



**University of  
Zurich** <sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2006

---

## **Regenerative compositions in the case of slow variation**

Barbour, A D ; Gnedin, A

DOI: <https://doi.org/10.1016/j.spa.2005.12.006>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-21591>

Journal Article

Accepted Version

Originally published at:

Barbour, A D; Gnedin, A (2006). Regenerative compositions in the case of slow variation. *Stochastic Processes and their Applications*, 116(7):1012-1047.

DOI: <https://doi.org/10.1016/j.spa.2005.12.006>

# Regenerative Compositions in the Case of Slow Variation

A.D. Barbour\* and A.V. Gnedin†

Universität Zürich and Universiteit Utrecht

## Abstract

For  $S$  a subordinator and  $\Pi_n$  an independent Poisson process of intensity  $ne^{-x}$ ,  $x > 0$ , we are interested in the number  $K_n$  of gaps in the range of  $S$  that are hit by at least one point of  $\Pi_n$ . Extending previous studies in [7, 10, 11] we focus on the case when the tail of the Lévy measure of  $S$  is slowly varying. We view  $K_n$  as the terminal value of a random process  $\mathcal{K}_n$ , and provide an asymptotic analysis of the fluctuations of  $\mathcal{K}_n$ , as  $n \rightarrow \infty$ , for a wide spectrum of situations.

## 1 Introduction

Let  $S = (S_t, t \geq 0)$  be an increasing Lévy process (subordinator) with  $S_0 = 0$ , zero drift and no killing. The closed range  $\mathcal{R}$  of  $S$  has zero Lebesgue measure, and defines a random division of the complement set  $\mathbb{R}_+ \setminus \mathcal{R}$  into open interval components, referred to as *gaps*. In this paper, we are concerned with the distribution of the number  $K_n$  of gaps hit by at least one point of an independent Poisson process  $\Pi_n$  with the inhomogeneous rate  $ne^{-x}$ ,  $x > 0$ , where  $n$  is a large parameter. We actually go into more detail. We view  $K_n$  as the terminal value  $\mathcal{K}_n(\infty)$  of the increasing process  $\mathcal{K}_n = (\mathcal{K}_n(T), T \geq 0)$ , where  $\mathcal{K}_n(T)$  is defined to be the number of jumps of the subordinator within  $[0, T]$  which cover one or more Poisson points; that is,  $\mathcal{K}_n(T)$  counts all instants  $t \in [0, T]$  which satisfy  $\Pi_n \cap ]S_{t-}, S_t[ \neq \emptyset$ . Our aim is to describe the random fluctuations of the process  $\mathcal{K}_n$ .

The general motivation for the setting stems from the study of the number of blocks in a random decomposable combinatorial structure, which in the case under focus is a *composition* (ordered partition)  $\mathcal{C}_n$  of some integer. Viewing  $\mathcal{C}_n$  as distribution of some number of balls in some collection of boxes, each gap may be interpreted as a box, which is hit by a particular ball with probability equal to the exponential measure of the gap. The parameter  $n$  controls the total number of balls, which is a Poisson variable, and the composition  $\mathcal{C}_n$  of this random number is defined as the consecutive record of nonzero occupancy numbers, in the natural ordering of the gaps. Our study fits into the recent

---

\*Angewandte Mathematik, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland; supported in part by Schweizer Nationalfonds Projekt Nr. 20-107935/1

†Mathematisch Instituut, PO Box 80010, 3508 TA Utrecht, Nederland

theory of sampling models called *regenerative composition structures*, which have a distinguished Markovian property resulting from the renewal features of  $\mathcal{R}$  combined with that of the exponential distribution [8, 9]. Concretely, the regeneration property of  $\mathcal{C}_n$  means that, for each  $t > 0$ , conditionally given the value  $s = S_t$ , the partial compositions appearing within  $[0, s]$  and  $[s, \infty[$  are independent and the latter has the same distribution as the composition  $\mathcal{C}_{ne^{-s}}$ .

The distribution of  $S$  is completely determined by a Lévy measure  $\nu_0$  on  $\mathbb{R}_+$ , which describes the intensity of the jumps of different sizes, and the behaviour of  $\mathcal{K}_n$  depends very much on the form of  $\nu_0$ . Qualitatively different modes of behaviour are known [7, 10, 11].

For  $\nu_0$  a finite measure,  $S$  is a compound Poisson process. Under mild additional assumptions, the two central moments are of the order of  $\log n$  and  $K_n$  is asymptotically normal. In this situation, the methods of renewal theory are adequate, since the process  $\mathcal{K}_n(T)$  essentially coincides with the process of jump epochs of  $S$  for  $T < \log n$ , while the contribution of larger times  $T > \log n$  to  $K_n$  is negligible, see [7]. In particular, when  $\nu_0$  is an exponential distribution, the induced composition follows the poissonised (ordered) Ewens sampling formula, in which case much finer results on  $K_n$  are available by combinatorial methods [1, 14].

If  $\nu_0$  is infinite, and its tail  $N_0(x) := \nu_0[x, \infty[$  is such that  $N(1/y)$  is regularly varying as  $y \rightarrow \infty$  with exponent  $\alpha$  (here and henceforth this means regular variation with  $0 < \alpha \leq 1$ ), then  $\mathbb{E}K_n$  is also regularly varying with the same exponent and  $K_n/\mathbb{E}K_n$  approaches a nondegenerate limit, which is not gaussian. The moments of  $\mathcal{K}_n(T)$  are then of the same order of magnitude as that of  $K_n$ , for each fixed  $T$ , see [10].

Between these two possibilities lies the setting in which  $N_0(1/y)$  is slowly varying as  $y \rightarrow \infty$ , but the Lévy measure is infinite, *i.e.*  $\lim_{y \rightarrow \infty} N_0(1/y) = \infty$ . Here, the special case with  $N_0(1/y) \sim c \log y$  has been studied in some detail. For these *gamma-like* subordinators, the proper formats for the two central moments of  $K_n$  are  $\log^2 n$  and  $\log^3 n$ , respectively, and the limiting distribution is again normal, see [11].

In this paper, we treat the case of slowly varying  $N_0$  in greater generality. As might be expected of a transitional régime between the finite and the regularly varying cases, there is a further wealth of possible modes of behaviour, and the discussion reveals how these are related to the time scales over which the significant variation in  $\mathcal{K}_n$  occurs. Our argument leading to a functional central limit theorem is very different from that in [7, 11], and is based on the observation that, to first order, the fluctuations of the counting process  $\mathcal{K}_n$  are dominated by those of its compensator  $A_n$ , defined in Proposition 2.1. The explicit representation of the random process  $A_n$  makes it possible to find approximations by rather direct arguments, and under relatively mild conditions. These are broadly speaking of two kinds. The first is expressed in Assumption A2, which puts a mild restriction on the way in which a certain transform  $L$  of the measure  $\nu_0$  can vary locally as a function of its parameter. Conditions of the second kind, appearing in different forms in (4.9), (4.19), (4.20) and (5.1), limit the global variability of  $L$ .

Our analysis of subordinators with slowly varying  $N_0$  distinguishes three basic modes. In the case of *moderate growth*, which includes the subordinators with logarithmic asymptotics  $N_0(1/y) \asymp (\log y)^\beta$ ,  $\beta > 0$ , including the gamma-like subordinators studied in [11], the random fluctuations of  $\mathcal{K}_n(T)$  occur more or less evenly on the scale  $T = v \log n$  in

$v \in [0, 1]$ . In the case of *fast growth*, well exemplified by  $N_0(1/y) \asymp \exp(\log^\beta y)$ ,  $0 < \beta < 1$ , almost everything happens at times of order  $L(n)$ , and  $L(n)$  is of smaller order than  $\log n$ . The third case is that of *slow growth*, as for example  $N_0(1/y) \asymp \log \log y$ , when significant contributions to the random fluctuations of  $K_n$  are only made at times very close to  $\log n$ , just as in the compound Poisson case [7].

**Notation.** We use  $\lambda, \lambda_j$  for positive constants whose value is not important and may depend on the context. The asymptotic relation  $a_n \asymp b_n$  means that  $a_n = O(b_n)$  and  $b_n = O(a_n)$ , while  $a_n \gg b_n$  means that  $b_n = o(a_n)$ . Asymptotic relations like  $X_n \sim Y_n$  or  $X_n \asymp Y_n$  for random quantities mean that they hold with probability one, unless otherwise specified.

## 2 The basic setting

### 2.1 Laplace exponents and the compensator

The Lévy measure  $\nu_0$  is uniquely determined by the Laplace exponent  $\Phi_0$ , defined for  $m \geq 0$  by

$$\Phi_0(m) = \int_0^\infty (1 - e^{-mx}) \nu_0(dx) = m \int_0^\infty e^{-mx} N_0(x) dx;$$

note that  $\nu_0$  must satisfy  $\Phi_0(1) < \infty$ . The distribution of the subordinator is determined by the Lévy-Khintchine formula for the Laplace transform

$$\mathbb{E} e^{-nS_t} = e^{-t\Phi_0(n)}, \quad n \geq 0, \quad (2.1)$$

see [2] as a general reference on the Lévy processes and see [3] especially for subordinators.

The function  $\Phi_0$  can be extended to an analytic function in the right half-plane, and hence, by Müntz's theorem,  $\Phi_0$  can be uniquely extrapolated from the values  $\Phi_0(m)$ ,  $m = 1, 2, \dots$ ; these also determine the poissonised version of  $\Phi_0$ , defined either by the series

$$\Phi(n) := e^{-n} \sum_{m=1}^{\infty} \frac{n^m}{m!} \Phi_0(m), \quad (2.2)$$

or by the integral

$$\Phi(n) = \int_0^\infty (1 - e^{-n(1-e^{-x})}) \nu_0(dx). \quad (2.3)$$

This latter transform is particularly useful to us, since it appears naturally in the definition of the compensator  $A_n$  of the counting process  $\mathcal{K}_n$ .

**Proposition 2.1** *With respect to the filtration  $(\mathcal{F}_{T,n}, T \geq 0)$ , defined by*

$$\mathcal{F}_{T,n} := \sigma \{ S_t, 0 \leq t \leq T; \Pi_n|_{[0, S_T]} \},$$

*the compensator of  $\mathcal{K}_n$  is the increasing process  $A_n$  given by the formula*

$$A_n(T) := \int_0^T \Phi(ne^{-S_t}) dt, \quad T \in [0, \infty]. \quad (2.4)$$

**Proof.** The subordinator gains an increment within  $[x, x + dx]$  at rate  $\nu_0(dx)$ . On the other hand,  $\Pi_n$  hits  $[S_{t-}, S_{t-} + x]$  with probability  $1 - \exp(-ne^{-s}(1 - e^{-x}))$  (where  $s = S_{t-}$ ), because the number of atoms in  $[s, s + x]$  has Poisson distribution with mean

$$n \int_s^{s+x} e^{-u} du = ne^{-s}(1 - e^{-x}).$$

Integrating over  $x$  yields the derivative  $dA_n(t)/dt = \Phi(ne^{-S_t})$ .  $\square$

We further assume that

$$\mathbb{E}S_t = t \quad \text{and} \quad \text{Var} S_t = t\sigma^2.$$

The former is the same as

$$\mathbb{E}S_1 = \Phi'_0(0) = \int_0^\infty x \nu_0(dx) = \int_0^\infty N_0(x) dx = 1. \quad (2.5)$$

This can always be achieved by a linear time-scaling, which does not affect the range  $\mathcal{R}$ . A consequence of the assumption is that  $N_0(x)$  is a probability density. It then follows for all  $m \geq 0$  that

$$\Phi_0(m) \leq m \int_0^\infty N_0(x) dx = m \quad \text{and} \quad \Phi(m) \leq m, \quad (2.6)$$

this last from (2.2).

With this scaling,  $N_0(x)$  is the density of a *delay* variable. If  $X$  has this density and is independent of  $S$ , then the process  $(X + S_t, t \geq 0)$  is a *stationary* subordinator, in the sense that its closed range  $X + \mathcal{R}$  may be extended to a random subset of  $\mathbb{R}$  invariant under all translations. In particular,  $X + S_t$  has the same overshoot distribution at every level  $s \geq 0$ .

It will be convenient to define all the Poisson processes  $(\Pi_n, n \geq 0)$  (which are independent of the subordinator  $S$ ) consistently on the same probability space. To this end, we take an inhomogeneous planar Poisson point process on  $\mathbb{R}_+^2$  with intensity measure  $e^{-y} dy dn$ , and we introduce  $\Pi_n$  as the projection on the  $y$ -axis of the planar process restricted to the strip  $[0, \infty] \times [0, n]$ . In this setting, the compositions  $\mathcal{C}_n$  are defined consistently for all  $n \geq 0$ : a decrease in  $n$  has the effect of thinning, *i.e.* removing some balls from the boxes; while as  $n$  increases more Poisson atoms are added, hence  $\mathcal{K}_n(T)$  and  $K_n$  are nondecreasing in  $n$ . Thus, in principle, our setting is 3-dimensional, with three parameters  $n, t, s$  meaning the intensity, the time and the range of subordinator.

For our analysis of  $A_n$ , it is also convenient to note that we can truncate the integral (2.4), which defines  $A_n(T)$ , at the first passage time

$$\tau_n := \min\{t: S_t \geq \log n\}, \quad (2.7)$$

with little loss.

**Lemma 2.2** *The jumps of  $S$  after  $\tau_n$  make only a bounded contribution to  $\mathcal{K}_n(T)$ , uniformly in  $T \leq \infty$ , and for  $\psi > 0$*

$$\mathbb{P} \left[ \sup_{T \geq 0} |A_n(T) - A_n(T \wedge \tau_n)| > \psi \right] \leq \frac{1}{\psi \Phi_0(1)}.$$

**Proof.**  $K_n - \mathcal{K}_n(\tau_n)$  cannot exceed the number of atoms of  $\Pi_n \cap [\log n, \infty[$ , which is Poisson distributed with mean 1. To estimate the contribution to the compensator, recalling (2.6), we have

$$\int_{\tau_n}^{\infty} \Phi(ne^{-S_t}) dt < n \int_{\tau_n}^{\infty} e^{-S_t} dt = ne^{-S_{\tau_n}} \int_0^{\infty} e^{-S'_u} du \leq \int_0^{\infty} e^{-S'_u} du,$$

where  $S'$  defined by  $S'_u := S_{\tau_n+u} - S_{\tau_n}$ ,  $u \geq 0$ , has the same distribution as  $S$ . To complete the proof use (2.1) and Markov's inequality.  $\square$

## 2.2 Slow variation

Our aim is to investigate the process  $\mathcal{K}_n$  in the intermediate setting, between that in which the tail  $N_0$  of  $\nu_0$  is regularly varying (with exponent  $0 < \alpha \leq 1$ ), and that in which it is bounded. We therefore assume that

**Assumption A1:**  $N_0(1/y)$  is an unbounded function of slow variation for  $y \rightarrow \infty$ . (2.8)

The condition can be equally stated in terms of  $\Phi_0$ , because by the Abel–Tauber theorem [6]:

$$\Phi_0(m) \sim N_0(1/m), \quad \text{as } m \rightarrow \infty.$$

In what follows, we prefer to work in terms of  $\Phi$ , because of (2.4), so that A1 is then more naturally expressed in the equivalent form:  $\Phi(m)$  is unbounded and slowly varying at infinity. While this equivalence is more or less clear from (2.2), it is useful for Section 3 to have a better idea of how close the functions  $\Phi$  and  $\Phi_0$  are to each other. To this end, we define a measure  $\nu$  on  $[0, 1]$  as the pushforward of  $\nu_0$  under the change of variable  $x \rightarrow 1 - e^{-x}$ . Then we have

$$\Phi_0(m) = \int_0^1 (1 - (1 - x)^m) \nu(dx) = m \int_0^1 (1 - x)^{m-1} N(x) dx,$$

where  $N(x) := \nu[x, 1]$ , and

$$\Phi(m) = \int_0^1 (1 - e^{-mx}) \nu(dx) = m \int_0^1 e^{-mx} N(x) dx.$$

So the substitution transforms a Laplace integral into a Mellin integral, while  $\Phi$  assumes the conventional form of a Laplace exponent (hence  $\Phi$  also corresponds to some subordinator, whose jump-sizes do not exceed 1). We can now use these representations to prove the following lemma.

**Lemma 2.3** *Under Assumption A1, we have*

$$|\Phi_0^{(l)}(m) - \Phi^{(l)}(m)| = o\left(\frac{\Phi(m)}{m^{l+1}}\right), \quad l \geq 0,$$

where  $f^{(l)}$  denotes the  $l$ th derivative of  $f$ .

