terms of state-space collapse, a well-known property in queueing theory.

State Space Collapse. Consider a sequence (S'_n) of positive random variables such that $\mathbb{E}_n(S'_n) \rightarrow \beta' \in (0, \infty)$, so taking for example $S'_n = S_n^*$ yields $\beta' = \beta$ by (2). Denote by \mathbb{P}'_n the law of the PS queue started with

ζ_n customers with i.i.d. service requirements distributed as S'_n . Let w_n the workload process associated to q_n and \mathcal{W}_n the rescaled process $\mathcal{W}_n(t) = w_n(n^2t)/n$. Then \mathcal{W}_n is equal in distribution to \underline{X}_n and so converges weakly to \underline{X} , which we write \mathcal{W} for clarity.

The state space collapse property states that the sequence (Q_n, \mathcal{W}_n) under $\mathbb{P}_n^{\zeta_n^*}$ converges to (Q, \mathcal{W}) which satisfy $Q = c\mathcal{W}$ for some constant $c > 0$. By the law of large numbers, $\mathcal{W}_n(0)$ under $\mathbb{P}_n^{\zeta_n^*}$ converges to $\zeta\beta$ while $Q_n(0)$ converges to ζ , which shows that $c = \beta^{-1}$. To understand the behavior of the processes under \mathbb{P}'_n , one needs to zoom in around time 0.

Define the *fluid limits* (\bar{q}_n) and (\bar{w}_n) of (q_n) and (w_n) as the rescaled processes $\bar{q}_n(t) = q_n(nt)/n$ and $\bar{w}_n(t) = w_n(nt)/n$. Fluid limits can be thought of as functional law of large numbers (whereas heavy-traffic approximations can be thought of as functional central limit theorems). By definition, at the critical point the amount of work that enters the queue is equal to the amount of work that exits it. Hence it is not surprising that (\bar{w}_n) under \mathbb{P}'_n converges to the deterministic function w with constant value $w_0 = \zeta\beta'$. Note that the workload process does not fluctuate on the fluid time scale n , while it does on the diffusion time scale n^2 .

Let q be the limit of (\bar{q}_n) under \mathbb{P}'_n , so that $q(0) = \zeta$ (see Gromoll et al. [13]). Moreover it is known that as t goes to infinity, $q(t)$ converges to an equilibrium point q_∞ . In steady state the residual service requirement of each customer has mean β , which suggests thanks to the law of large numbers that $q_\infty\beta$ must satisfy $q_\infty\beta = w_0$.

So it takes a time of order of n for the (scaled) queue length process to go from ζ to $q_\infty = w_0/\beta = \zeta\beta'/\beta$. Since the time scale n^2 of the heavy traffic approximation is orders of magnitude larger, this happens instantaneously on the diffusion time scale and causes a discontinuity when $\beta \neq \beta'$.

Once the process has reached the equilibrium point of the fluid limit, the state space collapse property applies. In particular, this show that (Q_n) under \mathbb{P}'_n should converge to a process Q such that $Q(0) = \zeta$ and $Q(t) = \beta^{-1}\mathcal{W}(t)$ for $t > 0$. In particular, $Q(0+) = \zeta\beta/\beta'$ is different from $Q(0)$ when $\beta' \neq \beta$, which provides yet another interpretation of the discontinuity of local times mentioned above. This separation of fluid and diffusion time scales is at the heart of state space collapse, see for instance Bramson [6].

It would be interesting to understand to what extent the above reasoning can be carried over to branching processes. In particular, the state space collapse is quite robust in the finite variance case and makes it possible to derive the heavy traffic limit of the PS queue with general inter-arrival times,

see Gromoll [12]. This suggests an approach to generalize the branching results of this paper to the case where the offspring process is a general renewal process.

APPENDIX A: PROOF OF LEMMA ??

