A Characterization of Small and Large Time Limit Laws for Self-Normalized Lévy Processes

Ross Maller* and David M. Mason†
Australian National University and University of Delaware

Abstract

We establish asymptotic distribution results for self-normalized Lévy processes at small and large times that are analogs of those of Chistyakov and Götze (2004) for self-normalized sums.

Dedicated to Friedrich Götze on the occasion of his sixtieth birthday.

Short title: Small and Large Time Limit Laws for Self-Normalized Lévy Processes
AMS 2000 Subject Classifications: 60F05, 60F17, 60G51.
Keywords: Lévy Process, Self-Normalized, Quadratic Variation, Domain of Attraction of Normal Distribution, Stable Laws, Small Times, Large Times.

1 Introduction and statements of main results

Let ξ, $\xi_1, \xi_2, \ldots,$ be i.i.d. nondegenerate random variables with common distribution function F. For each $n \geq 1$ let $\mathbb{S}_n = \xi_1 + \cdots + \xi_n$ and $\mathbb{V}_n = \xi_1^2 + \cdots + \xi_n^2$. Consider the self–normalized sum

$$\mathbb{T}_n := \mathbb{S}_n / \sqrt{\mathbb{V}_n}. \tag{1.1}$$

(Here and elsewhere $0/0 := 0$.) Giné, Götze and Mason (1997) proved that $\mathbb{T}_n$ converges in distribution to a standard normal random variable Z if and only if F is in the domain of attraction of a normal law, written $F \in D(N)$, and $E\xi = 0$. This verified part of a conjecture of Logan, Mallows, Rice and Shepp (1973). (Later Mason (2005) provided an alternate proof.) Chistyakov and Götze (2004) established the rest of the Logan et al. conjecture by completely characterizing when $\mathbb{T}_n$ converges in distribution to a non-degenerate random variable Y such that $P\{|Y| = 1\} \neq 1$. A bit later, as a by-product of the study of a seemingly unrelated problem, Mason and Zinn (2005) found a simple proof of the full Logan et al. conjecture assuming symmetry.

Theorem 1.1 of Chistyakov and Götze (2004) implies the following: one has $\mathbb{T}_n \xrightarrow{\mathrm{D}} Y$, where $P(|Y| = 1) = 0$, if and only if there exists a sequence of norming constants $b_n > 0$ such that either

$$b_n^{-1} \mathbb{S}_n \xrightarrow{\mathrm{D}} Z \tag{1.2}$$

*Research partially supported by ARC Grant DP1092502
†Research partially supported by NSF Grant DMS–0503908.

or for some $0 < \alpha < 2$,

$$b_n^{-1}\mathbb{S}_n \xrightarrow{\mathrm{D}} U^\alpha, \tag{1.3}$$

where U^α is a strictly stable random variable as defined in the Appendix. In case (1.2), $Y \stackrel{\mathrm{D}}{=} Z$ and in case (1.3), $Y \stackrel{\mathrm{D}}{=} U^\alpha/\sqrt{V^\alpha}$. The V^α random variable arises in the distributional limit in (1.4), which is implied by (1.3):

$$\left(b_n^{-1}\mathbb{S}_n, b_n^{-2}\mathbb{V}_n\right) \xrightarrow{\mathrm{D}} (U^\alpha, V^\alpha)\,. \tag{1.4}$$

For details see the Appendix.

Our aim is to prove analogs of the Chistyakov and Götze (2004) result for a Lévy process X_t, $t \ge 0$, at small times ($t \searrow 0$) and large times ($t \to \infty$). To state our results we must first fix notation. We abbreviate "infinitely divisible" to "inf. div." throughout. Let $(\Omega, \mathcal{F}, P)$ be a probability space carrying a real-valued Lévy process $(X_t)_{t\ge 0}$ having nondegenerate inf. div. characteristic function

$$Ee^{\mathrm{i}\theta X_t} = e^{t\Psi(\theta)}, \quad \theta \in \mathbb{R}, \tag{1.5}$$

where

$$\Psi(\theta) = \mathrm{i}\gamma\theta - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{\mathrm{i}\theta x} - 1 - \mathrm{i}\theta x \mathbf{1}_{\{|x|\le 1\}}\right)\Pi(\mathrm{d}x), \tag{1.6}$$

$\gamma \in \mathbb{R}$, $\sigma^2 \ge 0$, and Π is a measure on $\mathbb{R}\setminus\{0\}$ with $\int_{\mathbb{R}\setminus\{0\}}(x^2 \wedge 1)\Pi(\mathrm{d}x)$ finite.

We say that X_t has *canonical triplet* (γ, σ^2, Π) and X_1 is inf. div. with *triplet* (γ, σ^2, Π). The tails $\overline{\Pi}(x)$ and $\overline{\Pi}^{\pm}(x)$ of Π are defined by

$$\overline{\Pi}^-(x) = \Pi\{(-\infty, -x)\},\ \overline{\Pi}^+(x) = \Pi\{(x, \infty)\},\ \text{and}\ \overline{\Pi}(x) = \overline{\Pi}^+(x) + \overline{\Pi}^-(x), \quad x > 0. \tag{1.7}$$

Assume throughout that $\sigma^2 + \overline{\Pi}(0+) > 0$, otherwise X degenerates to a constant drift.

Let $(\Delta X_t)_{t\ge 0}$, with $\Delta X_t = X_t - X_{t-}$, $X_{0-} = 0$, denote the jump process of X, and consider the Lévy process

$$V_t = \sigma^2 t + \sum_{0<s\le t} (\Delta X_s)^2,\ t > 0. \tag{1.8}$$

V is a subordinator with drift σ^2 and Lévy measure satisfying $\overline{\Pi}_V(x) = \overline{\Pi}(\sqrt{x})$, $x > 0$. By Theorem 2.1 of Maller and Mason (2008) the joint characteristic function of (X_t, V_t) is given by

$$Ee^{\mathrm{i}(\theta_1 X_t + \theta_2 V_t)}$$
$$= \exp\left\{\mathrm{i}t\left(\theta_1\gamma + \theta_2\sigma^2\right) - t\theta_1^2\sigma^2/2 + t\int_{\mathbb{R}\setminus\{0\}}\left(e^{\mathrm{i}(\theta_1 x + \theta_2 x^2)} - 1 - \mathrm{i}\theta_1 x\mathbf{1}_{\{|x|\le 1\}}\right)\Pi(\mathrm{d}x)\right\}.$$

We shall say that (X_t, V_t) has triplet (γ, σ^2, Π).