**Proof.** Because  $N_0(1/y)$  is slowly varying as  $y \rightarrow \infty$ , the same is true of  $N(1/y)$ , and  $N(1/m) \sim \Phi(m)$ . Hence we immediately have

$$\Phi^{(l)}(m) = O\left(\frac{\Phi(m)}{m^l}\right), \quad l \geq 0. \quad (2.9)$$

We now define

$$D_0(m) := m^{-1}\{\Phi(m) - \Phi_0(m)\} = \int_0^1 \{e^{-mx} - e^{(m-1)\log(1-x)}\} N(x) dx.$$

Since

$$\begin{aligned} \int_0^1 \{e^{-mx} - e^{m\log(1-x)}\} N(x) dx &= \int_0^1 e^{-mx} (1 - \exp\{m[\log(1-x) + x]\}) N(x) dx \\ &\sim \int_0^1 \frac{1}{2} m x^2 e^{-mx} N(x) dx \sim m^{-2} \Phi(m), \end{aligned}$$

and

$$\int_0^1 e^{m\log(1-x)} \{1 - (1-x)^{-1}\} N(x) dx \sim - \int_0^1 x e^{-mx} N(x) dx \sim -m^{-2} \Phi(m),$$

it then follows that  $|D_0(m)| = o(m^{-2}\Phi(m))$ .

For the  $l$ th derivative  $D_l$  of  $D_0$ , we similarly have

$$D_l(m) = \int_0^1 \{(-x)^l e^{-mx} - \{\log(1-x)\}^l e^{(m-1)\log(1-x)}\} N(x) dx.$$

Once again, since

$$\begin{aligned} &\int_0^1 (-x)^l \{e^{-mx} - e^{m\log(1-x)}\} N(x) dx \\ &= (-1)^l \int_0^1 x^l e^{-mx} (1 - \exp\{m[\log(1-x) + x]\}) N(x) dx \\ &\sim (-1)^l \int_0^1 \frac{1}{2} m x^{l+2} e^{-mx} N(x) dx \sim (-1)^l \frac{\Phi(m)}{m^{l+2}} \frac{(l+2)!}{2}, \end{aligned}$$

and

$$\begin{aligned} &\int_0^1 e^{m\log(1-x)} \{(-x)^l - \{\log(1-x)\}^l / (1-x)\} N(x) dx \\ &\sim - \int_0^1 (-x)^{l+\frac{1}{2}} (l+2) x e^{-mx} N(x) dx \\ &\sim \frac{1}{2} (l+2) (-1)^{l+1} (l+1)! m^{-l-2} \Phi(m), \end{aligned}$$

it follows that  $|D_l(m)| = o(m^{-l-2}\Phi(m))$  also. The lemma now follows by expressing the differences  $\Phi_0^{(l)}(m) - \Phi^{(l)}(m)$  using the quantities  $(D_j(m), 0 \leq j \leq l)$ .  $\square$

In fact, the measure  $\nu$  is an object in its own right: it is the Lévy measure of the *geometric* or *multiplicative* subordinator  $\widehat{S}_t = 1 - \exp(-S_t)$ . In terms of  $\widehat{S}$  the composition  $\mathcal{C}_n$  is defined as a record of occupancy counts for the gaps in the range of  $\widehat{S}$  that are hit by at least one atom of a homogeneous Poisson process on  $[0, 1]$  with rate  $n$ .

## 2.3 Law of large numbers

For a law of large numbers, we begin by noting that

$$\mathbb{E}K_n = \mathbb{E}A_n(\infty) = \int_0^\infty \mathbb{E}\Phi(ne^{-S_t}) dt = \int_0^\infty \Phi(ne^{-s})U(ds) \quad (2.10)$$

where  $U$  is the potential measure of  $S$  (i.e.  $U[0, s]$  is the expected time  $S$  stays below  $s$ ). Now, since  $\Phi$  is slowly varying, it is plausible that

$$\mathbb{E}\Phi(ne^{-S_t}) \sim \Phi(ne^{-\mathbb{E}S_t}) = \Phi(ne^{-t}).$$

This motivates the introduction of

$$\Psi(n) := \int_0^\infty \Phi(ne^{-t}) dt = \int_0^n \Phi(t) \frac{dt}{t}, \quad (2.11)$$

as an approximation to  $K_n$ ; under A1 it follows that  $\Psi(n) \gg \log n$ . Our aim in this section is to show that in fact  $K_n \sim \Psi(n)$  for large  $n$ .

**Lemma 2.4** *Assumption A1 implies  $\Psi(m) \gg \Phi(m)$ .*

**Proof.** This is a standard consequence of the uniform convergence theorem [4], which states that slow variation implies  $\Phi(mu)/\Phi(m) \rightarrow 1$ , uniformly in  $u$  bounded away from 0 and  $\infty$ .  $\square$

It hence follows that, for any fixed  $T$ , and as  $n \rightarrow \infty$ ,

$$\frac{A_n(T)}{\Psi(n)} < \frac{T\Phi(n)}{\Psi(n)} \rightarrow 0 \quad \text{a.s.} \quad (2.12)$$

and also that

$$\frac{\int_0^T \Phi(ne^{-t}) dt}{\Psi(n)} < \frac{T\Phi(n)}{\Psi(n)} \rightarrow 0. \quad (2.13)$$

The convergence in (2.12) and (2.13) also holds if the fixed time  $T$  is replaced by an a.s. finite random time  $\tau$  which is measurable with respect to  $\sigma\{S_t, t \geq 0\}$ .

The next lemma explores the error caused by replacing  $U$  with the Lebesgue measure in (2.10).

**Lemma 2.5** *For an arbitrary subordinator  $S$  with  $\mathbb{E}S_1^2 < \infty$ ,*

$$\Psi(n) \leq \mathbb{E}K_n \leq \Psi(n) + \lambda\Phi(n) \quad (2.14)$$

for some constant  $\lambda > 0$ . It follows that, under A1,  $\mathbb{E}K_n/\Psi(n) \rightarrow 1$  as  $n \rightarrow \infty$ .

**Proof.** By renewal theory, there is a constant  $\lambda > 0$  such that  $s \leq U[0, s] < s + \lambda$  for  $s \geq 0$ . The bounds (2.14) follow from this, (2.10) and the monotonicity of  $\Phi$ . The convergence of  $\mathbb{E}K_n/\Psi(n)$  is now a consequence of Lemma 2.4.  $\square$

We now show that  $K_n$  is close to  $A_n(\infty)$ .



**Lemma 2.6** *Under Assumption A1*

$$\mathbb{E}(K_n - A_n(\infty))^2 \sim \Psi(n).$$

Furthermore, considering the whole path of  $\mathcal{K}_n - A_n$ , we have

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[ \sup_{0 \leq T \leq \infty} |\mathcal{K}_n(T) - A_n(T)| > b_n \right] = 0,$$

for all sequences  $b_n$  such that  $b_n^{-2} \Psi(n) \rightarrow 0$ .

**Proof.** The difference  $\mathcal{K}_n(T) - A_n(T)$  is a square-integrable martingale of locally bounded variation with respect to the filtration  $\mathcal{F}_{T,n}$ , and has all jumps of size 1; this yields the formula [12, Section 15.2]

$$\mathbb{E}(K_n - A_n(\infty))^2 = \mathbb{E}A_n(\infty) \sim \Psi(n), \quad (2.15)$$

proving the first part. The second follows from Kolmogorov's inequality, which gives

$$\mathbb{P} \left[ \sup_{0 \leq T \leq \infty} |\mathcal{K}_n(T) - A_n(T)| > b_n \right] \leq b_n^{-2} \mathbb{E}A_n(\infty). \quad \square \quad (2.16)$$

Next is the law of large numbers for the compensator.

**Proposition 2.7** *Under Assumption A1, as  $n \rightarrow \infty$ ,  $A_n(\infty)/\Psi(n) \rightarrow 1$  almost surely and in the mean.*

**Proof.** Fix  $\varepsilon$ , and define

$$\tau := \sup\{t \geq 0 : t^{-1}|S_t - t| > \varepsilon\},$$

finite almost surely. By the monotonicity of  $\Phi$ ,

$$\begin{aligned} (1 + \varepsilon)^{-1} \int_{(1+\varepsilon)\tau}^{\infty} \Phi(ne^{-s}) ds &= \int_{\tau}^{\infty} \Phi(ne^{-(1+\varepsilon)t}) dt < \int_{\tau}^{\infty} \Phi(ne^{-S_t}) dt \\ &< \int_{\tau}^{\infty} \Phi(ne^{-(1-\varepsilon)t}) dt = (1 - \varepsilon)^{-1} \int_{(1-\varepsilon)\tau}^{\infty} \Phi(ne^{-s}) ds. \end{aligned}$$

Dividing by  $\Psi(n)$  and using (2.12) and (2.13), we make the sandwich

$$\frac{1}{1 + \varepsilon} < \liminf_{n \rightarrow \infty} \frac{A_n(\infty)}{\Psi(n)} \leq \limsup_{n \rightarrow \infty} \frac{A_n(\infty)}{\Psi(n)} < \frac{1}{1 - \varepsilon} \quad \text{a.s.},$$

and now let  $\varepsilon \rightarrow 0$  to obtain almost sure convergence. Convergence in the mean then follows from  $\mathbb{E}K_n = \mathbb{E}A_n(\infty)$  and Lemma 2.5, together with the fact that  $A_n(\infty)/\Psi(n) \geq 0$  a.s.  $\square$

Finally, we have all ingredients to establish the law of large numbers for  $K_n$ .

**Theorem 2.8** *As  $n \rightarrow \infty$ ,  $K_n/\Psi(n) \rightarrow 1$  almost surely and in the mean.*

**Proof.** It follows from Lemma 2.6 that

$$\mathbb{E} \left( \frac{K_n}{\Psi(n)} - \frac{A_n(\infty)}{\Psi(n)} \right)^2 = \frac{1}{\Psi(n)} \mathbb{E} \left( \frac{A_n(\infty)}{\Psi(n)} \right) \sim \frac{1}{\Psi(n)},$$

and convergence in the mean follows from Proposition 2.7. Then, because  $\Psi(n)$  is increasing, continuous and unbounded, we can select  $n_j$  to satisfy  $\Psi(n_j) = j^2$ . Then from

$$\sum_{j=1}^{\infty} \left( \frac{K_{n_j}}{\Psi(n_j)} - \frac{A_{n_j}(\infty)}{\Psi(n_j)} \right)^2 < \infty$$

we conclude, in a standard way, that  $K_{n_j}/\Psi(n_j) \rightarrow 1$  a.s. along the subsequence. Now, because both  $K_n$  and  $\Psi(n)$  are increasing in  $n$ , the inequalities

$$\frac{K_{n_j}}{\Psi(n_{j+1})} \leq \frac{K_n}{\Psi(n)} \leq \frac{K_{n_{j+1}}}{\Psi(n_j)}$$

hold for  $n_j \leq n \leq n_{j+1}$ . The convergence almost surely follows from these relations and the trivial fact that  $\Psi(n_j)/\Psi(n_{j+1}) \rightarrow 1$ .  $\square$

Along the same lines,  $K_n \gg \mathcal{K}_n(T)$  for each fixed  $T$ . This property can be shown to be characteristic of the slow variation case.

## 2.4 The variance

The aim of this section is to derive asymptotics of the variance of  $K_n$  and  $A_n(\infty)$ . We start with the explicit formula

$$\text{Var } A_n(\infty) = 2 \int_0^\infty \Phi(ne^{-s})U(ds) \int_0^\infty \Phi(ne^{-s-v})d_v\{U[0, v] - U[s, s + v]\}, \quad (2.17)$$

where  $d_v$  indicates the active variable of integration. The formula is derived by using the representation of the path past  $t$  as  $(S_u, u \geq t) =_d (S_t + S'_u, u \geq 0)$  with  $S'$  an independent

copy of  $S$ , and using a familiar symmetrisation trick for squared integrals:

$$\begin{aligned}
\text{Var } A_n(\infty) &= \mathbb{E} \left( \int_0^\infty \Phi(ne^{-S_t}) dt \right)^2 - \left( \int_0^\infty \Phi(ne^{-s}) U(ds) \right)^2 \\
&= 2 \mathbb{E} \left( \int_0^\infty \Phi(ne^{-S_t}) dt \int_t^\infty \Phi(ne^{-S_u}) du \right) \\
&\quad - 2 \int_0^\infty \Phi(ne^{-s}) U(ds) \int_s^\infty \Phi(ne^{-v}) U(dv) \\
&= 2 \mathbb{E} \left( \int_0^\infty \Phi(ne^{-S_t}) dt \int_0^\infty \Phi(ne^{-S_t - S'_u}) du \right) \\
&\quad - 2 \int_0^\infty \Phi(ne^{-s}) U(ds) \int_0^\infty \Phi(ne^{-s-v}) U(s+dv) \\
&= 2 \int_0^\infty \Phi(ne^{-s}) U(ds) \int_0^\infty \Phi(ne^{-s-v}) U(dv) \\
&\quad - 2 \int_0^\infty \Phi(ne^{-s}) U(ds) \int_0^\infty \Phi(ne^{-s-v}) U(s+dv) \\
&= 2 \int_0^\infty \Phi(ne^{-s}) U(ds) \int_0^\infty \Phi(ne^{-s-v}) d_v \{U[0, v] - U[s, s+v]\}.
\end{aligned}$$

Next is a more informative asymptotic formula, very much in the spirit of the asymptotics for the expectation  $\mathbb{E}K_n \sim \Psi(n)$  derived before. To state it, we first define

$$\Psi_2(n) := \int_0^\infty \Phi^2(ne^{-s}) ds = \int_0^n \Phi^2(s) \frac{ds}{s}. \quad (2.18)$$

**Lemma 2.9** *For a subordinator such that  $\Phi$  is slowly varying at infinity,*

$$\text{Var } A_n(\infty) \sim \sigma^2 \Psi_2(n). \quad (2.19)$$

*If also A1 holds, then the same asymptotics hold for  $K_n$ :*

$$\text{Var } K_n \sim \text{Var } A_n(\infty) \sim \sigma^2 \Psi_2(n).$$

**Proof.** We aim to evaluate the integral (2.17). Renewal theory tells us that

$$0 \leq U[0, s] - s \rightarrow \sigma^2/2,$$

where the constant appears as the mean value of the delay variable  $X$ :

$$\mathbb{E}X = \int_0^\infty x N_0(x) dx = \frac{1}{2} \int_0^\infty x^2 \nu_0(dx) = \frac{1}{2} \text{Var } S_1 = \frac{\sigma^2}{2}$$

(the third equality follows from (2.1)). This general fact holds for arbitrary square-integrable subordinators, and it follows easily from the compound Poisson case treated in [6].

We can therefore, for  $\varepsilon$  given, select  $s_0$  and  $v_0$  so large that, for  $s > s_0$  and  $v > v_0$ ,

$$|U[0, v_0] - U[s, s+v_0] - \sigma^2/2| < \varepsilon; \quad |U[v_0, v] - U[s+v_0, s+v]| < \varepsilon. \quad (2.20)$$

Now, writing  $G_s(v) = U[0, v] - U[s, s + v]$ , we have

$$G_s(0) = 0, \quad \|G_s\| := \sup_{v>0} |G_s(v)| \leq \sigma^2 \quad \text{and} \quad |G_s(v) - \sigma^2/2| < \varepsilon$$

for  $s > s_0$  and  $v > v_0$ . These facts can be used to determine the asymptotics of the double integral (2.17).