In the sequel we say that a sequence of càdlàg functions (f_n) is C-relatively compact if it is relatively compact and any of its accumulation points is continuous. A straightforward adaptation of Proposition VI.3.26 in Jacod and Shiryaev [15], which gives a criterion for C-tightness, shows that a sequence (f_n) is C-relatively compact if and only if for every $m \geq 0$, $\sup_{n \geq 1} \|f_n\|_m$ is finite and

$$\lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow +\infty} w_m(h_n, \varepsilon) = 0$$

where from now on w_m is the modulus of continuity:

$$w_m(f, \varepsilon) = \sup \{|f(t) - f(s)| : 0 \leq s, t \leq m \text{ and } |t - s| \leq \varepsilon\}.$$

LEMMA A.1. *Let $f_n, f \in \mathcal{E}'$ and assume that the sequence (f_n) is C-relatively compact and that the sequence (T_{f_n}) is bounded. Then the sequence $(\mathcal{L}^{-1}(f_n))$ is C-relatively compact. If in addition $f_n \rightarrow f$ then we also have $\mathcal{L}^{-1}(f_n) \rightarrow \mathcal{L}^{-1}(f)$*

PROOF. In the rest of the proof let $h_n = \mathcal{L}^{-1}(f_n)$ and $\bar{t} = \sup_{n \geq 1} T_{f_n}$. To show that the sequence (h_n) is C-relatively compact, we show that $\sup_{n \geq 1} \|h_n\|_\infty$ is finite and that

$$(4) \quad \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow +\infty} w_\infty(h_n, \varepsilon) = 0$$

with $w_\infty(j, \delta) = \lim_{m \rightarrow +\infty} w_m(j, \delta)$ for any $j \in D$ and $\delta > 0$. By definition we have $h_n(t) = f_n(\int_0^t h_n)$ and so $\|h_n\|_\infty = \|f_n\|_\infty = \|f_n\|_{T_{f_n}} = \|f_n\|_{\bar{t}}$. Since (f_n) is C-relatively compact, $\sup_{n \geq 1} \|f_n\|_{\bar{t}}$ is finite and hence so is $\sup_{n \geq 1} \|h_n\|_\infty$.

As for $w_\infty(h_n, \varepsilon)$, we have $\int_s^t h_n \leq (t - s)\|h_n\|_\infty$ for any $0 \leq s \leq t$, and since $\|h_n\|_\infty = \|f_n\|_{\bar{t}}$ we obtain for any $0 \leq s \leq t \leq s + \varepsilon$

$$|h_n(t) - h_n(s)| = \left| f_n \left(\int_0^t h_n \right) - f_n \left(\int_0^s h_n \right) \right| \leq w_{\bar{t}}(f_n, \varepsilon \|f_n\|_{\bar{t}}).$$

Hence $w_\infty(h_n, \varepsilon) \leq w_{\bar{t}}(f_n, \varepsilon \|f_n\|_{\bar{t}})$, and so (4) follows from this inequality together with the fact that (f_n) is C-relatively compact.

We now prove that $h_n \rightarrow \mathcal{L}^{-1}(f)$ provided $f_n \rightarrow f$. Since the sequence (h_n) is C-relatively compact we only have to identify accumulation points, so

let now h be any continuous accumulation point of (h_n) and assume without loss of generality that $h_n \rightarrow h$. Let $t \geq 0$: then $h_n(t) \rightarrow h(t)$, while on the other hand from $f_n \rightarrow f$, $h_n \rightarrow h$ and the fact that f is continuous we obtain that $f_n(\int_0^t h_n) \rightarrow f(\int_0^t h)$. Since by definition $h_n(t) = f_n(\int_0^t h_n)$ this gives $h(t) = f(\int_0^t h)$ for every $t \geq 0$. Since the solution to this equation is unique because $f \in \mathcal{E}'$ we obtain that $h = \mathcal{L}^{-1}(f)$, hence the result. \square

LEMMA A.2. *Let $f_n, f \in \mathcal{E}$ and assume that $f_n \rightarrow f$, that f is continuous and that the sequence (T_{f_n}) is bounded. Then $\mathcal{L}(f_n) \rightarrow \mathcal{L}(f)$.*

PROOF. Let $c_n(t) = \int_0^t f_n$ and $c_n^{-1} : [0, \infty) \rightarrow [0, T_{f_n}]$ be such that $\int_0^{c_n^{-1}(t)} f_n = t$ for any $t < \int f_n$ and $c_n^{-1}(t) = T_{f_n}$ for $t \geq \int f_n$, and define similarly c and c^{-1} starting from f instead of f_n . Then for any $t \geq 0$, we have by definition $\mathcal{L}(f_n)(t) = f_n(c_n^{-1}(t))$ and $\mathcal{L}(f)(t) = f(c^{-1}(t))$. We are in the framework of Theorem 2.7 of Helland [14] but none of his cases applies here (notwithstanding the problem that we allow excursions to start at 0). We break the proof into two steps.