Here is our small time ($t \searrow 0$) analog of the Chistyakov and Götze (2004) result.

Theorem 1.1 *Let X_t, $t \ge 0$, be a Levy process satisfying $\overline{\Pi}(0+) = \infty$. Assume that*

$$X_t/\sqrt{V_t} \xrightarrow{\mathrm{D}} Y,\ \textit{as}\ t \searrow 0, \tag{1.9}$$

where Y is a finite rv with $P(|Y|=1)=0$. Then either Y is standard normal or

$$Y \stackrel{\mathrm{D}}{=} U^{\alpha}/\sqrt{V^{\alpha}}, \tag{1.10}$$

where (U^{α}, V^{α}) is a strictly stable pair of index α for some $0<\alpha<2$, as in (1.4).

Here is our large time analog ($t \to \infty$) of the Chistyakov and Götze (2004) result.

Theorem 1.2 *Let X_t, $t \geq 0$, be a Levy process satisfying $\sigma^2 + \overline{\Pi}(0+) > 0$. We have*

$$X_t/\sqrt{V_t} \stackrel{\mathrm{D}}{\longrightarrow} Y, \text{ as } t \to \infty, \tag{1.11}$$

where $P(|Y|=1)=0$, if and only if either X_1 has expectation 0 and is in the domain of attraction of a normal law, in which case $Y=Z$, or X_1 is in the domain of attraction of a strictly stable law in the sense that for a sequence of positive norming constants b_n,

$$b_n^{-1}\left\{X_{(1)} + \cdots + X_{(n)}\right\}$$

converges in distribution to a nondegenerate strictly stable law of index $0<\alpha<2$, where $X_{(1)}, X_{(2)}, \ldots,$ are i.i.d. as X_1, in which case Y is as in (1.10).

Remark 1. Chistyakov and Götze (2004) also show that $\mathbb{T}_n$ converges in distribution to a non-degenerate random variable Y such that $P\{|Y|=1\}=1$ if and only if $P\{|\xi|>x\}$ is slowly varying at infinity. We do not have such a complete picture for X_t. Assuming $\overline{\Pi}(0+)=\infty$, the proof of Lemma 5.3 of Maller and Mason (2010) shows that if $\overline{\Pi}$ is slowly varying at zero then $|X_t|/\sqrt{V_t} \stackrel{\mathrm{D}}{\longrightarrow} 1$. However, we do not know whether the converse is true, except in the case when X_t is symmetric. In this case, Maller and Mason (2008) prove that whenever $\overline{\Pi}(0+)=\infty$, $X_t/\sqrt{V_t} \stackrel{\mathrm{D}}{\longrightarrow} Y$, as $t \searrow 0$, where Y is equal to 1 or -1 with probability 1/2 each if and only if $\overline{\Pi}$ is slowly varying at zero. Further results are given in Theorem 3.4 of Maller and Mason (2010). Analogous statements can be said about the case $t \to \infty$.

1.1 Some needed technical results

To prove Theorems 1.1 and 1.2 we shall need to establish a number of technical results about $X_t/\sqrt{V_t}$, which are also of independent interest. To do this we must introduce some more notation and definitions. We will use some truncated mean and variance functions, defined for $x>0$ by

$$\nu(x) = \gamma - \int_{x<|y|\leq 1} y\Pi(\mathrm{d}y),\ V(x) = \sigma^2 + \int_{|y|\leq x} y^2\Pi(\mathrm{d}y), \text{ and } U(x) = \sigma^2 + 2\int_0^x y\overline{\Pi}(y)\mathrm{d}y. \tag{1.12}$$

These functions are finite for all $x>0$ by virtue of the properties of the Lévy measure Π, which further imply that $\lim_{x\searrow 0} x^2\overline{\Pi}(x)=0$, and $\lim_{x\searrow 0} x\nu(x)=0$.

By the *relative compactness* of a real-valued stochastic process $(S_t)_{t\geq 0}$, as $t\to\infty$, we will mean that it satisfies

$$\lim_{x\to\infty}\limsup_{t\to\infty} P(|S_t|>x)=0,$$

or, equivalently, every sequence $t_k \to \infty$ contains a subsequence $t_{k'} \to \infty$ with $S_{t_{k'}}$ converging in distribution to an a.s. finite rv. If in addition each such subsequential limit is not degenerate at a constant, we say that S_t is *stochastically compact*, as $t \to \infty$.

By the *Feller class at 0* we will mean the class of Lévy processes which are stochastically compact at 0 after norming and centering; that is, those for which there are nonstochastic functions $a(t)$, $b(t) > 0$ (where, *throughout, $b(t)$ will be assumed positive, but not, a priori, monotone),* such that every sequence $t_k \searrow 0$ contains a subsequence $t_{k'} \searrow 0$ with

$$\left(X_{t_{k'}} - a(t_{k'})\right)/b(t_{k'}) \xrightarrow{\mathrm{D}} Y', \text{ as } k' \to \infty, \tag{1.13}$$

where Y' is a finite nondegenerate rv, a.s. (The prime on Y' denotes that in general it will depend on the choice of subsequence $t_{k'}$.) We describe this kind of convergence as "$X_t \in FC$ at 0".

It was shown in Maller and Mason (2008) that when the relation (1.13) holds (with Y' not degenerate at a constant) then it must be the case that Y' is an inf. div. rv, and $b(t_{k'}) \to 0$ as $t_{k'} \searrow 0$.

Closely related is the *centered Feller class at 0.* This is the class of Lévy processes which are stochastically compact at 0, after norming, but with no centering function needed; that is, those for which there is a nonstochastic function $b(t) > 0$ such that every sequence $t_k \searrow 0$ contains a subsequence $t_{k'} \searrow 0$ with

$$X_{t_{k'}}/b(t_{k'}) \xrightarrow{\mathrm{D}} Y', \text{ as } k' \to \infty, \tag{1.14}$$

where Y' is a finite, nondegenerate, necessarily inf. div., rv, a.s. We describe this as "$X_t \in FC_0$ at 0".