First, by partial integration, the inner integral in (2.17) is at most  $\|G_s\|\Phi(ne^{-s})$ , so that truncating the external integral at the lower bound  $s_0$  yields an error of at most  $\lambda s_0 \Phi^2(n)$ , which is negligible when compared with the claimed asymptotics. Similarly, truncating the external integral at an upper bound  $\log n - \psi$  yields an error of at most

$$2\sigma^2 \int_{\log n - \psi}^{\infty} \Phi^2(ne^{-s}) U(ds) \leq 2e^{2\psi} \sigma^2 \left\{ \int_0^{\infty} e^{-2v} dv + U^* \right\} = 2e^{2\psi} \sigma^2 (1 + U^*),$$

using (2.6), where  $U^* := \sup_{s>0} |U[0, s] - s|$ . Then truncating the internal integral at the upper bound  $v_0$  results in an error estimated as

$$2 \int_{s_0}^{\infty} \Phi(ne^{-s}) U(ds) \int_{v_0}^{\infty} \Phi(ne^{-s-v}) d\{U[v_0, v] - U[s+v_0, s+v]\} < 2\varepsilon \int_{s_0}^{\infty} \Phi^2(ne^{-s}) U(ds), \quad (2.21)$$

while, as above,

$$\left| \int_0^{\infty} \Phi^2(ne^{-s})(U(ds) - ds) \right| \leq 2U^* \Phi^2(n),$$

so that

$$\int_0^{\infty} \Phi^2(ne^{-s}) U(ds) \sim \int_0^n \Phi^2(s) \frac{ds}{s}. \quad (2.22)$$

Thus we are reduced to evaluating

$$J_{\varepsilon, n} := 2 \int_{s_0}^{\log n - \psi_n} \Phi(ne^{-s}) U(ds) \int_0^{v_0} \Phi(ne^{-s-v}) d\{U[0, v] - U[s, s + v]\},$$

where we let  $\psi = \psi_n \rightarrow \infty$  slowly enough that

$$e^{2\psi_n} \ll \int_0^n \Phi^2(ne^{-s}) U(ds).$$

But in the range  $v < v_0$ ,  $s < \log n - \psi_n$  we have  $e^{-v}$  bounded from 0 and  $\infty$ , and  $ne^{-s} > e^{\psi_n} \rightarrow \infty$ ; hence, by the uniform convergence theorem for slowly varying functions [4],

$$\frac{\Phi(ne^{-s-v})}{\Phi(ne^{-s})} \rightarrow 1, \quad \text{as } n \rightarrow \infty,$$

uniformly in such  $s$  and  $v$ . With this substitution and using (2.20) and (2.22) we obtain

$$J_{\varepsilon, n} \sim (1 + O(\varepsilon)) \sigma^2 \int_{s_0}^{\infty} \Phi^2(ne^{-s}) U(ds) \sim (1 + O(\varepsilon)) \sigma^2 \int_0^n \Phi^2(s) \frac{ds}{s}.$$

Hence, recalling (2.21) and sending  $\varepsilon \rightarrow 0$ , the desired asymptotics follow.

Noting that, under A1,

$$\Psi_2(n) = \int_0^n \Phi^2(s) \frac{ds}{s} \gg \int_0^n \Phi(s) \frac{ds}{s} = \Psi(n), \quad (2.23)$$

it follows from Lemma 2.6 that the asymptotics of  $\text{Var} K_n$  are implied by those of  $\text{Var} A_n(\infty)$ .  $\square$

**Remarks.** A general subordinator  $\tilde{S}$  with

$$\mathfrak{m} := \mathbb{E}\tilde{S}_1 = \int_0^\infty x \tilde{\nu}_0(dx) \quad \text{and} \quad \tau^2 := \text{Var} \tilde{S}_1 = \int_0^\infty x^2 \tilde{\nu}_0(dx)$$

yields a subordinator  $S$  with  $\mathbb{E}S_1 = 1$ , by scaling time so that  $S_t = \tilde{S}_{t/\mathfrak{m}}$ ;  $S$  has  $\Phi = \mathfrak{m}^{-1}\tilde{\Phi}$  and  $\sigma^2 = \mathfrak{m}^{-1}\tau^2$ , and has the same quantities  $K_n$  and  $A_n(\infty)$  as  $\tilde{S}$ . Hence, for  $\tilde{S}$ , we have

$$\mathbb{E}A_n(\infty) \sim \int_0^n s^{-1}\Phi(s) ds = \mathfrak{m}^{-1} \int_0^n s^{-1}\tilde{\Phi}(s) ds$$

and

$$\text{Var} A_n(\infty) \sim \sigma^2 \int_0^n s^{-1}\Phi^2(s) ds = \tau^2 \mathfrak{m}^{-3} \int_0^n s^{-1}\tilde{\Phi}^2(s) ds.$$

For the gamma subordinator with  $\tilde{\nu}_0(dx) = ax^{-1}e^{-\theta x} dx$ , we have  $\tilde{\Phi}_0(s) = a \log(1 + s/\theta)$ ,  $\mathfrak{m} = a/\theta$  and  $\tau^2 = a/\theta^2$ . Hence

$$\text{Var} K_n \sim \text{Var} A_n(\infty) \sim \theta \int_0^n s^{-1} \log^2(1 + s/\theta) ds \sim \frac{\theta}{3} \log^3 n,$$

agreeing with the asymptotics for gamma-like subordinators obtained in [11] by a method based on the Mellin transform.

In the compound Poisson case, the asymptotics of  $\text{Var} A_n(\infty)$  and  $\text{Var} K_n$  are different, because  $\Psi(n)$  is no longer of smaller order than  $\Psi_2(n)$  as in (2.23). Instead, with the normalisation  $\nu_0[0, \infty] = 1$ , so that  $\lim_{n \rightarrow \infty} \tilde{\Phi}(n) = 1$ , we have

$$\text{Var} A_n(\infty) \sim \frac{\tau^2}{\mathfrak{m}^3} \log n; \quad \text{Var} K_n \sim \frac{\tau^2 - \mathfrak{m}^2}{\mathfrak{m}^3} \log n,$$

(see [7]), so that (2.19) is valid only for the compensator.

In the case of regular variation,  $\Psi_2(n)$  still gives the correct order of growth for the variances of both quantities, but the coefficients are not as in (2.19); see [10] for details.

In consequence of (2.19), (2.23) and Lemma 2.6, the fluctuations of the process  $\mathcal{K}_n - A_n$  are of smaller order than those of  $A_n$ , so that, when studying limit theorems for  $\mathcal{K}_n(t)$ , it is enough to consider  $A_n(t)$ .

### 3 The key assumption

The asymptotics of moments only required the monotonicity of  $\Phi$  and the property of slow variation. In order to progress to a finer description of the asymptotics of  $K_n$ , we need a

further assumption in addition to A1. To express it, we begin by associating with  $\Phi$  the function

$$L(s) := \frac{\Phi(s)}{s\Phi'(s)},$$

so that  $(sL(s))^{-1} = (\log \Phi(s))'$  and

$$\Phi(s) = \Phi(1) \exp \left\{ \int_1^s \frac{dz}{zL(z)} \right\}. \quad (3.1)$$

Assumption A1 forces  $\lim_{s \rightarrow \infty} L(s) = \infty$ , because the last formula is just an instance of the Karamata representation for slowly varying functions [4]. Constantly keep in mind that the faster  $L$ , the slower  $\Phi$ .

Our extra assumption on  $\Phi$  is expressed via  $L$ , and puts a limit on the way in which it can vary locally: we assume that there exist  $s_0 \geq 1$  and  $k > 0$  such that

$$\textbf{Assumption A2:} \quad \left| \frac{sL'(s)}{L(s)} \right| < \frac{k}{\log s} \quad \text{for all } s \geq s_0. \quad (3.2)$$

Because the right side in (3.2) goes to zero with  $s$ , the function  $L$  is itself slowly varying; under A1 the latter property is equivalent to the slow variation of  $s\Phi'(s)$ .

**Lemma 3.1** *If A1 holds and  $L$  is slowly varying, then for  $L_0(s) := \Phi_0(s)/(s\Phi'_0(s))$*

$$\left| \frac{sL'(s)}{L(s)} - \frac{sL'_0(s)}{L_0(s)} \right| = o(s^{-1}L^2(s));$$

*thus Assumption A2 can equivalently be stated using  $L_0$  in place of  $L$ .*

**Proof.** Direct calculation shows that

$$\frac{sL'(s)}{L(s)} - \frac{sL'_0(s)}{L_0(s)} = s \left\{ \left( \frac{\Phi'(s)}{\Phi(s)} - \frac{\Phi'_0(s)}{\Phi_0(s)} \right) - \left( \frac{\Phi''(s)}{\Phi'(s)} - \frac{\Phi''_0(s)}{\Phi'_0(s)} \right) \right\}.$$

Now we apply Lemma 2.3 to bound differences between the derivatives of  $\Phi_0$  and those of  $\Phi$ , and (2.9) to bound the derivatives themselves, and we also note that  $\Phi'(s) = \Phi(s)/sL(s)$ . The lemma follows.  $\square$

We note in passing that both functions  $1/L_0$  and  $1/L$  can be given various probabilistic interpretations. For instance, in the spirit of (2.1),

$$1/L_0(n) = \mathbb{E} \exp\{-n(S_\tau - \xi)\}$$

where  $\xi$  is an independent exponential level with rate  $n$  and  $\tau$  is the passage time across  $\xi$ , so that  $S_\tau - \xi$  is the overshoot at  $\xi$ , see [15, Corollary 1 (ii)]. The function  $1/L$  determines a conditional rate for creating singleton blocks of  $\mathcal{C}_n$ , meaning that  $1/L(s)$ , with  $s = ne^{-S_t}$ , is the conditional probability that a jump of  $S$  at time  $t$  covers exactly one Poisson point given at least one point is covered.

Although we regard A2 as a *local* condition, under circumstances it can restrict the *global* growth of  $L$ . For suppose that  $L$  is eventually increasing. Then, introducing  $h(m) := L(m)/(mL'(m))$ , we have in the usual way

$$L(m) = L(1) \exp\left(\int_1^m \frac{dz}{zh(z)}\right).$$

Now, because eventually  $L' \geq 0$ , Assumption A2 reads as  $h(m) > k^{-1} \log m$ , hence yielding a global bound  $L(m) < \lambda \log^k m$ . In the other direction, observe that, even if  $L$  is not monotone, the inequality inverse to (3.2),  $L(m) > \lambda \log^k m$  with some  $k > 1$ , would disagree with A1, because in this case  $\Phi$  would be bounded.

**Remark.** Of course, when (3.2) holds for some  $s_0$  and  $k$ , we can set  $s_0 = 1$  by taking  $k$  sufficiently large. This will suffice for our purposes, but gives a poor idea of the growth of  $L$ .

Assumption A2 implies that

$$\left(\frac{\log y}{\log x}\right)^{-k} < \frac{L(y)}{L(x)} < \left(\frac{\log y}{\log x}\right)^k, \quad s_0 < x < y. \quad (3.3)$$

To see this, for  $s_0 < x < y$ , observe that

$$\frac{L(y)}{L(x)} = \exp\left\{\int_x^y \frac{dz}{h(z)z}\right\} < \exp\left\{\int_x^y \frac{dz}{z} \left(\frac{k}{\log z}\right)\right\} = \left(\frac{\log y}{\log x}\right)^k,$$

and similarly that

$$\frac{L(y)}{L(x)} > \exp\left\{-\int_x^y \frac{dz}{z} \left(\frac{k}{\log z}\right)\right\} = \left(\frac{\log y}{\log x}\right)^{-k}.$$

Assumption A2 is a kind of ‘second order’ slow variation, in the sense that the function  $\log \Phi$  has a form of de Haan’s property (see [4, Section 3.0]):

$$\frac{\log \Phi(ne^{-s}) - \log \Phi(n)}{1/L(n)} \rightarrow -s, \quad (n \rightarrow \infty).$$

However, A2 is stronger than just this, and offers a better control on the variability of  $\Phi$ ; in particular we have the following estimate for the remainder.

**Lemma 3.2** *Under assumptions A1–A2, we have*

$$\left|\log\left\{\frac{\Phi(me^{-s})}{\Phi(m)}\right\} + \frac{s}{L(m)}\right| < \frac{\varkappa s^2}{L(m) \log m}, \quad (3.4)$$

for all  $m \geq 1$  and  $|s| < \frac{1}{2} \log m$ , where  $\varkappa = k2^k$ .

**Proof.** By Taylor's formula with the remainder in Lagrange's form,

$$\log \Phi(me^{-s}) = \log \Phi(m) - \frac{s}{L(m)} - \frac{s^2}{2} \frac{m^* L'(m^*)}{L(m^*)^2},$$

for some  $m^*$  such that

$$m \wedge (me^{-s}) \leq m^* \leq m \vee (me^{-s}).$$

Since  $|s| < (\log m)/2$ , we have  $m^{1/2} \leq m^* \leq m^{3/2}$ , and hence  $\frac{1}{2} \log m \leq \log m^* \leq \frac{3}{2} \log m$ ; and also, from A2,

$$\left| \frac{m^* L'(m^*)}{L(m^*)} \right| < \frac{k}{\log m^*} \leq \frac{2k}{\log m}.$$

On the other hand, for  $m^*$  in this range, (3.3) implies that  $L(m^*) > 2^{-k} L(m)$ . Hence

$$\frac{m^* L'(m^*)}{2L(m^*)^2} < \frac{k2^k}{L(m) \log m},$$

as required. □

**Remark.** So, loosely speaking, we are dealing with functions  $L$  that grow slowly enough, satisfying  $L(n^{1/2}) \asymp L(n)$ . Indeed, it can be shown that  $L(n) \gg L(n^{1/2})$  implies the convergence

$$\int_1^\infty \frac{ds}{sL(s)} < \infty,$$

in which case the Lévy measure is finite.

**Corollary 3.3** *Under Assumptions A1–A2, we have*

$$\Phi(n) e^{-3s/2L(n)} \leq \Phi(ne^{-s}) \leq \Phi(n) e^{-s/2L(n)}$$

for  $0 \leq s \leq \frac{1}{2(\varkappa\sqrt{1})} \log n$ .

**Proof.** Immediate from the above. □

**Corollary 3.4** *Under Assumptions A1–A2, if  $L(n) < \frac{1}{4(\varkappa\sqrt{1})} \log n$ , we have*

$$\begin{aligned} \Psi(n) &\leq 1 + \Phi(n)L(n) \left( 2 + \frac{4(\varkappa\sqrt{1})}{e} \right); \\ \frac{1}{3}(1 - e^{-6})\Phi^2(n)L(n) &\leq \Psi_2(n) \leq \frac{1}{2} + \Phi^2(n)L(n) \left\{ 1 + \frac{2(\varkappa\sqrt{1})}{e} \right\}. \end{aligned}$$

If  $L(n) \geq \frac{1}{4(\varkappa\sqrt{1})} \log n$ , we have

$$\begin{aligned} \Psi(n) &\leq 1 + \Phi(n) \log n \left( \frac{1}{2(\varkappa\sqrt{1})} + \frac{1}{e} \right); \\ \frac{1}{3}(1 - e^{-6})\Phi^2(n) \log n &\leq \Psi_2(n) \leq \frac{1}{2} + \Phi^2(n) \log n \left\{ \frac{1}{2(\varkappa\sqrt{1})} + \frac{1}{e^2} \right\}. \end{aligned}$$



**Proof.** Write  $k_n := \frac{1}{2(\varkappa \vee 1)} \log n$ . From the upper bound in Corollary 3.3, we have

$$\int_0^{k_n} \Phi(ne^{-t}) dt \leq \Phi(n)\{2L(n) \wedge k_n\},$$

and

$$\begin{aligned} \int_{k_n}^{\log n} \Phi(ne^{-t}) dt &\leq \Phi(n)e^{-k_n/2L(n)} \log n \\ &= 4(\varkappa \vee 1)\Phi(n)L(n) \left\{ \frac{k_n}{2L(n)} \exp\left(-\frac{k_n}{2L(n)}\right) \right\}; \end{aligned}$$

furthermore, from (2.6),

$$\int_{\log n}^{\infty} \Phi(ne^{-t}) dt \leq \int_{\log n}^{\infty} ne^{-t} dt = 1.$$

The bounds for  $\Psi(n)$  now follow from its definition, and because  $xe^{-x} \leq e^{-1}$  for  $x \geq 0$ . The proof of the upper bounds for  $\Psi_2(n)$  is analogous.

For the lower bound on  $\Psi_2(n)$ , integrate  $\Phi^2(ne^{-t})$  from 0 to  $\min\{k_n, 2L(n)\}$ , and then use the lower bound in Corollary 3.3.  $\square$

For the rest of this paper both assumptions A1 and A2 will be taken for granted, even if not explicitly mentioned.