First step. Let $t < \int f$: we prove that $\|c_n^{-1} - c^{-1}\|_t \rightarrow 0$. First, note that c_n^{-1} restricted to $[0, \int f_n]$ is the inverse of c_n restricted to $[0, T_{f_n}]$, and similarly for c and c^{-1} . Moreover, Lemma 2.1 implies that $T_f \leq \liminf_n T_{f_n}$ and since $\int_0^s f_n \rightarrow \int_0^s f$ for any $s \geq 0$ this implies that $\liminf_n \int f_n \geq \int f$. Since the inverse of a continuous and strictly increasing function is a continuous mapping (see for instance Theorem 7.1 in Whitt [32]), we get that $\|c_n^{-1} - c^{-1}\|_t \rightarrow 0$.

Second step. We now prove that $\mathcal{L}(f_n) \rightarrow \mathcal{L}(f)$. Let $\bar{t} = T_f \vee \sup_n T_{f_n}$, which is finite by assumption and is such that $c_n(t) \leq \bar{t}$ and $c(t) \leq \bar{t}$ for any $t \geq 0$ and $n \geq 1$. For $t < \int f$ we write

$$\begin{aligned} \|f_n \circ c_n^{-1} - f \circ c^{-1}\|_t &\leq \|f_n \circ c_n^{-1} - f \circ c_n^{-1}\|_t + \|f \circ c_n^{-1} - f \circ c^{-1}\|_t \\ &\leq \|f_n - f\|_{\bar{t}} + w_{\bar{t}}(f, \|c_n^{-1} - c^{-1}\|_t). \end{aligned}$$

Since $f_n \rightarrow f$, f is continuous and $\|c_n^{-1} - c^{-1}\|_t \rightarrow 0$ by the first step we see that the last upper bound vanishes. Consider now some arbitrary $t' < \int f \leq t$, then

$$\begin{aligned} \|f_n \circ c_n^{-1} - f \circ c^{-1}\|_t &\leq \|f_n \circ c_n^{-1} - f \circ c^{-1}\|_{t'} + \sup_{t' \leq s \leq t} f(c^{-1}(s)) \\ &\quad + \sup_{t' \leq s \leq t} f_n(c_n^{-1}(s)). \end{aligned}$$

The first term of this upper bound goes to 0 by the first step. The last term is equal to $\sup_{[c_n^{-1}(t'), \bar{t}]} f_n$ and since $c_n^{-1}(t') \rightarrow c_n(t')$ by the previous step and the supremum is a continuous function, we get

$$\limsup_{n \rightarrow +\infty} \|f_n \circ c_n^{-1} - f \circ c^{-1}\|_t \leq 2 \sup_{c^{-1}(t') \leq s \leq T_f} f(s).$$

Letting $t' \rightarrow T_f$ achieves the proof. \square

We now prove Lemma 2.5, so consider X_n, X random elements of \mathcal{E} such that the sequence (X_n) is C-tight and the sequence (T_{X_n}) is tight. Let $(u(n))$ be any subsequence.

Assume that $X_n \Rightarrow X$: to prove that $\mathcal{L}(X_n) \Rightarrow \mathcal{L}(X)$ it is enough to find a subsequence $(v(n))$ of $(u(n))$ such that $\mathcal{L}(X_{v(n)}) \Rightarrow \mathcal{L}(X)$. Since the sequence (T_{X_n}) is tight there exists such a subsubsequence such that $(X_{v(n)}, T_{X_{v(n)}}) \Rightarrow (X', T')$ for some continuous X' equal in distribution to X and some random variable T' . Lemma A.2 together with the continuous mapping theorem implies that $\mathcal{L}(X_{v(n)}) \Rightarrow \mathcal{L}(X')$, hence the result. The other statements of the lemma follow using similar arguments, invoking Lemma A.1 instead of Lemma A.2.