The classes FC and FC_0 "at infinity" are defined in exactly the same ways, but with the subsequences tending to infinity rather than to 0. Maller and Mason (2009, 2010) have carried out a thorough study of FC and FC_0 at 0 and infinity and have obtained a number of useful analytic equivalences in terms of the Lévy measure Π of X_t.

The following propositions connect the self-normalized and compactness ideas, and will be essential ingredients in the proofs of Theorems 1.1 and 1.2.

Proposition 1.1 *Suppose $X_t/\sqrt{V_t}$ is relatively compact as $t \searrow 0$ and no subsequential limit has positive mass at ± 1. Then $X \in FC_0$ as $t \searrow 0$, or, equivalently, by Theorem 2.3 of Maller and Mason (2010)*

$$\limsup_{x \searrow 0} \frac{x^2\overline{\Pi}(x) + x|\nu(x)|}{V(x)} < \infty. \tag{1.15}$$

Proposition 1.2 *Suppose $X_t/\sqrt{V_t}$ is relatively compact as $t \to \infty$ and no subsequential limit has positive mass at ± 1. Then $X \in FC_0$ as $t \to \infty$, or, equivalently, by Theorem 1 (ii) of Maller and Mason (2009)*

$$\limsup_{x \to \infty} \frac{x^2\overline{\Pi}(x) + x|\nu(x)|}{V(x)} < \infty. \tag{1.16}$$

These two propositions will be proved in a separate section.

2 Proofs of Theorems

The proofs will require the following two limit theorems, which we state as Lemmas 2.1 and 2.2. Recall that $(X_t)_{t\geq 0}$ is Lévy with canonical triplet (γ, σ^2, Π). Suppose that, for a sequence of integers $n_k \to \infty$,

$$X_{n_k}/B(n_k) \xrightarrow{\mathrm{D}} U, \tag{2.1}$$

where U is inf. div. with triplet (b, a, Λ). Notice that necessarily $B(n_k) \to \infty$.
Each random variable $X_{n_k}/B(n_k)$ is inf. div. with triplet $(b_{n_k}, a_{n_k}, \Lambda_{n_k})$, where

$$b_{n_k} = n_k\gamma/B(n_k), \; a_{n_k} = n_k\sigma^2/B^2(n_k) \; \text{ and } \Lambda_{n_k}(\mathrm{d}x) = n_k\Pi(\mathrm{d}x/B(n_k)).$$

Moreover, $\left(X_{n_k}/B(n_k), V_{n_k}/B^2(n_k)\right)$ has joint characteristic function

$$\begin{aligned} &Ee^{\mathrm{i}(\theta_1 X_{n_k}/B(n_k)+\theta_2 V_{n_k}/B^2(n_k))} \\ &= \exp\left\{\mathrm{i}(\theta_1 b_{n_k} + \theta_2 a_{n_k}) - \theta_1^2 a_{n_k}/2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{\mathrm{i}(\theta_1 x+\theta_2 x^2)} - 1 - \mathrm{i}\theta_1 x\mathbf{1}_{\{|x|\leq 1\}}\right)\Lambda_{n_k}(\mathrm{d}x)\right\}. \end{aligned}$$

Lemma 2.1 *Whenever (2.1) holds,*

$$\left(X_{n_k}/B(n_k), V_{n_k}/B^2(n_k)\right) \xrightarrow{\mathrm{D}} (U, W), \tag{2.2}$$

where (U, W) has joint characteristic function

$$\begin{aligned} &Ee^{\mathrm{i}(\theta_1 U+\theta_2 W)} \\ &= \exp\left\{\mathrm{i}(\theta_1 b + \theta_2 a) - \theta_1^2 a/2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{\mathrm{i}(\theta_1 x+\theta_2 x^2)} - 1 - \mathrm{i}\theta_1 x\mathbf{1}_{\{|x|\leq 1\}}\right)\Lambda(\mathrm{d}x)\right\}. \end{aligned} \tag{2.3}$$

Proof. For each $h > 0$ let

$$a^h = a + \int_{0<|x|\leq h} x^2\Lambda(\mathrm{d}x) \; \text{ and } b^h = b - \int_{h<|x|\leq 1} x\Lambda(\mathrm{d}x), \tag{2.4}$$

and let $a^h_{n_k}$ and $b^h_{n_k}$ be defined as in (2.4) with Λ replaced by Λ_{n_k}, a by a_{n_k}, and b by b_{n_k}. According to Theorem 15.14 of Kallenberg (2001), (2.1) happens if and only if

$$\Lambda_{n_k} \text{ converges vaguely to } \Lambda \text{ on } \mathbb{R}\setminus\{0\} \tag{2.5}$$

and for any $h > 0$ such that $\Lambda\{|x| = h\} = 0$,

$$a^h_{n_k} \to a^h \text{ and } b^h_{n_k} \to b^h. \tag{2.6}$$

By the vague convergence of Λ_{n_k} to Λ, and since $a^h_{n_k} \to a^h$, we also have, for any $r > 2$,

$$\int_{0<|x|\leq h} x^r\Lambda_{n_k}(\mathrm{d}x) \to \int_{0<|x|\leq h} x^r\Lambda(\mathrm{d}x). \tag{2.7}$$

To verify (2.7), take $r > 2$ and $0 < \delta < h$, with δ and h continuity points of Λ, and write

$$\int_{0<|x|\leq\delta} x^r \Lambda_{n_k}(dx) \leq \delta^{r-2} \int_{0<|x|\leq\delta} x^2 \Lambda_{n_k}(dx)$$

$$\leq \delta^{r-2} \left(a_{n_k} + \int_{0<|x|\leq h} x^2 \Lambda_{n_k}(dx) \right).$$

Thus

$$\lim_{\delta\searrow 0} \limsup_{k\to\infty} \int_{0<|x|\leq\delta} x^r \Lambda_{n_k}(dx) \leq \lim_{\delta\searrow 0} \delta^{r-2} a^h = 0.$$