## 4 The forward argument

In this section, under a wide range of circumstances in which  $L(n) = O(\log n)$ , we show that the quantity

$$A_n^*(T) := \int_0^{T \wedge \log n} \Phi(ne^{-t}) \left(1 - \frac{S_t - t}{L(ne^{-t})}\right) dt \quad (4.1)$$

is an adequate approximation to  $A_n(T \wedge \tau_n)$ , and hence, in view of Lemma 2.2, to  $A_n(T)$ . This is a very attractive result, because the random process  $S$  appears only linearly in  $A_n^*(T)$ , making it easier to determine the approximate behaviour of  $A_n(T)$  from knowledge of that of  $S$ . The way that the approximation is proved is to show that the quantity  $\sup_{T \geq 0} |A_n(T \wedge \tau_n) - A_n^*(T)|$  is asymptotically smaller than the scale of fluctuations of  $A_n$ . In view of Lemma 2.2 and Corollary 3.4, in order to achieve this when  $L(n) = O(\log n)$ , we need to prove that, with probability tending to 1,

$$\sup_{T \geq 0} |A_n(T \wedge \tau_n) - A_n^*(T)| = o(\Phi(n)\sqrt{L(n)}).$$

To this end, we define the centred process

$$Z_t := S_t - t$$

and restrict attention as far as possible to realisations of  $S$  for which the paths of  $Z$  are reasonably nice. This we make precise as follows. First, for any  $T, \varphi, \psi > 0$ , we define the events

$$B_0(n) := \left\{ \frac{1}{2} \log n \leq \tau_n \leq 2 \log n \right\}; \quad (4.2)$$

$$B_1(T, \varphi) := \left\{ \sup_{0 \leq t \leq T} (t \vee 1)^{-1/2} |Z_t| \leq \varphi \right\}; \quad (4.3)$$

$$B_2(\psi) := \left\{ \sup_{t \geq \psi} 2t^{-1} |Z_t| \leq 1 \right\}. \quad (4.4)$$

The paths of  $Z$  are well behaved if  $B_1(T, \varphi)$  holds for  $\varphi$  not too large and for large enough  $T$ , and if  $B_2(\psi)$  holds for  $\psi$  not too large. With reference to these desiderata, we have the following lemma.

**Lemma 4.1** *For  $T, \varphi, \psi > 0$ , we have*

$$\mathbb{P}[B_0^c(n)] \leq 8\sigma^2 / \log n; \quad (4.5)$$

$$\mathbb{P}[B_1^c(T, \varphi)] \leq 2\sigma^2 \varphi^{-2} \lceil \log_2 T \rceil \quad (T > 2); \quad (4.6)$$

$$\mathbb{P}[B_2^c(\psi)] \leq 32\sigma^2 \psi^{-1}, \quad (4.7)$$

and also  $B_0(n) \supset B_1(2 \log n, \frac{1}{2} \log^\alpha n)$  for any  $0 \leq \alpha \leq \frac{1}{2}$  and  $n \geq 3$ .

**Proof.** First, by Kolmogorov's inequality for the centred, independent increments process  $Z$ , we have

$$\mathbb{P}[B_0^c(n)] \leq \mathbb{P} \left[ \sup_{0 \leq u \leq 2 \log n} |Z_u| > \frac{1}{2} \log n \right] \leq 8\sigma^2 / \log n.$$

The remaining statements are proved by combining Kolmogorov's inequality with geometric dissection, in a rather standard fashion. For the second inequality, we have

$$\begin{aligned} \mathbb{P}[B_1^c(T, \varphi)] &\leq \sum_{r=1}^{\lceil \log_2 T \rceil} \mathbb{P} \left[ \max_{0 \leq t \leq 2^r} (2^{r-1})^{-1/2} |Z_t| > \varphi \right] \\ &\leq \sum_{r=1}^{\lceil \log_2 T \rceil} \frac{2^r \sigma^2}{\varphi^2 2^{r-1}} = 2\sigma^2 \varphi^{-2} \lceil \log_2 T \rceil. \end{aligned}$$

For the third, we have

$$\begin{aligned} \mathbb{P}[B_2^c(\psi)] &\leq \sum_{r \geq \lceil \log_2 \psi \rceil} \mathbb{P} \left[ \max_{0 \leq t \leq 2^r} 2|Z_t| > 2^{r-1} \right] \\ &\leq \sum_{r \geq \lceil \log_2 \psi \rceil} 16\sigma^2 2^{-r} \leq 32\sigma^2 \psi^{-1}. \end{aligned}$$

□

The following corollary needs no proof.

**Corollary 4.2** For any positive sequences  $T_n, \varphi_n, \psi_n$  we have

- (i)  $\lim_{n \rightarrow \infty} \mathbb{P}[B_1(T_n, \varphi_n)] = 1$  if  $\varphi_n^{-2} \log T_n \rightarrow 0$ ;
- (ii)  $\lim_{n \rightarrow \infty} \mathbb{P}[B_2(\psi_n)] = 1$  if  $\psi_n \rightarrow \infty$ .

For the further argument, we distinguish two cases, relating to the global pattern of growth of  $\Phi(n)$ , each of which needs separate treatment. The idea of the distinction can be seen from the following formula for the variance of the linearised compensator when  $T < \log n$ :

$$\text{Var } A_n^*(T) = \sigma^2 \int_0^T \{\Phi(ne^{-t}) - \Phi(ne^{-T})\}^2 dt, \quad (4.8)$$

which is derived by writing (4.1) for the centred  $A_n^*$  as a stochastic integral:

$$A_n^*(T) - \mathbb{E}A_n^*(T) = - \int_0^T \{\Phi(ne^{-t}) - \Phi(ne^{-T})\} d(S_t - t),$$

and using the independence of increments. So, when  $\Phi$  is a function like a power of logarithm, the difference  $\Phi(ne^{-t}) - \Phi(ne^{-T})$  is of constant order over the whole time-range from 0 to  $\log n$ . On the other hand, if  $\Phi$  grows fast enough, the first term will dominate, and the principal contribution to the integral will come from times  $t = o(\log n)$ , as is also the case if  $\Phi$  is regularly varying.

#### 4.1 Moderately growing $\Phi$

We begin with the boundary case, which includes the gamma-like subordinators [11], when  $\Phi(n)$  grows more or less like a power of  $\log n$ . Here, all times  $t$  between 0 and  $\log n$  contribute more or less evenly to the fluctuations of  $A_n$ . This case is defined by a global condition on the function  $L$ ; that, for some  $1 \leq c_2 < \infty$  and for some  $m_0$ , and with  $c_1 := \{3(\varkappa \vee 1)\}^{-1}$ ,

$$\frac{c_1 \log m}{6 \log \log m} \leq L(m) \leq c_2 \log m, \quad m \geq m_0. \quad (4.9)$$

The next lemma is a preliminary to proving that, under these circumstances,  $A_n^*$  is a good approximation to  $A_n$ . It enables us to truncate the integrals defining  $A_n(T)$  and  $A_n^*(T \wedge \tau_n)$  close to  $\log n$ , when the paths of  $Z$  are nice enough.

**Lemma 4.3** On the event  $B_1(2 \log n, \varphi_n)$ , and for  $0 < \psi_n \leq \log n - \varphi_n \sqrt{\log n}$ , we have, for all  $T > 0$ ,

$$0 \leq \int_0^{\tau_n \wedge T} \Phi(ne^{-S_t}) dt - \int_0^{\psi_n \wedge T} \Phi(ne^{-S_t}) dt \leq \eta_n; \quad (4.10)$$

$$\int_{\psi_n \wedge T}^{T \wedge \log n} \Phi(ne^{-t}) \left| 1 - \frac{Z_t}{L(ne^{-t})} \right| dt \leq \eta_n, \quad (4.11)$$

where

$$\eta_n = (\log n - \psi_n + \varphi_n \sqrt{2 \log n}) \Phi(n \exp\{-\psi_n + \varphi_n \sqrt{\psi_n}\}).$$

**Proof.** On  $B_1(2 \log n, \varphi_n)$ , we have

$$\log n - \varphi_n \sqrt{\log n} \leq \tau_n \leq \log n + \varphi_n \sqrt{2 \log n},$$

and so  $\psi_n \leq \min\{\tau_n, \log n\}$ . Hence (4.10) and (4.11) are both zero if  $T \leq \psi_n$ . The first part of the lemma then merely uses the fact that

$$\psi_n \leq \tau_n \leq \log n + \varphi_n \sqrt{2 \log n},$$

combined with the largest possible value of the integrand in this range. For the second part, we observe that

$$\int_{\psi_n}^{\log n} \Phi(ne^{-t}) dt \leq \Phi(ne^{-\psi_n})(\log n - \psi_n),$$

and from (4.3)

$$\int_{\psi_n}^{\log n} \Phi(ne^{-t}) \frac{|Z_t|}{L(ne^{-t})} dt \leq \Phi(ne^{-\psi_n}) \varphi_n \sqrt{\log n}. \quad \square$$

It follows from A2, (4.9) and the definition of  $L$  that

$$\Phi(ne^{-t}) = \Phi(n) \exp \left\{ - \int_{ne^{-t}}^n \frac{dy}{yL(y)} \right\} \leq \Phi(n) \left\{ 1 - \frac{t}{\log n} \right\}^{1/c_2} \quad (4.12)$$

for all  $n$  and  $t$  such that  $ne^{-t} \geq m_0$ . Thus, taking

$$\psi_n = \log n - 2u_n \sqrt{\log n}, \quad (4.13)$$

for  $u_n \geq \varphi_n$ , the quantity  $\eta_n$  in Lemma 4.3 is, for all  $n$  large enough, at most

$$\Phi(n) 4u_n \sqrt{\log n} \{3u_n / \sqrt{\log n}\}^{1/c_2}.$$

This is in turn at most

$$12 \sqrt{18(\varkappa \vee 1)} \Phi(n) \sqrt{L(n)} \{ \sqrt{\log \log n} \log^{-\beta/2} n \}$$

if we take  $u_n = \log^\beta n$  for  $\beta = 1/\{4(1 + c_2)\}$ .

**Theorem 4.4** *Suppose that Assumptions A1–A2 and (4.9) hold, fix  $3\alpha = \beta = \frac{1}{4(c_2+1)}$ , and set  $u_n = \log^\beta n$ ,  $\varphi_n = \log^\alpha n$ . Then, on  $B_1(2 \log n, \varphi_n)$ , we have*

$$\sup_{T \geq 0} |A_n(T \wedge \tau_n) - A_n^*(T)| = \varepsilon(n) \left( \Phi(n) \sqrt{L(n)} \right),$$

where  $\lim_{n \rightarrow \infty} \varepsilon(n) = 0$  uniformly in  $1 \leq c_2 \leq C$ , for each  $1 < C < \infty$ . Furthermore,

$$\lim_{n \rightarrow \infty} \mathbb{P}[B_1(2 \log n, \varphi_n)] = 0.$$

**Proof.** By the argument just completed, it is enough to examine the integrated difference

$$\int_0^{\psi_n \wedge T} \left| \Phi(ne^{-s_t}) - \Phi(ne^{-t}) \left\{ 1 - \frac{Z_t}{L(ne^{-t})} \right\} \right| dt,$$

where  $\psi_n = \log n - 2u_n\sqrt{\log n}$ .

To this end, we use Lemma 3.2 with  $ne^{-t}$  for  $m$  and  $Z_t$  for  $s$ . On  $B_1(2\log n, \varphi_n)$ , and since, for  $0 \leq t \leq \psi_n$ , we have  $ne^{-t} \geq \exp\{2u_n\sqrt{\log n}\}$ , it follows that

$$\begin{aligned} |Z_t| &\leq \varphi_n\sqrt{\log n} = \frac{1}{2} \log\{\exp(2\varphi_n\sqrt{\log n})\} \\ &\leq \frac{1}{2} \log\{\exp(2u_n\sqrt{\log n})\} \leq \frac{1}{2} \log\{ne^{-t}\}, \end{aligned}$$

so that the lemma can be applied. It then follows that

$$\begin{aligned} &|\Phi(ne^{-s_t}) - \Phi(ne^{-t}) \exp\{-Z_t/L(ne^{-t})\}| \\ &\leq \Phi(ne^{-t}) \exp\{-Z_t/L(ne^{-t})\} X(n, t) \exp\{X(n, t)\}, \end{aligned} \quad (4.14)$$

where

$$X(n, t) = \frac{\varkappa Z_t^2}{L(ne^{-t})(\log n - t)}. \quad (4.15)$$

It is then also immediate from  $e^{-x} - 1 + x < e^{|x|}x^2/2$  that

$$\begin{aligned} &|\Phi(ne^{-t})| \exp\{-Z_t/L(ne^{-t})\} - 1 + Z_t/L(ne^{-t})| \\ &\leq \frac{1}{2} \Phi(ne^{-t}) \exp\{|Z_t|/L(ne^{-t})\} \{Z_t/L(ne^{-t})\}^2. \end{aligned} \quad (4.16)$$

Now, on  $B_1(2\log n, \varphi_n)$ , and for  $0 \leq t \leq \psi_n$ , we have

$$X(n, t) \leq (6\varkappa/c_1)(\varphi_n/u_n)^2 \log \log n \leq \lambda_1,$$

and also

$$|Z_t|/L(ne^{-t}) \leq (6/c_1)(\varphi_n/u_n) \log \log n \leq \lambda_2;$$

more precisely, in this range of  $t$ , by (4.9)

$$X(n, t) \leq \frac{6\varkappa\varphi_n^2 \log n \log \log n}{c_1(\log n - t)^2} \quad \text{and} \quad \frac{|Z_t|}{L(ne^{-t})} \leq \frac{6\varphi_n\sqrt{\log n} \log \log n}{c_1(\log n - t)}.$$

We also have the bound (4.12) for  $\Phi(ne^{-t})$ . Combining these, it follows that

$$\begin{aligned} &\left| \int_0^{\psi_n} \Phi(ne^{-s_t}) dt - \int_0^{\psi_n} \Phi(ne^{-t}) \left\{ 1 - \frac{Z_t}{L(ne^{-t})} \right\} dt \right| \\ &\leq \int_0^{\psi_n} \Phi(ne^{-t}) \left\{ e^{\lambda_1 + \lambda_2} X(n, t) + \frac{1}{2} e^{\lambda_2} (Z_t/L(ne^{-t}))^2 \right\} dt \\ &\leq \frac{6}{c_1^2} (\varkappa c_1 e^{\lambda_1 + \lambda_2} + 3e^{\lambda_2}) \varphi_n^2 \Phi(n) (\log \log n)^2 \int_{2u_n/\sqrt{\log n}}^1 u^{-2+1/c_2} du \end{aligned} \quad (4.17)$$

$$\begin{aligned} &\leq \lambda \left( \Phi(n) \sqrt{L(n)} \right) (\log \log n)^{5/2} \varphi_n^2 / u_n \\ &= \lambda \left( \Phi(n) \sqrt{L(n)} \right) \{ (\log \log n)^{5/2} \log^{-\beta/3} n \}, \end{aligned} \quad (4.18)$$

for some  $\lambda > 0$ . This completes the proof. Note that  $c_2$  enters the bound implicitly, in the value of  $\beta$ , and hence in  $\lambda_1$  and  $\lambda_2$ .  $\square$

**Remark.** The restrictions imposed by (4.9) can be relaxed somewhat, to allow a little more freedom in both lower and upper bounds. For instance, the same proof can be used under the condition

$$(\log m)^{1-\gamma(c_2(m))} \leq L(m) \leq c_2(m) \log m \quad \text{for all } m \geq m_0, \quad (4.19)$$

for any increasing function  $c_2$  satisfying  $c_2(m) = o(\log \log m / \log \log \log m)$ , where  $\gamma(c) := 1/\{32(c+1)\}$ . Suitable choices of the parameters are now

$$\beta = \beta(n) := 8\gamma(c_2(n)) \quad \text{and} \quad \alpha = \alpha(n) := \beta(n)/3;$$

note that, with these definitions and with  $\varphi_n = \log^{\alpha(n)} n$ , we still have  $\mathbb{P}[B_1^c(2 \log n, \varphi_n)] \rightarrow 0$ . This extra freedom enables the main transition, between the behaviour in the case of moderately growing  $\Phi(n)$  and that when  $\Phi(n)$  grows either faster or more slowly, to be understood in greater detail.

## 4.2 Fast growing $\Phi$

We turn to the setting in which slowly varying  $\Phi(n)$  grows faster than any power of  $\log n$ . In this case most of the random fluctuation in  $A_n$  takes place at times of order  $L(n)$ , where  $L(n)$  goes to infinity (as required by A1) but slower than  $\log n$ . Our global condition determining this régime is

$$6L(n) \log L(n) \leq c_1 \log n, \quad (4.20)$$

where  $c_1 = \{3(\varkappa \vee 1)\}^{-1}$  is as before. Note that, if  $L(n) \leq \frac{c_1 \log n}{6 \log \log n}$ , then (4.20) is satisfied; the condition given in (4.9) was chosen to match neatly, though in view of the remark at the end of the previous section, this was not really necessary. Here, we first need a modification of Lemma 4.3, in order to be able to truncate the integrals defining  $A_n(T \wedge \tau_n)$  and  $A_n^*(T)$  as far as we need to.