REFERENCES

- [1] BERTOIN, J. (1996). *Lévy processes*. Cambridge Tracts in Mathematics, Vol. **121**. Cambridge University Press, Cambridge. [MR1406564 \(98e:60117\)](#)
- [2] BILLINGSLEY, P. (1999). *Convergence of probability measures*, Second ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York. [MR1700749 \(2000e:60008\)](#)
- [3] BLUMENTHAL, R. M. (1992). *Excursions of Markov processes*. Probability and its Applications. Birkhäuser Boston Inc., Boston, MA. [MR1138461 \(93b:60159\)](#)
- [4] BORODIN, A. N. (1981). The asymptotic behavior of local times of recurrent random walks with finite variance. *Teor. Veroyatnost. i Primenen.* **26**, 4, 769–783. [MR636771 \(83a:60120\)](#)
- [5] BORODIN, A. N. (1986). On the character of convergence to Brownian local time. I, II. *Probab. Theory Relat. Fields* **72**, 2, 231–250, 251–277. <http://dx.doi.org/10.1007/BF00699105>. [MR836277 \(87i:60087\)](#)
- [6] BRAMSON, M. (1998). State space collapse with application to heavy traffic limits for multiclass queueing networks. *Queueing Systems: Theory and Applications* **30**, 1–2 (November), 89–140.
- [7] CABALLERO, M. E., LAMBERT, A., AND URIBE BRAVO, G. (2009). Proof(s) of the Lamperti representation of continuous-state branching processes. *Probab. Surv.* **6**, 62–89. <http://dx.doi.org/10.1214/09-PS154>. [MR2592395](#)
- [8] CSÖRGŐ, M. AND RÉVÉSZ, P. (1985). On strong invariance for local time of partial sums. *Stochastic Process. Appl.* **20**, 1, 59–84. [http://dx.doi.org/10.1016/0304-4149\(85\)90017-1](http://dx.doi.org/10.1016/0304-4149(85)90017-1). [MR805116 \(87b:60115\)](#)

- [9] ETHIER, S. N. AND KURTZ, T. G. (1986). *Markov Processes, Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics.
- [10] GRIMVALL, A. (1973). On the transition from a Markov chain to a continuous time process. *Stochastic Processes Appl.* **1**, 335–368. [MR0372933 \(51 #9137\)](#)
- [11] GRISHECHKIN, S. (1992). On a Relationship between Processor-Sharing Queues and Crump-Mode-Jagers Branching Processes. *Advances in Applied Probability* **24**, 3, 653–698. <http://www.jstor.org/stable/1427484>.
- [12] GROMOLL, H. C. (2004). Diffusion approximation for a processor sharing queue in heavy traffic. *Ann. Appl. Probab.* **14**, 2, 555–611. <http://dx.doi.org/10.1214/105051604000000035>. [MR2052895 \(2005b:60239\)](#)
- [13] GROMOLL, H. C., PUHA, A. L., AND WILLIAMS, R. J. (2002). The fluid limit of a heavily loaded processor sharing queue. *Ann. Appl. Probab.* **12**, 3, 797–859. <http://dx.doi.org/janus.lib.tue.nl/10.1214/aoap/1031863171>. [MR1925442 \(2003h:60135\)](#)
- [14] HELLAND, I. S. (1978). Continuity of a class of random time transformations. *Stochastic Processes Appl.* **7**, 1, 79–99. [MR0488203 \(58 #7765\)](#)
- [15] JACOD, J. AND SHIRYAEV, A. N. (2003). *Limit theorems for stochastic processes*, Second ed. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. **288**. Springer-Verlag, Berlin. [MR1943877 \(2003j:60001\)](#)
- [16] KALLENBERG, O. (2002). *Foundations of modern probability*, Second ed. Probability and its Applications (New York). Springer-Verlag, New York. [MR1876169 \(2002m:60002\)](#)
- [17] KELLY, F. P. (1979, reprinted 1987, 1994). *Reversibility and Stochastic Networks*. Wiley, Chichester.
- [18] KENDALL, D. G. (1951). Some problems in the theory of queues. *Journal of the Royal Statistical Society. Series B (Methodological)* **13**, 2, 151–185. <http://www.jstor.org/stable/2984059>.
- [19] KHOSHNEVISAN, D. (1993). An embedding of compensated compound Poisson processes with applications to local times. *Ann. Probab.* **21**, 1, 340–361. <http://www.jstor.org/stable/2244765>. [MR1207230 \(94d:60121\)](#)
- [20] KITAYEV, M. Y. AND YASHKOV, S. F. (1979). Analysis of a single-channel queueing system with the discipline of uniform sharing of a device. *Engineering Cybernetics* **17**, 42–49. [MR598335 \(82d:60184\)](#)
- [21] KNIGHT, F. B. (1963). Random walks and a sojourn density process of Brownian motion. *Trans. Amer. Math. Soc.* **109**, 56–86. [MR0154337 \(27 #4286\)](#)
- [22] LAMBERT, A. (2010). The contour of splitting trees is a Lévy process. *Ann. Probab.* **38**, 1, 348–395.
- [23] LAMPERTI, J. (1967a). Continuous state branching processes. *Bull. Amer. Math. Soc.* **73**, 382–386. [MR0208685 \(34 #8494\)](#)
- [24] LAMPERTI, J. (1967b). The limit of a sequence of branching processes. *Z. Wahrsch. Verw. Gebiete* **7**, 271–288. [MR0217893 \(36 #982\)](#)
- [25] LIMIC, V. (2001). A LIFO queue in heavy traffic. *Ann. Appl. Probab.* **11**, 2, 301–331. [MR1843048 \(2002e:60151\)](#)
- [26] PARDOUX, E. AND WAKOLBINGER, A. (2011). From exploration paths to mass excursions – variations on a theme of Ray and Knight. In *Surveys in Stochastic Processes*, J. Blath, P. Imkeller, and S. Roelly, Eds. EMS Series of Congress Reports. 87–106.