We also have by vague convergence of Λ_{n_k} to Λ,

$$\lim_{k\to\infty} \int_{\delta<|x|\leq h} x^r \Lambda_{n_k}(\mathrm{d}x) = \int_{\delta<|x|\leq h} x^r \Lambda(\mathrm{d}x).$$

These two convergences imply that for all $r > 2$, we have (2.7).
Also $a_{n_k}^h \to a^h$ implies that

$$c_{n_k}^h = c_{n_k} - \int_{h<|x|\leq 1} x^2 \Lambda_{n_k}(\mathrm{d}x) \to c - \int_{h<|x|\leq 1} x^2 \Lambda(\mathrm{d}x),$$

where

$$c_{n_k} = a_{n_k} + \int_{0<|x|\leq 1} x^2 \Lambda_{n_k}(\mathrm{d}x) \text{ and } c = a + \int_{0<|x|\leq 1} x^2 \Lambda(\mathrm{d}x).$$

Write $\nu = \Lambda \circ T^{-1}$, where $T(x) = \left(x, x^2\right)$. Now on account of (2.5) we can readily infer that $\Lambda_{n_k} \circ T^{-1}$ converges vaguely to ν on $\overline{\mathbb{R}}^2 \setminus \{(0,0)\}$. Thus by using the bivariate version of Theorem 15.14 of Kallenberg (2001), we get after a little algebra that (2.2) holds with (U, W) having characteristic function

$$\exp\left\{ -\frac{a\theta_1^2}{2} + i\left(b^h\theta_1 + a\theta_2\right) + \int_{\mathbb{R}\setminus\{0\}} \left(\exp\left(i\left(\theta_1 x + \theta_2 x^2\right)\right) - 1 - i\theta_1 x 1\{|x| \leq h\}\right) \Lambda(\mathrm{d}x) \right\}, \tag{2.8}$$

for any $h > 0$ such that $\Lambda\{|x| = h\} = 0$. Note that in applying the bivariate version of Theorem 15.14 of Kallenberg (2001), we get, using $a_{n_k}^h \to a^h$ and (2.7), that

$$\begin{pmatrix} a_{n_k}^h & \int_{0<|x|\leq h} x^3 \Lambda_{n_k}(\mathrm{d}x) \\ \int_{0<|x|\leq h} x^3 \Lambda_{n_k}(\mathrm{d}x) & \int_{0<|x|\leq h} x^4 \Lambda_{n_k}(\mathrm{d}x) \end{pmatrix} \to$$

$$\begin{pmatrix} a^h & \int_{0<|x|\leq h} x^3 \Lambda(\mathrm{d}x) \\ \int_{0<|x|\leq h} x^3 \Lambda(\mathrm{d}x) & \int_{0<|x|\leq h} x^4 \Lambda(\mathrm{d}x) \end{pmatrix}$$

and

$$\begin{pmatrix} b_{n_k}^h \\ a_{n_k} + \int_{0<|x|\leq 1} x^2 \Lambda_{n_k}(\mathrm{d}x) \end{pmatrix} \to \begin{pmatrix} b^h \\ a + \int_{0<|x|\leq 1} x^2 \Lambda(\mathrm{d}x) \end{pmatrix}.$$

We see then that the resulting limiting infinitely divisible vector has in its defining characteristic function the Lévy measure $\nu = \Lambda \circ T^{-1}$ on $\overline{\mathbb{R}}^2 \setminus \{(0,0)\}$ and the matrix and constant vector, respectively,

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} b \\ a + \int_{0<|x|\le 1} x^2 \Lambda\left(\mathrm{d}x\right) \end{pmatrix}.$$

For very similar details see the proof of Lemma 4 of Giné and Mason (1998). Since

$$b^h = b - \int_{h<|x|\le 1} x\Lambda\left(\mathrm{d}x\right),$$

we get that the characteristic function (2.8) is equal to (2.3). □

As above, let $\xi_1, \xi_2, \ldots,$ be i.i.d. nondegenerate random variables with cumulative distribution function F, and for each integer $n \ge 1$ denote the sums $\mathbb{S}_n = \sum_{i=1}^n \xi_i$ and $\mathbb{V}_n = \sum_{i=1}^n \xi_i^2$. Suppose that there exist a subsequence $\{n_k\} \subset \{n\}$ and norming constants $B\left(n_k\right)$ such that

$$\mathbb{S}_{n_k}/B\left(n_k\right) \xrightarrow{\mathrm{D}} U, \tag{2.9}$$

where U is an inf. div. random variable with triplet (b, a, Λ).

Lemma 2.2 *Whenever (2.9) holds,*

$$\left(\mathbb{S}_{n_k}/B\left(n_k\right), \mathbb{V}_{n_k}/B^2\left(n_k\right)\right) \xrightarrow{\mathrm{D}} \left(U, W\right), \tag{2.10}$$

where (U, W) *has joint characteristic function (2.3).*

Proof. By Corollary 15.16 of Kallenberg (2001), (2.9) occurs if and only if

$$n_k\mathcal{L}\left(\xi/B\left(n_k\right)\right) \text{ converges vaguely to } \Lambda \text{ on } \mathbb{R}\setminus\{0\} \tag{2.11}$$

and for any $h > 0$ such that $\Lambda\{|x| = h\} = 0$

$$n_k E\left[\left(\xi/B\left(n_k\right)\right)^2 1\left\{\left|\xi/B\left(n_k\right)\right| \le h\right\}\right] \to a^h \tag{2.12}$$

and

$$n_k E\left[\left(\xi/B\left(n_k\right)\right) 1\left\{\left|\xi/B\left(n_k\right)\right| \le h\right\}\right] \to b^h.$$

Also, as in the proof of Lemma 2.1, for every $r > 2$,

$$n_k E\left[\left(\xi/B\left(n_k\right)\right)^r 1\left\{\left|\xi/B\left(n_k\right)\right| \le h\right\}\right] \to \int_{0<|x|\le h} x^r \Lambda\left(\mathrm{d}x\right),$$

and similarly as in the proof of Lemma 2.1, $n_k\mathcal{L}\left(\xi/B\left(n_k\right), \xi^2/B^2\left(n_k\right)\right)$ converges vaguely to ν on $\overline{\mathbb{R}}^2 \setminus \{(0,0)\}$. □

Remark 2. In the sequel we shall only need the special case of Lemma 2.2 when $\xi_1, \xi_2, \ldots,$ are i.i.d $\xi = X_1$, where X_1 is inf. div. with canonical triplet (γ, σ^2, Π).