**Lemma 4.5** *Suppose that  $\psi_n$  is such that  $6L(n) \log L(n) \leq \psi_n \leq c_1 \log n$ , and that  $n$  is large enough to satisfy  $L(n) \geq e^6$ . Then, on the event  $B_2(\psi_n)$ , we have*

$$0 \leq \int_0^{\tau_n \wedge T} \Phi(ne^{-St}) dt - \int_0^{\psi_n \wedge T} \Phi(ne^{-St}) dt \leq 4\Phi(n)\{e^{-3} + c_1^{-1}\}; \quad (4.21)$$

$$\int_{\psi_n \wedge T}^{T \wedge \log n} \Phi(ne^{-t}) \left| 1 - \frac{Z_t}{L(ne^{-t})} \right| dt \leq \Phi(n)\{2e^{-3} + \frac{3}{2}c_1^{-1} + 6e^{-9}2^k\}. \quad (4.22)$$

**Proof.** On  $B_2(\psi_n)$ , we have  $\frac{2}{3} \log n \leq \tau_n \leq 2 \log n$ , implying immediately that  $\psi_n \leq \min\{\tau_n, \log n\}$ . Hence, if  $T \leq \psi_n$ , both of the quantities to be bounded in the lemma are zero. Note also, in preparation, that for  $l \geq e^6$  and for any  $x \geq 6l \log l$ , we have

$$xe^{-x/4l} \leq 6l \log l \exp\{-3 \log l/2\} = 6l^{-1/2} \log l \leq 2. \quad (4.23)$$

For the bound (4.21), since  $S_t \geq t/2$  for  $t \geq \psi_n$  on  $B_2(\psi_n)$  and since  $c_1 \leq 1/2(\varkappa \vee 1)$ , we can apply Corollary 3.3 to give

$$\begin{aligned} \int_{\psi_n}^{c_1 \log n} \Phi(ne^{-S_t}) dt &\leq \Phi(n) \int_{\psi_n}^{c_1 \log n} e^{-t/4L(n)} dt \\ &\leq 4L(n)\Phi(n)e^{-\psi_n/4L(n)} \leq 4L(n)^{-1/2}\Phi(n), \end{aligned}$$

by the definition of  $\psi_n$ . Then we also have

$$\int_{c_1 \log n}^{2 \log n} \Phi(ne^{-S_t}) dt \leq 2 \log n \Phi(n) \exp\{-c_1 \log n/4L(n)\} \leq 4c_1^{-1}\Phi(n),$$

this last by (4.23).

The argument for (4.22) is very similar. First, bounding  $\int_{\psi_n}^{\log n} \Phi(ne^{-t}) dt$ , it follows from Corollary 3.3 that

$$\int_{\psi_n}^{c_1 \log n} \Phi(ne^{-t}) dt \leq 2L(n)\Phi(n)e^{-\psi_n/4L(n)} \leq 2L(n)^{-1/2}\Phi(n),$$

and then that

$$\int_{c_1 \log n}^{\log n} \Phi(ne^{-t}) dt \leq \log n \Phi(n) \exp\{-c_1 \log n/2L(n)\} \leq c_1^{-1}\Phi(n).$$

For the remaining term, we first have

$$\begin{aligned} \int_{\psi_n}^{c_1 \log n} \frac{\Phi(ne^{-t})|Z_t|}{L(ne^{-t})} dt &\leq \frac{\Phi(n)}{L(n^{1-c_1})} \int_{\psi_n}^{c_1 \log n} \frac{1}{2}te^{-t/2L(n)} dt \\ &\leq \frac{2L(n)^2\Phi(n)}{L(n^{1-c_1})} \int_{\psi_n/2L(n)}^{\infty} ue^{-u} du. \end{aligned} \quad (4.24)$$

Now, for any  $y \geq 1/2$ ,

$$\int_y^{\infty} ue^{-u} du = \int_0^{\infty} (y+v)e^{-y-v} dv = e^{-y}(y+1) \leq 3ye^{-y},$$

so that (4.24) can be bounded, using (3.3) and (4.23), by

$$\frac{3L(n)\Phi(n)}{L(n^{1-c_1})} \psi_n e^{-\psi_n/2L(n)} \leq \frac{6L(n)\Phi(n)}{L(n^{1/2})} e^{-\psi_n/4L(n)} \leq \frac{6 \cdot 2^k \Phi(n)}{\{L(n)\}^{3/2}}.$$

Finally, using (5.3) and  $c_1 < 1/2(\varkappa \vee 1)$ , we have

$$\begin{aligned} \int_{c_1 \log n}^{\log n} \frac{\Phi(ne^{-S_t})|Z_t|}{L(ne^{-t})} dt &\leq \frac{1}{2} \log n \Phi(ne^{-c_1 \log n}) \\ &\leq \frac{1}{2} \log n \Phi(n) e^{-c_1 \log n/2L(n)} \leq \frac{1}{2c_1} \Phi(n), \end{aligned}$$

again using (4.23). This completes the proof.  $\square$

**Theorem 4.6** *Under Assumptions A1–A2 and (4.20), set  $\psi_n = 6L(n) \log L(n)$  and  $\varphi_n = L(n)^{1/6}$ . Then, on the event  $B_1(\psi_n, \varphi_n) \cap B_2(\psi_n)$ , and if  $L(n) \geq e^6$ , we have*

$$\sup_{T \geq 0} |A_n(T \wedge \tau_n) - A_n^*(T)| \leq \varepsilon(L(n)) \left( \Phi(n) \sqrt{L(n)} \right),$$

where  $\lim_{m \rightarrow \infty} \varepsilon(m) = 0$ . Furthermore,

$$\lim_{n \rightarrow \infty} \mathbb{P}[B_1^c(\psi_n, \varphi_n)] = \lim_{n \rightarrow \infty} \mathbb{P}[B_2^c(\psi_n)] = 0.$$

**Proof.** As before, on  $B_2(\psi_n)$ , we have  $\frac{2}{3} \log n \leq \tau_n \leq 2 \log n$ , implying immediately that  $\psi_n \leq \min\{\tau_n, \log n\}$ . By Lemma 4.5, it is enough to bound the difference

$$\int_0^{\psi_n} \left| \Phi(ne^{-st}) - \Phi(ne^{-t}) \left\{ 1 - \frac{Z_t}{L(ne^{-t})} \right\} \right| dt.$$

By (3.3), we can use the inequality  $L(ne^{-t}) \geq 2^{-k} L(n)$  for  $0 \leq t \leq c_1 \log n$ . Hence, on the event  $B_1(\psi_n, \varphi_n)$ , and noting that  $\log n - \psi_n \geq \frac{1}{2} L(n)$  because of (4.20), we can bound the quantities  $X(n, t)$  and  $\{Z_t/L(ne^{-t})\}^2$  appearing in the proof of Theorem 4.4 by

$$2^{2k} \varphi_n^2 \psi_n L(n)^{-2} = 6 \cdot 2^{2k} L(n)^{-2/3} \log L(n) \leq 36 e^{-4} 2^{2k},$$

in the range  $t \leq \psi_n$ ; thus they are both uniformly bounded in  $n$ , and asymptotically small as  $n \rightarrow \infty$ . Hence, using (4.14) and (4.16), it follows that

$$\begin{aligned} & \int_0^{\psi_n} \left| \Phi(ne^{-st}) - \Phi(ne^{-t}) \left\{ 1 - \frac{Z_t}{L(ne^{-t})} \right\} \right| dt \\ & \leq \lambda L(n)^{-2/3} \log L(n) \int_0^{\psi_n} \Phi(ne^{-t}) dt \end{aligned}$$

for some  $\lambda < \infty$ . But now, from Corollary 3.3, it follows that

$$\begin{aligned} L(n)^{-2/3} \log L(n) \int_0^{\psi_n} \Phi(ne^{-t}) dt & \leq 2L(n)^{1/3} \Phi(n) \log L(n) \\ & = 2L(n)^{-1/6} \log L(n) \left( \Phi(n) \sqrt{L(n)} \right), \end{aligned}$$

proving the main assertion. The last statement follows from Corollary 4.2.  $\square$

## 5 The backward argument

We now turn to the case of functions  $\Phi(n)$  that grow more slowly than any power of  $\log n$ . Here, the argument required and the approximations obtained are of rather different character to those of the previous section. In particular, we make use of properties of the Lévy process when looking backwards in time. Our setting is defined by requiring that  $\lim_{n \rightarrow \infty} \Phi(n) = \infty$ , but that  $L$  satisfies the following global condition:

$$L(m) = c_2(m) \log m, \quad \text{where} \quad \lim_{m \rightarrow \infty} c_2(m) = \infty. \quad (5.1)$$



To agree with A1,  $c_2(m)$  must grow slowly enough, meaning that the integral in (3.1),

$$\int_2^\infty \frac{dm}{c_2(m)m \log m},$$

must diverge, a condition which excludes functions like  $c_2(m) = \log^\varepsilon m$  for any  $\varepsilon > 0$ . One can think of  $c_2(m) = \log \log m$  for  $m \geq m_0$ , as one possible example, in which case  $\Phi(n) \asymp \log \log n$ . Here, we no longer have Lemma 4.3 to help us. However, the argument of Theorem 4.4 is still good, if we restrict to taking the supremum over  $0 \leq T \leq (1 - \delta_n) \log n$ , for some  $\delta_n \rightarrow 0$  sufficiently slowly, and this gives us the following approximation of  $A_n$  by  $A_n^*$ .

**Lemma 5.1** *Take  $\alpha = 1/8$ ,  $\varphi_n = \log^\alpha n$  and  $\delta_n = \log^{-1/8} n$ . Then, on the event  $B_1(2 \log n, \varphi_n)$ , it follows that*

$$\sup_{0 \leq T \leq (1 - \delta_n) \log n} |A_n(T \wedge \tau_n) - A_n^*(T)| \leq \lambda \frac{1}{\log^{1/8} n} \frac{\Phi(n) \sqrt{\log n}}{c_2^*(n^{\delta_n})},$$

for some  $\lambda > 0$  and  $c_2^*(m) = \inf_{r \geq m} c_2(r)$ .

**Proof.** We argue as for Theorem 4.4, now with  $\psi_n = (1 - \delta_n) \log n$ , noting that, for  $t \leq \psi_n$ ,

$$|X(n, t)| \leq \frac{\varkappa \varphi_n^2 \log n}{(\log n - t)^2 c_2^*(n^{\delta_n})} \leq \frac{\varkappa}{\varphi_n^2 c_2^*(n^{\delta_n})},$$

since  $\delta_n > \varphi_n^2 / \sqrt{\log n}$ , and that

$$\frac{|Z_t|}{L(ne^{-t})} \leq \frac{\varphi_n \sqrt{\log n}}{(\log n - t) c_2^*(n^{\delta_n})} \leq \frac{1}{\varphi_n c_2^*(n^{\delta_n})},$$

both of which are small in  $n$ . Then, arguing as for (4.17), and using the crude bound  $\Phi(ne^{-t}) \leq \Phi(n)$ , we have

$$\begin{aligned} |A_n(T \wedge \tau_n) - A_n^*(T)| &\leq \lambda \frac{\Phi(n) \varphi_n^2}{c_2^*(n^{\delta_n})} \int_{\delta_n}^1 u^{-2} du \\ &\leq \lambda \frac{\Phi(n) \varphi_n^2}{\delta_n c_2^*(n^{\delta_n})} \leq \lambda \frac{1}{\log^{1/8} n} \frac{\Phi(n) \sqrt{\log n}}{c_2^*(n^{\delta_n})}, \end{aligned}$$

for  $T \leq (1 - \delta_n) \log n$ , as required.  $\square$

To see that differences of this order are relatively small, we now make some variance calculations, for which we introduce the notation

$$-W(v) := \int_0^{v \log n} g(n, t) (S_t - t) dt, \quad (5.2)$$

where

$$g(n, t) := \Phi(ne^{-t}) / L(ne^{-t}) = ne^{-t} \Phi'(ne^{-t}) = -\frac{d}{dt} \{ \Phi(ne^{-t}) \}. \quad (5.3)$$

It thus follows that

$$W(v) = A_n^*(v \log n) - \mathbb{E}\{A_n^*(v \log n)\} \quad (5.4)$$

for  $v \in [0, 1]$ .

**Lemma 5.2** For  $0 \leq v \leq 1$  and  $n$  large enough, we have

$$\text{Var } A_n^*(v \log n) \geq \lambda(v \wedge \frac{1}{2})^3 \Phi(n)^2 \log n / c_2(n)^2;$$

for  $0 < \delta < 1$  and for  $0 \leq v \leq (1 - \delta)$ , we have

$$\text{Var } A_n^*(v \log n) \leq \lambda \Phi(n)^2 \log n \frac{\log(1/\delta)}{c_2^*(n^\delta)},$$

where  $c_2^*(m) = \inf_{r \geq m} c_2(r)$ .

**Proof.** From Lemma 3.2, it follows that, for  $0 \leq v \leq \frac{1}{2}$  and  $t = v \log n$ ,

$$\Phi(ne^{-t}) \geq \Phi(n) e^{-t/L(n)} \exp\{-\varkappa v^2 / c_2(n)\},$$

and, from (3.3), that  $L(n)/L(ne^{-t}) \geq 2^{-k}$ , implying that, for  $n$  so large that  $\varkappa/\{4c_2(n)\} \leq 1$ , we have

$$g(n, t) \geq 2^{-k} e^{-1} \Phi(n) e^{-t/L(n)} / L(n),$$

where  $g(n, t)$  is as in (5.3). Now  $Z_t = S_t - t$  has independent increments with zero means, and  $\text{Var } Z_t = \sigma^2 t$ . Hence, for any  $0 \leq v \leq 1/2$ , recalling (5.4), we have

$$\begin{aligned} \text{Var } A_n^*(v \log n) &= 2 \int_0^{v \log n} \int_0^t g(n, t) g(n, u) \sigma^2 u \, du \, dt \\ &\geq 2^{1-2k} e^{-2} \Phi(n)^2 \int_0^{v/c_2(n)} \int_0^w e^{-w-z} \sigma^2 z L(n) \, dz \, dw \\ &\geq 2^{1-2k} e^{-2} \sigma^2 \Phi(n)^2 c_2(n) \log n \cdot \frac{1}{6} (v/c_2(n))^3 e^{-1/c_2(n)} \\ &\geq 2^{1-2k} \sigma^2 \Phi(n)^2 c_2(n) \log n \cdot \frac{1}{6} (v/c_2(n))^3 e^{-3}, \end{aligned}$$

for all  $n$  large enough. This proves the first inequality, since this lower bound with  $v = 1/2$  is a lower bound for larger  $v$  also.

For the second part, we recall (4.8):

$$\text{Var } A_n^*(v \log n) = \sigma^2 \int_0^{v \log n} \{\Phi(ne^{-t}) - \Phi(n^{1-v})\}^2 \, dt,$$

whenever  $0 \leq v \leq 1$ . Now, from the representation (3.1), it follows that, for  $0 < T \leq (1 - \delta) \log n$ ,

$$\begin{aligned} &\int_0^T \{\Phi(ne^{-t}) - \Phi(ne^{-T})\}^2 \sigma^2 \, dt \\ &\leq \int_0^T \Phi^2(ne^{-t}) \left\{ 1 - \exp\left(-\frac{1}{c_2^*(n^\delta)} \int_{ne^{-T}}^{ne^{-t}} \frac{dy}{y \log y}\right) \right\}^2 \, dt \\ &= \int_0^T \Phi^2(ne^{-t}) \left\{ 1 - \left(\frac{1 - T/\log n}{1 - t/\log n}\right)^{1/c_2^*(n^\delta)} \right\}^2 \, dt \\ &\leq T \Phi^2(n) \{1 - \delta^{1/c_2^*(n^\delta)}\}, \end{aligned}$$

and the second part is proved.  $\square$

In particular, the lower bound shows that the standard deviation of  $A_n^*(T)$  is at least as big as a constant times  $\Phi(n)\sqrt{\log n}/c_2(n)$  for  $T \geq \frac{1}{2} \log n$ . By comparison, the differences in Lemma 5.1 are typically much smaller, because of the factor  $\log^{-1/8} n$ ; recall that  $c_2(n)$  grows rather slowly with  $n$ , and certainly not as fast as a power of  $\log n$ .