- [27] PERKINS, E. (1982). Weak invariance principles for local time. *Z. Wahrsch. Verw. Gebiete* **60**, 4, 437–451. <http://dx.doi.org/10.1007/BF00535709>. MR665738 (83h:60080)
- [28] RAY, D. (1963). Sojourn times of diffusion processes. *Illinois J. Math.* **7**, 615–630. MR0156383 (27 #6306)
- [29] ROBERT, P. (2003). *Stochastic Networks and Queues*. Stochastic Modelling and Applied Probability Series. Springer-Verlag, New York. xvii+398 pp.
- [30] SAGITOV, S. (1995). A key limit theorem for critical branching processes. *Stochastic Process. Appl.* **56**, 1, 87–100. [http://dx.doi.org/10.1016/0304-4149\(94\)00046-V](http://dx.doi.org/10.1016/0304-4149(94)00046-V). MR1324323 (96d:60134)
- [31] SAGITOV, S. M. (1994). General branching processes: convergence to Irzhina processes. *J. Math. Sci.* **69**, 4, 1199–1206. Stability problems for stochastic models (Kirillov, 1989). MR1284307
- [32] WHITT, W. (1980). Some useful functions for functional limit theorems. *Math. Oper. Res.* **5**, 1, 67–85. <http://dx.doi.org/10.1287/moor.5.1.67>. MR561155 (81e:60035)
- [33] WILLIAMS, R. J. (1998). Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems: Theory and Applications* **30**, 1–2 (November), 27–88.

LABORATOIRE DE PROBABILITÉS ET MODÈLES ALÉATOIRES
 UMR 7599 CNRS AND UPMC UNIV PARIS 06
 CASE COURRIER 188
 4 PLACE JUSSIEU
 F-75252 PARIS CEDEX 05, FRANCE
 E-MAIL: amaury.lambert@upmc.fr

CWI
 123 SCIENCE PARK
 1098 XG, AMSTERDAM
 THE NETHERLANDS
 E-MAIL: florian.simatos@cw.nl

CWI
 123 SCIENCE PARK
 1098 XG, AMSTERDAM
 THE NETHERLANDS
 E-MAIL: bert.zwart@cw.nl