2.1 Proof of Theorem 1.2

It is more efficient to prove Theorem 1.2 first. By Proposition 1.2, whenever

$$X_t/\sqrt{V_t} \xrightarrow{\mathrm{D}} Y, \text{ as } t \to \infty,$$

where Y does not put positive mass ± 1, then X_t is centered stochastically compact at infinity with a norming function b_t. This implies by Lemmas 2.1 and 2.2 that if $\xi_1, \ldots, \xi_m$ are i.i.d. X_1, there exists a positive norming b_m such that

$$\left(X_m/b_m, V_m/b_m^2\right) \text{ and } \left(\mathbb{S}_m/b_m, \mathbb{V}_m/b_m^2\right)$$

have the same nondegenerate subsequential distributional limits. Thus both

$$X_m/\sqrt{V_m} \text{ and } \mathbb{S}_m/\sqrt{\mathbb{V}_m}$$

converge in distribution to the same nondegenerate random variable Y that does not put positive mass on ± 1. The proof of Theorem 1.2 now follows from the Chistyakov and Götze (2004) result. □

2.2 Proof of Theorem 1.1

Assume that (1.9) holds, where $P\{|Y| = 1\} = 0$. We know by Proposition 1.1 that this forces X_t to be centered stochastically compact at 0. Thus there exists a norming function a_t such that every subsequence t_k converging to zero contains a further subsequence s_n with

$$\left(X_{ts_n}/a_{s_n}, V_{ts_n}/a_{s_n}^2\right)_{t\geq 0} \xrightarrow{\mathrm{D}} (U_t, W_t)_{t\geq 0}, \text{ as } n \to \infty, \tag{2.13}$$

where the Lévy process (U, W), which may depend on the subsequence s_n, has joint characteristic function

$$Ee^{\mathrm{i}(\theta_1 U_t + \theta_2 W_t)}$$
$$= \exp\left\{ \mathrm{i}t\left(\theta_1 a + \theta_2 b\right) - t\theta_1^2 b/2 + t\int_{\mathbb{R}\setminus\{0\}} \left(e^{\mathrm{i}(\theta_1 x + \theta_2 x^2)} - 1 - \mathrm{i}\theta_1 x \mathbf{1}_{\{|x|\leq 1\}}\right) \Lambda(\mathrm{d}x) \right\},$$

with $a \in \mathbb{R}$, $b > 0$ and Λ being a Lévy measure on $\mathbb{R} \setminus \{0\}$. See Maller and Mason (2010) Theorem 2.3 for the functional convergence in (2.13).

We first claim that if U contains a normal component, i.e. $b \neq 0$, this forces Y to be standard normal. This will be a consequence of the following lemma.

Lemma 2.3 *If U contains a normal component ($a \neq 0$), then one can find a subsequence t' $\searrow 0$ such that*

$$\left(X_{t'}/a_{t'}, V_{t'}/a_{t'}^2\right) \xrightarrow{\mathrm{D}} (aZ,\ a), \tag{2.14}$$

where Z is standard normal.

Proof. Using (2.13), we see that, for each fixed $m > 1$,

$$\left(X_{s_n/m}/\left(a_{s_n}/\sqrt{m}\right), V_{s_n/m}/\left(a_{s_n}^2/m\right)\right) \xrightarrow{\mathrm{D}} (\sqrt{m}U_{1/m}, mW_{1/m}), \text{ as } n \to \infty,$$

where $(\sqrt{m}U_{1/m}, mW_{1/m})$ has characteristic function

$$= \exp\left\{\mathrm{i}\left(\frac{\theta_1 b}{\sqrt{m}} + \theta_2 a\right) - \theta_1^2 a/2 + \frac{1}{m}\int_{\mathbb{R}\setminus\{0\}}\left(e^{\mathrm{i}(\sqrt{m}\theta_1 x + m\theta_2 x^2)} - 1 - \mathrm{i}\sqrt{m}\theta_1 x \mathbf{1}_{\{|x|\le 1\}}\right)\Lambda(\mathrm{d}x)\right\},$$

which as we will show converges as $m \to \infty$ to $\exp\left\{-\theta_1^2 a/2 + \mathrm{i}\theta_2 a\right\}$. Thus with some abuse of notation we can extract a sequence s_{n_k}/m_k converging to 0 so that as $k \to \infty$,

$$\left(X_{s_{n_k}/m_k}/\left(a_{s_{n_k}}/\sqrt{m_k}\right), V_{s_{n_k}/m_k}/\left(a_{s_{n_k}}^2/m_k\right)\right) \xrightarrow{\mathrm{D}} (aZ,\ a), \tag{2.15}$$

having characteristic function $\exp\left\{-\theta_1^2 a/2 + \mathrm{i}\theta_2 a\right\}$. Actually to show (2.15) it remains to prove that

$$\lim_{m\to\infty}\frac{1}{m}\int_{\mathbb{R}\setminus\{0\}}\left(e^{\mathrm{i}(\sqrt{m}\theta_1 x + m\theta_2 x^2)} - 1 - \mathrm{i}\sqrt{m}\theta_1 x \mathbf{1}_{\{|x|\le 1\}}\right)\Lambda(\mathrm{d}x) = 0. \tag{2.16}$$

To see why (2.16) is true notice that

$$\limsup_{m\to\infty}\frac{1}{m}\left|\int_{1/\sqrt{m}\le|x|<\infty}\left(e^{\mathrm{i}(\sqrt{m}\theta_1 x + m\theta_2 x^2)} - 1\right)\Lambda(\mathrm{d}x)\right| \le \limsup_{m\to\infty}\frac{2}{m}\Lambda\left(1/\sqrt{m} \le |x| < \infty\right) = 0.$$