Note also that, if  $\delta = \delta_n \rightarrow 0$  sufficiently slowly, the upper bound can be made to grow more slowly than  $\Phi^2(n) \log n$ . For example, with  $c_2(m) = \log \log m$  and therefore  $\Phi(n) \asymp \log \log n$ , one could take  $\delta_m = 1/\log \log m$ , giving an upper bound of order

$$O(\log n \log \log n \log \log \log n) = o(\log n \{\log \log n\}^2).$$

In general, taking  $\delta = \delta_n$  to be the solution of the equation  $\log(1/\delta) = \sqrt{c_2^*(n^\delta)}$  gives both  $\delta_n \rightarrow 0$  and  $\text{Var } A_n((1-\delta_n) \log n) = o(\Phi^2(n) \log n)$ . Thus, *almost* up to the time  $\log n$ , the compensator  $A_n$  behaves very much like the simpler integral process  $A_n^*$ , but the common scale of their fluctuations is of smaller order than that of  $A_n(\infty)$ , which, by Corollary 3.4, has variance of order  $\Psi_2(n) \asymp \Phi^2(n) \log n$ .

We now turn to approximating  $A_n(\infty)$ . As before, it is enough to consider  $A_n(\tau_n)$ , which we can write in the form

$$A_n(\tau_n) = \int_0^{\tau_n} \Phi(ne^{-S_t}) dt = \int_0^{\tau_n} \Phi(ne^{-S_{\tau_n-v}}) dv. \quad (5.5)$$

We now define the process  $\widehat{Z}_n$  by the equation

$$\widehat{Z}_n(v) := \begin{cases} S_{\tau_n-} - v - S_{(\tau_n-v)-} & \text{for } v < \tau_n, \\ S_{\tau_n-} - \tau_n & \text{for } v \geq \tau_n, \end{cases} \quad (5.6)$$

and we look for a suitable approximation to  $A_n(\tau_n)$  when the paths of  $\widehat{Z}_n$  are ‘nice’.

Very much as before, we define good events, for  $\varphi, \psi > 0$ ,

$$\widehat{B}_1(\varphi, n) := \left\{ \sup_{0 \leq v \leq 2 \log n} (v \vee 1)^{-1/2} |\widehat{Z}_n(v)| \leq \varphi \right\}; \quad (5.7)$$

$$\widehat{B}_2(\psi, n) := \{\log n - S_{\tau_n-} \leq \psi\}; \quad (5.8)$$

$$\widehat{B}_3(T, \psi, n) := \left\{ \int_0^T v^{-1} |\widehat{Z}_n(v)| dv \leq \psi \right\}, \quad (5.9)$$

whose probabilities we wish to show are large. The next two lemmas make this precise; we recall the definition (4.2) of the event  $B_0(n)$ .

**Lemma 5.3** *For any  $T, \varphi, \psi > 0$ , we have*

$$\begin{aligned} \mathbb{P}[\widehat{B}_1^c(\varphi, n) \cap B_0(n)] &\leq \sigma^2 \varphi^{-2} (72 + 29 \log \log n); \\ \mathbb{P}[\widehat{B}_3^c(T, \psi, n) \cap B_0(n)] &\leq 6\sigma \psi^{-1} \sqrt{T}. \end{aligned}$$

**Proof.** In order to make the calculations, it is convenient to exploit the explicit Itô construction of the process  $S$  [3, Proposition 1.3]. For  $H$  a Poisson point process on  $\mathbb{R}_+^2$  with intensity measure  $dt \nu_0(dx)$  we can define

$$S_t := \int_{]0,t] \times \mathbb{R}_+} x H(dt dx); \quad S_t^{(2)} := \int_{]0,t] \times \mathbb{R}_+} x^2 H(dt dx),$$

$S$  being a copy of our original subordinator. We also define the family of random point measures  $\mu_t$  on  $\mathbb{R}_+$  by

$$\mu_t[x, \infty] := H(]0, t[ \times [x, \infty]).$$

We then define the family of  $\sigma$ -fields

$$\mathcal{F}_{-t} := \sigma\{\mu_t, H|_{]t, \infty[ \times \mathbb{R}_+}\}, \quad t \geq 0,$$

so that  $\mathcal{F}_s \subset \mathcal{F}_{s'}$  whenever  $s \leq s' \leq 0$ . Then direct calculations show that the processes  $(M^{(l)}(t), t > 0)$ ,  $l = 1, 2, 3$ , are reversed martingales with respect to the filtration  $\{\mathcal{F}_s, s < 0\}$ , with means 1,  $\sigma^2$  and zero, respectively, where

$$M^{(1)}(t) := t^{-1}S_{t-}, \quad M^{(2)}(t) := t^{-1}S_{t-}^{(2)} \quad \text{and} \quad M^{(3)}(t) := (t^{-1}S_{t-} - 1)^2 - t^{-2}S_{t-}^{(2)}.$$

Thus it is immediate from the optional sampling theorem that

$$\mathbb{E}\{\tau_n^{-1}S_{\tau_n-}\} \leq 1; \quad \mathbb{E}\{\tau_n^{-1}S_{\tau_n-}^{(2)}\} \leq \sigma^2. \quad (5.10)$$

It also follows that  $\mathbb{E}M^{(3)}(t \vee \tau_n) = 0$  for any  $t > 0$ , which, taking  $t = \frac{1}{2} \log n$ , implies that

$$\mathbb{E}\{(\tau_n^{-1}S_{\tau_n-} - 1)^2 \mathbf{1}\{\tau_n \geq \frac{1}{2} \log n\}\} \leq 2\sigma^2 / \log n. \quad (5.11)$$

Furthermore, for  $v < \tau_n$ , the equality  $\mathbb{E}\{M^{(3)}(\tau_n - v) | \mathcal{F}_{-\tau_n}\} = M^{(3)}(\tau_n)$  a.s. also implies that, for such  $v$ ,

$$\mathbb{E}\{U_n(v)^2 | \mathcal{F}_{-\tau_n}\} = v \frac{\tau_n^{-1}S_{\tau_n-}^{(2)}}{\tau_n(\tau_n - v)}, \quad (5.12)$$

where

$$U_n(v) := (\tau_n - v)^{-1}S_{(\tau_n - v)-} - \tau_n^{-1}S_{\tau_n-}.$$

We thus have the expression

$$\widehat{Z}_n(v) = (v \wedge \tau_n)\{\tau_n^{-1}S_{\tau_n-} - 1\} - (\tau_n - v)_+ U_n(v), \quad v \geq 0, \quad (5.13)$$

as an alternative representation for  $\widehat{Z}_n$ , in addition to (5.6). Taking expectations conditional of  $\mathcal{F}_{-\tau_n}$ , we thus obtain

$$\begin{aligned} \mathbb{E}\{|\widehat{Z}_n(v)| | \mathcal{F}_{-\tau_n}\} &\leq (v \wedge \tau_n)|\tau_n^{-1}S_{\tau_n-} - 1| + (\tau_n - v)_+ \mathbb{E}\{|U_n(v)| | \mathcal{F}_{-\tau_n}\} \\ &\leq v|\tau_n^{-1}S_{\tau_n-} - 1| + \sqrt{v} \sqrt{\tau_n^{-1}S_{\tau_n-}^{(2)}}, \end{aligned}$$

the last inequality from (5.12). Multiplying by  $\mathbf{1}\{\tau_n \geq \frac{1}{2} \log n\}$  and taking expectations thus yields

$$\mathbb{E}(|\widehat{Z}_n(v)| \mathbf{1}\{\tau_n \geq \frac{1}{2} \log n\}) \leq v\sigma\sqrt{2/\log n} + \sigma\sqrt{v} \leq 3\sigma\sqrt{v}, \quad (5.14)$$

for  $0 \leq v \leq 2 \log n$ , in view of (5.10) and (5.11). The second inequality now follows from Markov's inequality, because

$$\mathbb{E} \left\{ \int_0^T v^{-1} |\widehat{Z}_n(v)| dv \mathbf{1}\{B_0(n)\} \right\} \leq 3\sigma \int_0^T v^{-1/2} dv.$$

It also follows from (5.13) that, for any  $\varphi > 0$  and for  $v < \tau_n$ ,

$$\{|\widehat{Z}_n(v)| > \varphi\sqrt{v \vee 1}\} \subset \{|\tau_n^{-1}S_{\tau_n-} - 1| > \frac{1}{2}\varphi v^{-1/2}\} \cup \{(\tau_n - v)|U_n(v)| > \frac{1}{2}\varphi v^{1/2}\}.$$

The first event happens for some  $v < \tau_n$  only if  $|\tau_n^{-1}S_{\tau_n-} - 1| > \frac{1}{2}\varphi\tau_n^{-1/2}$ , and the probability of this happening on the event  $B_0(n)$  is at most

$$\begin{aligned} & \mathbb{P} \left[ |\tau_n^{-1}S_{\tau_n-} - 1| \mathbf{1}\{\tau_n \geq \frac{1}{2} \log n\} > \frac{1}{2}\varphi(2 \log n)^{-1/2} \right] \\ & \leq \frac{2\sigma^2}{\log n} \cdot \frac{8 \log n}{\varphi^2} = 16\sigma^2\varphi^{-2}, \end{aligned}$$

by (5.11). For the second, using Kolmogorov's inequality much as in the proof of Lemma 4.1, for  $r \geq 1$  such that  $2^r \leq \frac{1}{2}\tau_n$ , we have

$$\begin{aligned} & \mathbb{P} \left[ \sup_{2^{r-1} \leq v \leq 2^r} v^{-1/2}(\tau_n - v)|U_n(v)| > \frac{1}{2}\varphi \mid \mathcal{F}_{-\tau_n} \right] \\ & \leq \mathbb{P} \left[ \sup_{0 \leq v \leq 2^r} |U_n(v)| > \frac{1}{2}\tau_n^{-1}\varphi 2^{(r-1)/2} \mid \mathcal{F}_{-\tau_n} \right] \\ & \leq \frac{4\mathbb{E}\{|U_n(2^r)|^2 \mid \mathcal{F}_{-\tau_n}\} \tau_n^2}{\varphi^2 2^{r-1}} \\ & \leq 16\varphi^{-2}\tau_n^{-1}S_{\tau_n-}^{(2)}, \end{aligned}$$

using (5.12). Adding over all such  $r$ , and including the  $v$ -intervals  $]0, 1[$  and  $]2^r, \frac{1}{2}\tau_n]$ , it follows that

$$\mathbb{P} \left[ \sup_{0 \leq v \leq \tau_n/2} v^{-1/2}(\tau_n - v)|U_n(v)| \mathbf{1}\{B_0(n)\} > \frac{1}{2}\varphi \right] \leq 20\sigma^2\varphi^{-2}(1 + \lceil \log_2 \log n \rceil), \quad (5.15)$$

from (5.10). For  $\frac{1}{2} \log n < v < \tau_n$ , we use (5.6) to give

$$\widehat{Z}_n(v) = Z_{\tau_n-} - Z_{(\tau_n-v)-},$$

so that

$$B_0(n) \cap \left\{ \sup_{\frac{1}{2}\tau_n \leq v \leq \tau_n} v^{-1/2}|\widehat{Z}_n(v)| > \varphi \right\} \subset \left\{ \sup_{0 \leq u \leq 2 \log n} |Z_u| > \frac{1}{4}\varphi\sqrt{\log n} \right\},$$

the latter event, by Kolmogorov's inequality, having probability at most  $32\sigma^2\varphi^{-2}$ . Finally, again by Kolmogorov's inequality,

$$\mathbb{P}[B_0(n)^c] \leq \mathbb{P}\left[\sup_{0 \leq u \leq 2\log n} |Z_u| > \frac{1}{2} \log n\right] \leq 8\sigma^2/\log n.$$

From these last two bounds and from (5.15), the lemma follows.  $\square$

**Lemma 5.4** *If  $\psi_n \rightarrow \infty$ , then  $\lim_{n \rightarrow \infty} \mathbb{P}[\widehat{B}_2^c(\psi_n, n)] = 0$ .*

**Proof.** Simply note that  $\widehat{B}_2^c(x, n) \subset \widehat{B}_2^c(y, n)$  whenever  $x > y$ , so that then

$$\pi_n(x) := \mathbb{P}[\widehat{B}_2^c(x, n)] \leq \pi_n(y),$$

and that

$$\lim_{n \rightarrow \infty} \pi_n(x) = \frac{\int_x^\infty N_0(u) du}{\int_0^\infty N_0(u) du} =: \pi(x),$$

by the renewal theorem [2, p. 99], with  $\lim_{x \rightarrow \infty} \pi(x) = 0$ . Hence, given  $\varepsilon > 0$ , pick  $x$  so that  $\pi(x) < \varepsilon/2$ , and then  $n_x$  such that  $\pi_n(x) \leq \varepsilon$  and  $\psi_n \geq x$  for all  $n \geq n_x$ ; it then follows that  $\pi_n(\psi_n) \leq \pi_n(x) \leq \varepsilon$  for all  $n \geq n_x$ .  $\square$

With these preparations, we are now in a position to approximate the behaviour of  $A_n(\tau_n)$ , and indeed of the whole process  $A_n(t \wedge \tau_n)$ .

**Theorem 5.5** *Suppose that Assumptions A1–A2 and (5.1) hold. Fix  $\alpha = 1/8$ ,  $\beta = 1/4$ , and set  $v_n := 4 \log^{2\alpha} n$ . Then, on the event*

$$B_0(n) \cap \widehat{B}_1(\log^\alpha n, n) \cap \widehat{B}_2(\log^\beta n, n) \cap \widehat{B}_3(2 \log n, \sqrt{c_2^*(e^{v_n}) \log n}, n),$$

it follows that

$$\left(\Phi(n)\sqrt{\log n}\right)^{-1} \sup_{t \geq 0} \left| A_n(t \wedge \tau_n) - \int_{(\tau_n - t)_+}^{\tau_n} \Phi(e^v) dv \right| \rightarrow 0.$$

Furthermore,

$$\lim_{n \rightarrow \infty} \mathbb{P}[B_0(n) \cap \widehat{B}_1(\log^\alpha n, n) \cap \widehat{B}_2(\log^\beta n, n) \cap \widehat{B}_3(2 \log n, \sqrt{c_2^*(e^{v_n}) \log n}, n)] = 1.$$

Here,  $c_2^*(m)$  is defined as in Lemma 5.1.

**Proof.** Recalling (5.5), we can write

$$\begin{aligned} A_n(t \wedge \tau_n) &= \int_0^{(t \wedge \tau_n)} \Phi(ne^{-S_u}) du \\ &= \int_{(\tau_n - t)_+}^{\tau_n} \Phi(e^v) dv + \int_{(\tau_n - t)_+}^{\tau_n} \{\Phi(e^{v+D_n}) - \Phi(e^v)\} dv \\ &\quad + \int_{(\tau_n - t)_+}^{\tau_n} \{\Phi(e^{v+D_n+\widehat{Z}_n(v)}) - \Phi(e^{v+D_n})\} dv, \end{aligned} \tag{5.16}$$

where  $D_n := \log n - S_{\tau_n-} \geq 0$ . The second of the integrals in (5.16) is nonnegative, and no larger than

$$\begin{aligned} \int_0^{\tau_n} \{\Phi(e^{v+D_n}) - \Phi(e^v)\} dv &\leq \int_{\tau_n}^{\tau_n+D_n} \Phi(e^v) dv \\ &\leq \Phi(n^3) \log^\beta n, \end{aligned} \quad (5.17)$$

on the event  $B_0(n) \cap \widehat{B}_2(\log^\beta n, n)$ . Note also that, for any  $r \geq 1$  and  $n$  such that  $c_2^*(n) \geq 1$ ,

$$\begin{aligned} 1 &\leq \frac{\Phi(n^r)}{\Phi(n)} = \exp \left\{ \int_n^{n^r} \frac{dy}{yL(y)} \right\} \\ &\leq \exp \{(\log \log(n^r) - \log \log n)/c_2^*(n)\} = r^{1/c_2^*(n)} \leq r; \end{aligned} \quad (5.18)$$

hence, from (5.17), the second of the integrals in (5.16) is of smaller order than  $\Phi(n)\sqrt{\log n}$  on the event  $B_0(n) \cap \widehat{B}_2(\log^\beta n, n)$ .