Also for all $0 < \delta < 1$

$$\begin{aligned}
\limsup_{m\to\infty}\frac{1}{m}\left|\int_{1/\sqrt{m}\le|x|\le 1}\mathrm{i}\sqrt{m}x\mathbf{1}_{\{|x|\le 1\}}\Lambda(\mathrm{d}x)\right| &\le \limsup_{m\to\infty}\frac{1}{\sqrt{m}}\int_{1/\sqrt{m}\le|x|\le\delta}|x|\,\mathbf{1}_{\{|x|\le 1\}}\Lambda(\mathrm{d}x) \\
&\quad + \limsup_{m\to\infty}\frac{1}{\sqrt{m}}\left|\int_{\delta\le|x|\le 1}x\mathbf{1}_{\{|x|\le 1\}}\Lambda(\mathrm{d}x)\right| \\
&\le \limsup_{m\to\infty}\int_{1/\sqrt{m}\le|x|\le\delta}x^2\mathbf{1}_{\{|x|\le 1\}}\Lambda(\mathrm{d}x) \\
&= \int_{0<|x|\le\delta}x^2\mathbf{1}_{\{|x|\le 1\}}\Lambda(\mathrm{d}x).
\end{aligned}$$

Since $\delta > 0$ can be made arbitrarily small we get

$$\lim_{m\to\infty}\frac{1}{m}\left|\int_{1/\sqrt{m}\le|x|\le 1}\mathrm{i}\sqrt{m}x\mathbf{1}_{\{|x|\le 1\}}\Lambda(\mathrm{d}x)\right| = 0. \tag{2.17}$$

Thus to complete the proof of (2.16) it suffices to show that

$$\lim_{m\to\infty}\frac{1}{m}\int_{0<|x|\le 1/\sqrt{m}}\left(e^{\mathrm{i}(\sqrt{m}\theta_1 x + m\theta_2 x^2)} - 1 - \mathrm{i}\sqrt{m}\theta_1 x \mathbf{1}_{\{|x|\le 1\}}\right)\Lambda(\mathrm{d}x) = 0. \tag{2.18}$$

Now the LHS of (2.18) does not exceed

$$\frac{1}{m}\int_{0<|x|\le 1/\sqrt{m}}\left|e^{\mathrm{i}(\sqrt{m}\theta_1 x + m\theta_2 x^2)} - 1 - \mathrm{i}\left(\sqrt{m}\theta_1 x\mathbf{1}_{\{|x|\le 1\}} + m\theta_2 x^2\right)\right|\Lambda(\mathrm{d}x)$$

$$+\int_{0<|x|\le 1/\sqrt{m}} \theta_2 x^2 \Lambda(\mathrm{d}x),$$

and for some $C>0$

$$\frac{1}{m}\int_{0<|x|\le 1/\sqrt{m}} \left| e^{\mathrm{i}(\sqrt{m}\theta_1 x + m\theta_2 x^2)} - 1 - \mathrm{i}\left(\sqrt{m}\theta_1 x \mathbf{1}_{\{|x|\le 1\}} + m\theta_2 x^2\right)\right| \Lambda(\mathrm{d}x)$$

$$\le \frac{C}{m}\int_{0<|x|\le 1/\sqrt{m}} \left|\sqrt{m}\theta_1 x + m\theta_2 x^2\right|^2 \Lambda(\mathrm{d}x),$$

which for some $D>0$ depending on θ_1 and θ_2 is

$$\le D\int_{0<|x|\le 1/\sqrt{m}} x^2 \Lambda(\mathrm{d}x).$$

Since the limit of this as $m\to\infty$ is 0, we have shown (2.18), which together with (2.2) and (2.17) gives (2.16). □

Hence if there exists a nonzero normal component in the characteristic function of U in (2.13) for convergence of $(X_t/a_t, V_t/a_t^2)$ along a subsequence s_n, then Y must be standard normal. Since in this case for some subsequence t' we have $X_{t'}/\sqrt{V_{t'}} \stackrel{\mathrm{D}}{\longrightarrow} Z$, as $t' \searrow 0$ we get $Y=Z$ in (1.9). From now on we shall assume that Y is not standard normal, which means that $a=0$ in the characteristic function of the (U,W) appearing in (2.13) for convergence of $(X_t/a_t, V_t/a_t^2)$ along a subsequence s_n. Also note that (U,W) may be different for different subsequences. However, in all cases, by assumption (1.9),

$$U_t/\sqrt{W_t} \stackrel{\mathrm{D}}{=} Y, \text{ for each } t>0.$$

Note that Theorem 2.1 (iii) of Maller and Mason (2010) implies that $P\{W>0\}=1$. Moreover, it can be shown that for any integer $m\ge 1$,

$$\frac{U_m}{\sqrt{W_m}} \stackrel{\mathrm{D}}{=} \frac{U_{(1)}+\cdots+U_{(m)}}{\sqrt{W_{(1)}+\cdots+W_{(m)}}} \stackrel{\mathrm{D}}{=} Y, \tag{2.19}$$

where $\left(U_{(1)},W_{(1)}\right), \ldots, \left(U_{(m)},W_{(m)}\right)$ are i.i.d. (U_1,W_1). To see this, observe that for any fixed integer $m\ge 1$,

$$\left(X_{ms_n}/a_{s_n}, V_{ms_n}/a_{s_n}^2\right) \stackrel{\mathrm{D}}{\longrightarrow} \left(U_{(1)}+\cdots+U_{(m)}, W_{(1)}+\cdots+W_{(m)}\right) \stackrel{\mathrm{D}}{=} \left(U_m, V_m\right).$$

We claim that

$$U_t/\sqrt{W_t} \stackrel{\mathrm{D}}{\longrightarrow} Y, \text{ as } t\to\infty. \tag{2.20}$$

This follows from (2.19) combined with the facts that

$$W_{m+1}-W_m \stackrel{\mathrm{D}}{=} W_{(1)} = O_p(1),\ W_{(1)}+\cdots+W_{(m+1)} \stackrel{\mathrm{P}}{\longrightarrow} \infty$$

and

$$\sup_{m<t\le m+1} |U_t-U_m| \stackrel{\mathrm{D}}{=} \sup_{0<t\le 1} |U_t| = O_p(1),$$

which together imply that

$$\sup_{m<t\le m+1}\left|U_t/\sqrt{W_t}-U_m/\sqrt{W_m}\right| \xrightarrow{\mathrm{P}} 0.$$

Therefore we can apply Theorem 1.2 combined with the fact that Y is not standard normal to conclude that U is in the domain of attraction of a strictly stable law of index $0<\alpha<2$, and thus $Y \stackrel{\mathrm{D}}{=} U^\alpha/\sqrt{V^\alpha}$. □

Remark 3. We do not have such a complete picture of the distributional limits of $X_t/\sqrt{V_t}$ as $t \searrow 0$ as Chistyakov and Götze (2004) obtained for self-normalized sums in their Theorem 1.1. All we can say is that if

$$X_t/\sqrt{V_t} \xrightarrow{\mathrm{D}} Y, \text{ as } t \searrow 0,$$

where Y does not place positive mass on any constant then either $Y \stackrel{\mathrm{D}}{=} Z$ or $Y \stackrel{\mathrm{D}}{=} U^\alpha/\sqrt{V^\alpha}$. Only in the case when $Y \stackrel{\mathrm{D}}{=} Z$ do we know that this happens if and only if for some norming function b_t,

$$X_t/b_t \xrightarrow{\mathrm{D}} Z, \text{ as } t \searrow 0.$$

This was proved in Theorem 2.4 of Maller and Mason (2010). The story for the case $Y \stackrel{\mathrm{D}}{=} U^\alpha/\sqrt{V^\alpha}$ is not complete. Presently all that we can infer is that if for some $0<\alpha<2$,

$$X_t/b_t \xrightarrow{\mathrm{D}} U^\alpha, \text{ as } t \searrow 0,$$

then

$$X_t/\sqrt{V_t} \xrightarrow{\mathrm{D}} U^\alpha/\sqrt{V^\alpha}, \text{ as } t \searrow 0.$$

However, right now, we cannot go the other way, except under the assumption of symmetry. See Maller and Mason (2008).

3 Proofs of Propositions 1.1 and 1.2 .

3.1 Proof of Proposition 1.1

Suppose $X_t/\sqrt{V_t}$ is relatively compact as $t\downarrow 0$ and no subsequential limit has positive mass at ±1. Then by Theorem 3.1 of Maller and Mason (2010) we have

$$\limsup_{x\downarrow 0}\frac{x|\nu(x)|}{U(x)} = \limsup_{x\downarrow 0}\frac{x|\nu(x)|}{x^2\overline{\Pi}(x)+V(x)} < \infty, \tag{3.1}$$

while by Proposition 5.5 of Maller and Mason (2010) we have

$$\limsup_{x\downarrow 0}\frac{x^2\overline{\Pi}(x)}{V(x)} < \infty, \tag{3.2}$$

since if (3.2) fails then there is a subsequential limit rv of $X_t/\sqrt{V_t}$ (as $t\downarrow 0$) with positive mass at ±1. Now (3.1) and (3.2) imply

$$\limsup_{x\downarrow 0}\frac{x|\nu(x)|}{V(x)} < \infty$$

which together with (3.2) gives (1.15). □

3.2 Proof of Proposition 1.2

The proof of Proposition 1.2 will be a consequence of the following two propositions and theorem, which are the large time analogs of their small time versions given in Propositions 5.1 and 5.5 and Theorem 3.1 of Maller and Mason (2010).

Recall the definition of $U(x)$ given in (1.12). Note that, after integrating by parts,

$$U(x) = V(x) + x^2\overline{\Pi}(x), \;\; x > 0. \tag{3.3}$$

The function $U(x)$ is continuous, in fact, differentiable, at each $x > 0$, with

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{U(x)}{x^2}\right) = \frac{-2\left(U(x) - x^2\overline{\Pi}(x)\right)}{x^3}.$$

Further,

$$U(x) - x^2\overline{\Pi}(x) = \sigma^2 + 2\int_0^x y\left(\overline{\Pi}(y) - \overline{\Pi}(x)\right)\mathrm{d}y \geq 2\int_0^x y\left(\overline{\Pi}(y) - \overline{\Pi}(x)\right)\mathrm{d}y.$$

The right-hand side here could be 0 only if $\overline{\Pi}(y)$ is constant on $(0, x]$, and since $\overline{\Pi}(\infty) = 0$, as long as $\overline{\Pi}(x) > 0$ for some $x > 0$, we see that $x^{-2}U(x)$ is strictly decreasing for large enough x, and $x^{-2}U(x) \to \infty\mathbf{1}_{\{\sigma^2>0\}} + \overline{\Pi}(0+)\mathbf{1}_{\{\sigma^2=0\}} > 0$ as $x \searrow 0$, while $x^{-2}U(x) \searrow 0$ as $x \nearrow \infty$.

In view of the monotonicity of $x^{-2}U(x)$ just established, for each $\lambda > 0$, once t is large enough, depending on λ, for $x^{-2}U(x) < \infty\mathbf{1}_{\{\sigma^2>0\}} + \overline{\Pi}(0+)\mathbf{1}_{\{\sigma^2=0\}}$, the function

$$b_\lambda(t) := \inf\{x > 0 : x^{-2}U(x) \leq (\lambda t)^{-1}\}$$

is finite, positive, is such that $b_\lambda(t) \to \infty$ as $t \to \infty$, and is such that

$$\frac{tU(b_\lambda(t))}{b_\lambda^2(t)} = \frac{1}{\lambda}. \tag{3.4}$$

Further, $x^{-2}U(x)$ has no intervals of constancy once x is large enough, because of its strict monotonicity, so $b_\lambda(t)$ is continuous and strictly increasing for each $\lambda > 0$, for large enough t.