To control the third of the integrals in (5.16), we bound

$$\int_0^{\tau_n} \left| \Phi(e^{v+D_n+\widehat{Z}_n(v)}) - \Phi(e^{v+D_n}) \right| dv. \quad (5.19)$$

On  $\widehat{B}_1(\log^\alpha n, n)$ , we have  $v^{-1}|\widehat{Z}_n(v)| \leq 1/2$  if  $v \geq v_n$ . So split the range of the integral into  $0 < v \leq v_n$  and  $v_n \leq v \leq \tau_n$ . In the lower range, on  $\widehat{B}_1(\log^\alpha n, n) \cap \widehat{B}_2(\log^\beta n, n)$ , the exponents  $v + D_n$  and  $v + D_n + \widehat{Z}_n(v)$  are bounded above by

$$v_n + \log^\beta n + \log^\alpha n \sqrt{v_n} \leq 7 \log n,$$

implying, together with (5.18), that (5.19) is bounded above by  $7\Phi(n)v_n$  for all  $n$  large enough, and this is  $o(\Phi(n)\sqrt{\log n})$  by choice of  $\alpha$ . In the upper range, we can apply Lemma 3.2 to  $\Phi(e^{v+D_n+\widehat{Z}_n(v)})$ , very much as in the proof of Theorem 4.4, because here  $|\widehat{Z}_n(v)| \leq \frac{1}{2}(v + D_n)$ . The quantity  $\widehat{X}(n, v)$ , analogous to  $X(n, t)$  of (4.15), is bounded for  $v \geq v_n$  by

$$\widehat{X}(n, v) \leq \frac{\varkappa |\widehat{Z}_n(v)|^2}{c_2(e^{v+D_n})v(v + D_n)} \leq \frac{\varkappa \log^{2\alpha} n}{vc_2^*(e^{v_n})} \leq \frac{\varkappa}{c_2^*(e^{v_n})},$$

and

$$\frac{|\widehat{Z}_n(v)|}{L(e^{v+D_n})} \leq \frac{\log^\alpha n}{c_2^*(e^{v_n})\sqrt{v}} \leq \frac{1}{2c_2^*(e^{v_n})},$$

giving

$$\begin{aligned} &\int_{v_n}^{\tau_n} \left| \Phi(e^{v+D_n+\widehat{Z}_n(v)}) - \Phi(e^{v+D_n}) \right| dv \\ &\leq \int_{v_n}^{\tau_n} \Phi(e^{v+D_n}) \frac{|\widehat{Z}_n(v)|}{vc_2^*(e^v)} dv + \lambda \left( \frac{\Phi(e^{\tau_n+D_n}) \log^{2\alpha} n}{c_2^*(e^{v_n})} \int_{v_n}^{\tau_n} v^{-1} dv \right), \end{aligned} \quad (5.20)$$

for some positive constant  $\lambda < \infty$ . On  $B_0(n) \cap \widehat{B}_2(\log^\beta n, n)$ , and from (5.18), we have  $\Phi(e^{\tau_n+D_n}) \leq 3\Phi(n)$  for all  $n$  large enough, so that the second term in (5.20) is of order  $o(\Phi(n)\sqrt{\log n})$ . The first term is bounded on

$$B_0(n) \cap \widehat{B}_2(\log^\beta n, n) \cap \widehat{B}_3(2 \log n, \sqrt{c_2^*(e^{v_n}) \log n}, n)$$

by

$$\Phi(n^3)\sqrt{\log n}/\sqrt{c_2^*(e^{v_n})} = o(\Phi(n)\sqrt{\log n}),$$

again by (5.18). This completes the proof of the main statement. The final assertion follows from Lemmas 4.1, 5.3 and 5.4.  $\square$

## 6 Approximation theorems

We can now build on the results of the previous sections to derive central limit approximations for  $\mathcal{K}_n$ . The starting point is the functional central limit theorem for the Lévy process itself. Defining the process  $W_m$  by  $W_m(t) := \sigma^{-1}m^{-1/2}Z_{mt}$ , it follows that

$$W_m \rightarrow_d W \quad \text{in } D_1[0, \infty[ \quad \text{as } m \rightarrow \infty, \quad (6.1)$$

where  $W$  is standard Brownian motion and  $D_1[0, \infty[$  denotes the space of càdlàg functions  $x: [0, \infty[ \rightarrow \mathbb{R}$  satisfying  $\lim_{t \rightarrow \infty} t^{-1}x(t) = 0$ , endowed with the metric  $\rho_1(x, y) := \sup_{t \geq 0} (t \vee 1)^{-1}|x(t) - y(t)|$  (Müller [13], Satz 1). As a consequence of the central limit theorem for the renewal processes [6, Section XI.5], it also follows that

$$U_n := (\tau_n - \log n)/\{\sigma\sqrt{\log n}\} \rightarrow_d \mathcal{N}(0, 1) \quad \text{as } n \rightarrow \infty. \quad (6.2)$$

We shall also be interested in approximations which are not given in the form of limit theorems, but are instead expressed in terms of bounds on a distance between the distributions of the processes considered, taken here to be the appropriate bounded Wasserstein distances. For probability measures  $Q$  and  $Q'$  on a metric space  $(\mathcal{X}, \rho)$ , the bounded Wasserstein distance  $d_{BW}(Q, Q')$  is defined to be  $\sup_{f \in \mathcal{W}} |\int f dQ - \int f dQ'|$ , where  $\mathcal{W}$  denotes the bounded Lipschitz functions on  $\mathcal{X}$ :

$$\mathcal{W} := \mathcal{W}_{\mathcal{X}, \rho} := \{f: \mathcal{X} \rightarrow \mathbb{R} : \|f\| \leq 1, L(f) \leq 1\},$$

and  $L(f) := \sup_{x \neq x' \in \mathcal{X}} |f(x) - f(x')|/\rho(x, x')$ . The distance  $d_{BW}$  metrises weak convergence in  $(\mathcal{X}, \rho)$  (Dudley [5], Theorem 8.3). Note also that if, for each  $n \geq 1$ , the random elements  $X_n$  and  $Y_n$  of  $(\mathcal{X}, \rho)$  are on the same probability space, then  $d_{BW}(\mathcal{L}(X_n), \mathcal{L}(Y_n)) \rightarrow 0$  if, for each  $\varepsilon > 0$ ,  $\mathbb{P}[\rho(X_n, Y_n) > \varepsilon] \rightarrow 0$ . If  $\mathcal{X}$  is the space  $D_1[0, \infty[$  defined above, we shall refer to  $\mathcal{W}^1$  and  $d_{BW}^1$ ; if  $\mathcal{X}$  is the space  $D_0[0, \infty[$  of càdlàg functions  $x: [0, \infty[ \rightarrow \mathbb{R}$  having finite limits as  $t \rightarrow \infty$ , endowed with the metric  $\rho_0(x, y) := \sup_{t \geq 0} |x(t) - y(t)|$ , we shall refer to  $\mathcal{W}^0$  and  $d_{BW}^0$ , and, if  $\mathcal{X} = \mathbb{R}$ , we shall write  $d_{BW}^{\mathbb{R}}$ .

### 6.1 Moderate growth

We begin with a setting of moderate growth, in which  $L(n) \asymp \log n$ , so that (4.9) is in force. In order to describe the behaviour of  $\mathcal{K}_n$ , we first define a centred and normalized version  $\mathcal{K}_n^{(1)}$  of the process by

$$\mathcal{K}_n^{(1)}(u) := \left(\Phi(n)\sqrt{\log n}\right)^{-1} \left\{ \mathcal{K}_n(u \log n) - \log n \int_0^{(u \wedge 1)} \Phi(n^{1-v}) dv \right\},$$



whose distribution we approximate by that of  $Y_n^{(1)}$ , where

$$Y_n^{(1)}(u) := \sigma \int_0^{(u \wedge 1)} h_n^{(1)}(v) W(v) dv,$$

with

$$h_n^{(1)}(u) := \frac{\Phi(n^{1-u}) \log n}{\Phi(n)L(n^{1-u})}.$$

Note that  $h_n^{(1)}(u) \geq 0$  for all  $u$ , and that, from (5.3),

$$\int_0^1 h_n^{(1)}(u) du = 1 - \Phi(1)/\Phi(n) \leq 1 \quad (n \geq 1).$$

**Theorem 6.1** *If Assumptions A1–A2 hold, and  $L(n) \asymp \log n$ , then*

$$d_{BW}(\mathcal{L}(\mathcal{K}_n^{(1)}), \mathcal{L}(Y_n^{(1)})) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

**Proof.** We begin by writing

$$\begin{aligned} \mathcal{K}_n^{(1)}(u) - \tilde{Y}_n^{(1)}(u) &= r_n \left\{ (\mathcal{K}_n(u \log n) - A_n(u \log n)) \right. \\ &\quad \left. + (A_n(u \log n) - A_n(\{u \log n\} \wedge \tau_n)) + (A_n(\{u \log n\} \wedge \tau_n) - A_n^*(u \log n)) \right\}, \end{aligned}$$

where  $r_n := (\Phi(n)\sqrt{\log n})^{-1}$  and

$$\begin{aligned} \tilde{Y}_n^{(1)} &:= A_n^*(u \log n) - \log n \int_0^{u \wedge 1} \Phi(n^{1-v}) dv \\ &= -\sigma \int_0^{(u \wedge 1)} h_n^{(1)}(v) W_{\log n}(v) dv. \end{aligned}$$

Now we have  $r_n \sup_{u \geq 0} |\mathcal{K}_n(u \log n) - A_n(u \log n)| \rightarrow_p 0$  by Lemma 2.6 and Corollary 3.4, then  $r_n \sup_{u \geq 0} |A_n(u \log n) - A_n(\{u \log n\} \wedge \tau_n)| \rightarrow_p 0$  by Lemma 2.2, and finally, by Theorem 4.4,  $r_n \sup_{u \geq 0} |A_n(\{u \log n\} \wedge \tau_n) - A_n^*(u \log n)| \rightarrow_p 0$ . Hence it follows that

$$d_{BW}^0(\mathcal{L}(\mathcal{K}_n^{(1)}), \mathcal{L}(\tilde{Y}_n^{(1)})) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

To conclude the proof, we now need to show that  $\sup_{f \in \mathcal{W}^0} |\mathbb{E}f(\tilde{Y}_n^{(1)}) - \mathbb{E}f(Y_n^{(1)})| \rightarrow 0$  as  $n \rightarrow \infty$ . To do so, for any  $f \in \mathcal{W}^0$ , define  $f_n: D_1[0, \infty[ \rightarrow \mathbb{R}$  by  $f_n(w) := f(H_n(w))$ , where  $H_n(w)(u) := \int_0^{(u \wedge 1)} h_n^{(1)}(v) w(v) dv$ . Note that, for  $w, w' \in D_1[0, \infty[$  and any  $u \geq 0$ ,

$$\begin{aligned} |H_n(w)(u) - H_n(w')(u)| &= \left| \int_0^{(u \wedge 1)} h_n^{(1)}(v) (w(v) - w'(v)) dv \right| \\ &\leq \rho_1(w, w') \int_0^1 h_n^{(1)}(v) dv \leq \rho_1(w, w'). \end{aligned}$$

Hence, for any  $f \in \mathcal{W}^0$ , it follows that  $f_n \in \mathcal{W}^1$ , and hence that

$$|\mathbb{E}f(\tilde{Y}_n^{(1)}) - \mathbb{E}f(Y_n^{(1)})| = |\mathbb{E}f_n(W_{\log n}) - \mathbb{E}f_n(W)| \leq d_{BW}^1(\mathcal{L}(W_{\log n}), \mathcal{L}(W)).$$

The theorem now follows from (6.1). □

**Theorem 6.2** *Under the assumptions of Theorem 6.1, if in addition  $L(n) \sim \gamma \log n$  for some  $0 < \gamma < \infty$ , then*

$$\mathcal{K}_n^{(1)} \rightarrow_d Y^{(1)} \quad \text{in } D_0[0, \infty[ \quad \text{as } n \rightarrow \infty,$$

where

$$Y^{(1)}(u) := \sigma \int_0^{(u \wedge 1)} \gamma^{-1} (1-v)^{(1-\gamma)/\gamma} W(v) dv.$$

**Proof.** If  $L(n) \sim \gamma \log n$ , then  $h_n^{(1)}(u) \rightarrow \gamma^{-1} (1-u)^{(1-\gamma)/\gamma}$  uniformly in  $0 \leq u \leq 1 - \delta$ , for any  $\delta > 0$ ; furthermore,

$$\limsup_{n \rightarrow \infty} \int_{1-\delta}^1 h_n^{(1)}(v) dv = \limsup_{n \rightarrow \infty} \{\Phi(n^\delta) - \Phi(1)\} / \Phi(n) \leq \delta^{\gamma'}$$

for any  $\gamma' < \gamma$ , and  $\int_{1-\delta}^1 h^{(1)}(v) dv = \delta^{\gamma'}$ . Hence

$$\mathbb{E} \left\{ \sup_{u \geq 0} |Y_n^{(1)}(u) - Y^{(1)}(u)| \right\} \leq \sqrt{\frac{2}{\pi}} \int_0^1 |h_n^{(1)}(v) - h(v)| dv \rightarrow 0,$$

proving the theorem. □

**Examples.** Suppose, for some  $0 < \gamma < \infty$ , that  $\tilde{S}$  is a subordinator such that  $\tilde{L}(n) \sim \gamma \log n$  and  $\tilde{\Phi}(n) \sim c \log^{1/\gamma} n$ ; as at the end of Section 2.4, we do not assume that  $\mathfrak{m} = \mathbb{E}\tilde{S}_1$  takes the value 1, and we write  $\tau^2 = \text{Var} \tilde{S}_1$ . Theorem 6.2 entails a gaussian limit for  $(K_n - \mu_n)/\sigma_n$ , with

$$\mu_n \sim \log n \int_0^1 \Phi(n^{1-v}) dv \sim \frac{c \log^{1+1/\gamma} n}{\mathfrak{m}(1 + 1/\gamma)}$$

and

$$\sigma_n^2 \sim \Phi^2(n) \log n \sigma^2 \text{Var} \left\{ \int_0^1 \gamma^{-1} (1-v)^{1/\gamma-1} W(v) dv \right\} \sim \frac{c^2 \tau^2 \log^{1+2/\gamma} n}{\mathfrak{m}^3 (1 + 2/\gamma)},$$

where, as before,  $\Phi(n) = \mathfrak{m}^{-1} \tilde{\Phi}(n)$  and  $\sigma^2 = \mathfrak{m}^{-1} \tau^2$ . Note also that  $\mu_n \sim \Psi(n) \sim \mathfrak{m}^{-1} \tilde{\Psi}(n)$  and that  $\sigma_n^2 \sim \sigma^2 \Psi_2(n) \sim \tau^2 \mathfrak{m}^{-3} \tilde{\Psi}_2(n)$ , as is to be expected.

For the classical gamma subordinator [2, p. 73], scaling so that  $\mathbb{E}S_1 = 1$ , we have  $\nu_0(dx) = \theta e^{-\theta x} dx/x$ ,  $\Phi_0(n) = \theta \log(1 + n/\theta)$ , and  $\sigma^2 = 1/\theta$ . Hence the CLT in [11] agrees with Theorem 6.2. Note that one parameter  $\theta > 0$  is enough, since, for the Lévy measure  $a\nu_0$ , the distribution of  $K_n$  does not depend on the scale parameter  $a$ .

In the case  $\gamma = 1$ , Theorem 6.2 covers a somewhat larger family of gamma-like subordinators than that considered in [11]. The extension is that the condition of exponential decay for  $N_0(x)$  as  $x \rightarrow \infty$  required in [11] is replaced now by a weaker condition  $\sigma^2 < \infty$ . The constraints on the behaviour of  $N(x)$  at  $x \rightarrow 0$  are also slightly weaker here.

## 6.2 Fast growth

We now turn to the setting in which  $L(n) \rightarrow \infty$  but  $L(n)/\log n \rightarrow 0$ ; hence  $\Phi$  grows faster than any power of the logarithm. In order to apply the previous theorems, we need to suppose either that (4.9) is in force, albeit with  $L(n) = o(\log n)$ , or that (4.20) holds. The analogue of  $\mathcal{K}_n^{(1)}$  is now  $\mathcal{K}_n^{(2)}$ , defined by

$$\mathcal{K}_n^{(2)}(u) := \left( \Phi(n) \sqrt{L(n)} \right)^{-1} \left\{ \mathcal{K}_n(uL(n)) - L(n) \int_0^{(u \wedge l_n)} \Phi(ne^{-vL(n)}) dv \right\},$$

where  $l_n := \log n / L(n)$ . Here, we approximate the distribution of  $\mathcal{K}_n^{(2)}$  by that of  $Y^{(2)}$ , where

$$Y^{(2)}(u) := \sigma \int_0^u e^{-v} W(v) dv.$$

**Theorem 6.3** *If Assumptions A1–A2 hold, and  $L(n)/\log n \rightarrow 0$ , with either (4.9) or (4.20) satisfied, then*

$$d_{BW}(\mathcal{L}(\mathcal{K}_n^{(2)}), \mathcal{L}(Y^{(2)})) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where  $d_{BW}$  is as before.

**Proof.** If (4.9) is satisfied, we argue as in the proof of Theorem 6.1 to show that

$$\sup_{u \geq 0} |\mathcal{K}_n^{(2)}(u) - \tilde{Y}_n^{(2)}(u)| \xrightarrow{p} 0 \quad \text{as } n \rightarrow \infty, \quad (6.3)$$

where

$$\tilde{Y}_n^{(2)}(u) := \sigma \int_0^{(u \wedge l_n)} h_n^{(2)}(v) W_{L(n)}(v) dv,$$

and

$$h_n^{(2)}(u) := \frac{\Phi(ne^{-uL(n)})L(n)}{\Phi(n)L(ne^{-uL(n)})}.$$

If (4.20) is satisfied, (6.3) is still true, using Theorem 4.6 in place of Theorem 4.4 in the proof. Once again,  $h_n^{(2)}(u) \geq 0$  for all  $u$ , and

$$\int_0^\infty h_n^{(2)}(u) du = 1.$$

The next step is to approximate  $\tilde{Y}_n^{(2)}$  by  $Y_n^{(2)}$ , where

$$Y_n^{(2)}(u) := \sigma \int_0^u e^{-v} W_{L(n)}(v) dv.$$

Here, it is immediate that

$$\mathbb{E} \left\{ \sup_{u \geq 0} \sigma^{-1} |\tilde{Y}_n^{(2)}(u) - Y_n^{(2)}(u)| \right\} \leq \int_0^{l'_n} |h_n^{(2)}(v) - e^{-v}| \sqrt{v} dv + \int_{l'_n}^\infty \{h_n^{(2)}(v) + e^{-v}\} \sqrt{v} dv \quad (6.4)$$

for any  $l'_n \leq l_n$ ; we take  $l'_n = \min\{l_n^{1/2}, \frac{1}{2(\varkappa \vee 1)}l_n\}$ . Now, from (3.3), for  $0 \leq v \leq l'_n$ , we have

$$\left| \frac{L(n)}{L(ne^{-vL(n)})} - 1 \right| \leq \lambda_1 l_n^{-1/2},$$

for some  $\lambda_1 < \infty$ , and hence, by Corollary 3.3, that

$$\sqrt{v} |h_n^{(2)}(v) - \Phi^*(n, v)| \leq \lambda_2 \sqrt{v} e^{-v/2} l_n^{-1/2},$$

where

$$\Phi^*(n, v) := \frac{\Phi(ne^{-vL(n)})}{\Phi(n)}.$$

Then Lemma 3.2 gives

$$\sqrt{v} |\Phi^*(n, v) - e^{-v}| \leq \lambda_3 \{\exp(\varkappa v^2/l_n) - 1\} \sqrt{v} e^{-v} \leq \lambda_4 l_n^{-1} v^{5/2} e^{-v},$$

for different constants  $\lambda_3, \lambda_4$ , again in  $0 \leq v \leq l'_n$ . Hence it follows that

$$\lim_{n \rightarrow \infty} \int_0^{l'_n} |h_n^{(2)}(v) - e^{-v}| \sqrt{v} dv = 0. \quad (6.5)$$

It is also immediate that  $\lim_{n \rightarrow \infty} \int_{l'_n}^{\infty} e^{-v} \sqrt{v} dv = 0$ . Hence, to show that the right hand side of (6.4) is small in the limit, it remains only to consider

$$\int_{l'_n}^{\infty} h_n^{(2)}(v) \sqrt{v} dv = \sqrt{l'_n} \Phi^*(n, l'_n) + \int_{l'_n}^{\infty} \frac{1}{2} v^{-1/2} \Phi^*(n, v) dv. \quad (6.6)$$

Here, the first term tends to zero as  $n \rightarrow \infty$  by Corollary 3.3, as does

$$\int_{l'_n}^{\alpha_n} v^{-1/2} \Phi^*(n, v) dv \leq \int_{l'_n}^{\alpha_n} v^{-1/2} e^{-v/2} dv,$$

where  $\alpha_n := l_n / \{2(\varkappa \vee 1)\}$ . Then, splitting the remaining integral at  $2l_n$ , we have

$$\int_{\alpha_n}^{\infty} v^{-1/2} \Phi^*(n, v) dv \leq 2l_n \exp\{-\alpha_n/2\} + \frac{1}{n\Phi(n)L(n)},$$

the final term following from (2.6). Combining these bounds, we have now also shown that  $\lim_{n \rightarrow \infty} \int_{l'_n}^{\infty} h_n^{(2)}(v) \sqrt{v} dv = 0$ ; hence, from (6.3) and (6.4), it follows that

$$d_{BW}^0(\mathcal{L}(\mathcal{K}_n^{(2)}), \mathcal{L}(Y_n^{(2)})) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Finally, if  $H: D_1[0, \infty[ \rightarrow D_0[0, \infty[$  is defined by  $H(w)(u) := \int_0^u e^{-v} w(v) dv$  and  $f$  is in  $\mathcal{W}^0$ , then  $(1 + e^{-1})^{-1} f \circ H \in \mathcal{W}^1$ , from which  $d_{BW}^0(\mathcal{L}(Y_n^{(2)}), \mathcal{L}(Y^{(2)})) \rightarrow 0$  as  $n \rightarrow \infty$  follows immediately, and the theorem is proved.  $\square$

### 6.3 Slow growth

If  $\Phi$  grows very slowly to infinity, with  $L(n)/\log n \rightarrow \infty$ , the arguments culminating in Theorem 5.5 show that the key quantity describing the process  $\mathcal{K}_n$  is the family of integrals

$$\int_{(\tau_n-t)_+}^{\tau_n} \Phi(e^v) dv, \quad t \geq 0.$$

Here, the randomness enters only through the hitting time  $\tau_n$ , which is asymptotically normally distributed, as recorded in (6.2). The process thus has a quite different qualitative behaviour to that of the previous cases.

Since  $\tau_n$  takes values fairly close to  $\log n$ , it makes sense to describe the random behaviour of  $\mathcal{K}_n(t)$  by first subtracting  $\int_{(\log n-t)_+}^{\log n} \Phi(e^v) dv$ , and then dividing by  $\Phi(n)\sqrt{\log n}$ . This leads us to define the process  $\mathcal{K}_n^{(3)}$  for  $t \geq 0$  by

$$\mathcal{K}_n^{(3)}(t) := \left(\Phi(n)\sqrt{\log n}\right)^{-1} \left\{ \mathcal{K}_n(t) - \int_{(\log n-t)_+}^{\log n} \Phi(e^v) dv \right\}.$$

Then, defining  $G_n: \mathbb{R} \rightarrow D_0[0, \infty[$  by

$$G_n[u](t) := \sigma u - \left(\Phi(n)\sqrt{\log n}\right)^{-1} \int_{(\log n-t)_+}^{(\log n-t+\sigma u\sqrt{\log n})_+} \Phi(e^v) dv, \quad t \geq 0, \quad (6.7)$$

for each  $u \in \mathbb{R}$ , we define our approximating process to be  $Y_n^{(3)} = G_n[U]$ , where  $U$  is a standard normal random variable.

**Theorem 6.4** *Suppose that Assumptions A1–A2 and (5.1) hold. Then it follows that*

$$d_{BW}^0(\mathcal{L}(\mathcal{K}_n^{(3)}), \mathcal{L}(Y_n^{(3)})) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

**Proof.** Once again, we combine Lemmas 2.2 and 2.6 and Corollary 3.4, this time with Theorem 5.5, showing that

$$d_{BW}^0(\mathcal{L}(\mathcal{K}_n^{(3)}), \mathcal{L}(\tilde{Y}_n^{(3)})) \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where

$$\begin{aligned} \tilde{Y}_n^{(3)}(t) &:= \left(\Phi(n)\sqrt{\log n}\right)^{-1} \left\{ \int_{(\tau_n-t)_+}^{\tau_n} \Phi(e^v) dv - \int_{(\log n-t)_+}^{\log n} \Phi(e^v) dv \right\} \\ &= \left(\Phi(n)\sqrt{\log n}\right)^{-1} \left\{ \int_{\log n}^{\tau_n} \Phi(e^v) dv - \int_{(\log n-t)_+}^{(\tau_n-t)_+} \Phi(e^v) dv \right\}. \end{aligned}$$

Now, from (5.18), on the set  $B_0(n)$ , we have

$$2^{-1/c_2^*(\sqrt{n})} \leq \Phi(e^v)/\Phi(n) \leq 2^{1/c_2^*(n)}$$

for all  $v$  between  $\tau_n$  and  $\log n$ , and hence

$$\mathbf{1}\{B_0(n)\} \left(\Phi(n)\sqrt{\log n}\right)^{-1} \left| \int_{\log n}^{\tau_n} \Phi(e^v) dv - \int_{\log n}^{\tau_n} \Phi(n) dv \right| \leq \sigma |(2^{1/c_2^*(\sqrt{n})} - 1)U_n|;$$

this tends to zero as  $n \rightarrow \infty$  because of (6.2) together with  $\lim_{n \rightarrow \infty} c_2^*(n) = \infty$ . On the other hand,  $\lim_{n \rightarrow \infty} \mathbb{P}[B_0(n)] = 0$ , by Lemma 4.1. Hence  $d_{BW}^0(\mathcal{L}(\tilde{Y}_n^{(3)}), \mathcal{L}(G_n[U_n])) \rightarrow 0$  as  $n \rightarrow \infty$ .

To complete the proof, we just have to show that the distributions of  $G_n[U_n]$  and  $G_n[U]$  are close. To do so, we first define  $\tilde{G}_n: \mathbb{R} \rightarrow D_0[0, \infty[$  by  $\tilde{G}_n[u] = G_n[u \wedge \sigma^{-1}\sqrt{\log n}]$ . Then, from the definition of  $G_n$  and from (5.18), it is immediate that

$$\begin{aligned} \sup_{v \geq 0} |\tilde{G}_n[u'](v) - \tilde{G}_n[u](v)| &\leq \sigma|u' - u| + \left(\Phi(n)\sqrt{\log n}\right)^{-1} \Phi(n^2)\sigma\sqrt{\log n}|u' - u| \\ &\leq 3\sigma|u' - u|; \end{aligned}$$

hence  $\lim_{n \rightarrow \infty} \sup_{f \in \mathcal{W}^0} |\mathbb{E}f(\tilde{G}_n[U_n]) - \mathbb{E}f(\tilde{G}_n[U])| = 0$ , in view of (6.2). Finally,

$$\begin{aligned} &\sup_{f \in \mathcal{W}^0} \{|\mathbb{E}f(G_n[U_n]) - \mathbb{E}f(\tilde{G}_n[U_n])| + |\mathbb{E}f(G_n[U]) - \mathbb{E}f(\tilde{G}_n[U])|\} \\ &\leq 2\mathbb{P}[U_n > \sigma^{-1}\sqrt{\log n}] + 2\mathbb{P}[U > \sigma^{-1}\sqrt{\log n}] \rightarrow 0, \end{aligned}$$

and the theorem follows.  $\square$

The process  $Y_n^{(3)}$  starts close to zero, and, as indicated by Lemmas 5.1 and 5.2, remains close to zero until  $1 - t/\log n$  becomes small. It reaches its final value  $\sigma U$  at time  $\log n + \sigma U\sqrt{\log n}$  if  $U \geq 0$ , and at time  $\log n$  if  $U < 0$ .

Its behaviour can also be understood in terms of the overlapping representation provided under the condition (4.19), when  $c_2(n)$  is allowed to tend to infinity, but not too fast. Here, the approximation to the random fluctuations is expressed in terms of the process

$$\left(\Phi(n)\sqrt{\log n}\right)^{-1} \int_0^{T \wedge \log n} \Phi(ne^{-t}) \frac{Z_t}{L(ne^{-t})} dt,$$

which at first sight looks very different. Here, however, as already observed at the start of Section 5,

$$\left(\Phi(n)\sqrt{\log n}\right)^{-1} \int_0^{(u \log n) \wedge \log n} \Phi(ne^{-t}) \frac{Z_t}{L(ne^{-t})} dt$$

is of small order whenever  $u$  is bounded away from 1, and even for choices of  $u = u(n) \rightarrow 1$  such that  $(1 - u(n))c_2(n^{1-u(n)}) \rightarrow \infty$ . On the other hand, for  $u$  closer to 1, the remaining contribution is approximately

$$\begin{aligned} &\left(\Phi(n)\sqrt{\log n}\right)^{-1} Z_{\log n} \int_{u(n)\log n}^{u \log n} \frac{\Phi(ne^{-t})}{L(ne^{-t})} dt \\ &= \left(\Phi(n)\sqrt{\log n}\right)^{-1} Z_{\log n} \{\Phi(n^{1-u(n)}) - \Phi(n^{1-u})\} \\ &\approx \left\{1 - \frac{\Phi(n^{1-u})}{\Phi(n)}\right\} \sigma W_{\log n}(1), \end{aligned}$$

whose randomness is determined only by the value of  $W_{\log n}(1) \sim -(\tau_n - \log n)/\sigma\sqrt{\log n}$ . To match this with the corresponding formula for  $Y_n^{(3)}(t)$ , note that, under (4.19), the

second term in  $G_n[U_n]$ ,

$$\left(\Phi(n)\sqrt{\log n}\right)^{-1} \int_{(\log n - t)_+}^{(\log n - t + \sigma U_n \sqrt{\log n})_+} \Phi(e^v) dv,$$

is small for  $\log n - t = O(\sqrt{\log n})$ , and that, for larger values of  $\log n - t = (1 - u) \log n$ , one can replace  $\Phi(e^v)$  by  $\Phi(n^{1-u})$  in the integral.

**Remark.** Setting formally  $\Phi(n) = \text{const}$  in the above formulas suggests that  $K_n \sim \tau_n$  in the case of bounded  $\nu_0$ . The latter is indeed true and, moreover,  $|K_n - \tau_n|$  remains bounded with all moments as  $n$  grows; the reason for this behaviour in the compound Poisson case is just that essentially all gaps within  $\mathcal{R} \cap [0, \log n]$  are hit by the atoms of  $Y_n$ , hence  $K_n$  is close to the number of renewals on  $[0, \log n]$ .

## Acknowledgement

We wish to express our warm thanks to the referee for a careful and constructive reading of the paper.

## References

- [1] R. Arratia, A.D. Barbour and S. Tavaré. *Logarithmic Combinatorial Structures: A Probabilistic Approach*, European Math. Soc. Monographs in Math., v. 1, 2003.
- [2] J. Bertoin, *Lévy Processes*, Cambridge University Press, 1996.
- [3] J. Bertoin, *Subordinators: Examples and Applications*, Springer Lecture Notes in Math. vol. 1727, 1996.
- [4] N.H. Bingham, C.M. Goldie and J.L. Teugels. *Regular Variation*, Cambridge University Press, 1987.
- [5] R.M. Dudley. *Probabilities and metrics*, Lecture Notes Series **45**, Aarhus Universitet, 1976.
- [6] W. Feller, *An Introduction to Probability Theory and its Applications*, volume II, Wiley, 2nd edition, 1971.
- [7] A.V. Gnedin. The Bernoulli sieve, *Bernoulli* 10: 79–96, 2004.
- [8] A.V. Gnedin and J. Pitman. Regenerative composition structures, *Ann. Probab.* 33: 445–479, 2005.
- [9] A.V. Gnedin and J. Pitman. Regenerative partition structures, *Elec. J. Combin.* 11(2), paper #R12, 2004/2005.

- [10] A.V. Gnedenko, J. Pitman and M. Yor. Asymptotic laws for compositions derived from transformed subordinators, *Ann. Probab.* (to appear), <http://arxiv.org/abs/math.PR/0403438>, 2004.
- [11] A.V. Gnedenko, J. Pitman and M. Yor. Asymptotic laws for regenerative compositions: gamma subordinators and the like, *Probab. Th. Rel. Fields* 2005.
- [12] G. Last and A. Brandt. *Marked Point Processes on the Real Line: The Dynamic Approach*, Springer, NY, 1995.
- [13] D.W. Müller. Verteilungs-Invarianzprinzipien für das starke Gesetz der großen Zahl, *Z. Wahrscheinlichkeitstheorie verw. Geb.* 10: 173–192, 1968.
- [14] J. Pitman, *Combinatorial Stochastic Processes*, Springer Lecture Notes Math. (to appear), 2005.
- [15] M. Winkel, Electronic foreign exchange markets and level passage events of multivariate subordinators, *J. Appl. Prob.* 42: 138–152, 2005.