In the sequel we shall often use the following decomposition, or a variant of it (see Sato (1999), Theorem 19.2 p.120, or Doney and Maller (2002), Eq. (6.1)):

$$X_t = t\nu(b) + \sigma Z_t + X_t^{(S,b)} + X_t^{(B,b)}, \;\; t \geq 0, \tag{3.5}$$

where Z_t is a standard Brownian motion, $X_t^{(S,b)}$ is the compensated sum of "small" jumps, i.e.

$$X_t^{(S,b)} = \text{a.s.}\lim_{\varepsilon\downarrow 0}\left(\sum_{0<s\leq t}\Delta X_s 1_{\{\varepsilon<|\Delta X_s|\leq b\}} - t\int_{\varepsilon<|x|\leq b} x\Pi(\mathrm{d}x)\right), \;\; t \geq 0,$$

and $X_t^{(B,b)}$ is the "big" jumps, i.e.,

$$X_t^{(B,b)} = \sum_{0<s\leq t}\Delta X_s 1_{\{|\Delta X_s|>b\}}, \;\; t \geq 0.$$

Further, the processes $(Z_t)_{t\geq 0}$, $(X_t^{(S,b)})_{t\geq 0}$ and $(X_t^{(B,b)})_{t\geq 0}$ are all independent.

Theorem 3.1 *We have that*

$$\frac{X_t}{\sqrt{V_t}} \text{ is relatively compact as } t \to \infty \text{ if and only if } \limsup_{x\to\infty} \frac{x|\nu(x)|}{U(x)} < \infty. \tag{3.6}$$

We will deduce Theorem 3.1 from the following analogue of Theorem 2 of Griffin (2002) and Corollary 3.1 is immediate from it.

Proposition 3.1 *There is a nonstochastic function $a(t)$ such that*

$$(X_t - a(t))/\sqrt{V_t} \text{ is relatively compact as } t \to \infty \tag{3.7}$$

if and only if

$$\limsup_{t\to\infty} \frac{|t\nu(b_\lambda(t)) - a(t)|}{b_\lambda(t)} < \infty, \tag{3.8}$$

for all small, and hence, all, $\lambda > 0$.

Corollary 3.1 (Corollary to Proposition 3.1) *(i) $(X_t - t\nu(b_\lambda(t))/\sqrt{V_t}$ is always relatively compact as $t \to \infty$, for any $\lambda > 0$.*

(ii) If X_t is symmetric, then $X_t/\sqrt{V_t}$ is always relatively compact as $t \to \infty$.

Proof of Proposition 3.1: (i) First suppose $EX_1^2 < \infty$. From (3.4) we see that $b_\lambda(t) \asymp \sqrt{t}$ as $t \to \infty$. The convergences

$$\frac{X_t - tEX_1}{\sqrt{t\text{Var}X_1}} \xrightarrow{\text{D}} N(0,1)$$

and

$$V_t/t \xrightarrow{\text{P}} EX_1^2,$$

as $t \to \infty$ can be found in Bertoin (1996), Sato (1999) or see Doney and Maller (2002). So (3.7) holds with $a(t) = tEX_1$, and with this choice,

$$\begin{aligned} \frac{t\nu(b_\lambda(t)) - a(t)}{b_\lambda(t)} &= \frac{t\int_{|y|>b_\lambda(t)} y\Pi(\text{d}y)}{b_\lambda(t)} \\ &\le \frac{t\int_{|y|>0} y^2\Pi(\text{d}y)}{b_\lambda^2(t)} = O(1). \end{aligned}$$

So Proposition 3.1 is true when $EX_1^2 < \infty$.

(ii) Next suppose $EX_1^2 = \infty$. In particular, this means $\overline{\Pi}(x) > 0$ for all $x > 0$. Fix $\lambda > 0$ and then take $t > 0$ big enough, depending on λ, for $b_\lambda(t) > 1$.

From (3.5) with $b = 1$ we have, for $t > 0$,

$$\begin{aligned} X_t &= t\nu(1) + \sigma Z_t + X_t^{(S,1)} + X_t^{(B,1)} \\ &= t\gamma + X_t^{(B,1)} + O_P(\sqrt{t}), \text{ as } t \to \infty, \end{aligned}$$

because $X_t^{(S,1)}$ is a mean 0, finite variance Lévy process. Also

$$\begin{aligned} V_t &= \sigma^2 t + \sum_{0<s\le t} (\Delta X_s)^2 \mathbf{1}_{\{|\Delta X_s|\le 1\}} + \sum_{0<s\le t} (\Delta X_s)^2 \mathbf{1}_{\{|\Delta X_s|>1\}} \\ &=: \sigma^2 t + V_t^{(S,1)} + V_t^{(B,1)} \\ &= V_t^{(B,1)} + O_P(t), \text{ as } t \to \infty. \end{aligned}$$

This is true because $V_t^{(S,1)}$ is a Lévy process with finite mean, so $V_t^{(S,1)}/t = O_P(1)$ as $t \to \infty$ by the weak law of large numbers. But $V_t^{(B,1)}/t \xrightarrow{\mathrm{P}} \infty$ as $t \to \infty$ since $EX_1^2 = \infty$, so $V_t/V_t^{(B,1)} \xrightarrow{\mathrm{P}} 1$ as $t \to \infty$, and thus from

$$\frac{X_t^{(B,1)} + t\gamma - a(t)}{\sqrt{V_t^{(B,1)}}} = \frac{X_t - a(t)}{\sqrt{V_t}} \sqrt{\frac{V_t}{V_t^{(B,1)}}} + O_P\left(\sqrt{\frac{t}{V_t^{(B,1)}}}\right)$$
$$= \frac{X_t - a(t)}{\sqrt{V_t}} (1 + o_P(1)) + o_P(1), \text{ as } t \to \infty, \tag{3.9}$$

we see that (3.7) holds if and only if

$$\left(X_t^{(B,1)} - \widetilde{a}(t)\right) / \sqrt{V_t^{(B,1)}} \text{ is relatively compact as } t \to \infty,$$

for some $\widetilde{a}(t) = a(t) - t\gamma$. So we can ignore small jumps in X and assume X is compound Poisson with no drift, no Brownian component, and all jumps exceeding 1 in magnitude.

Thus in (1.12) we take $\gamma = \sigma^2 = 0$, and can write

$$X_t = \sum_{i=1}^{N_t} J_i, \tag{3.10}$$

and

$$V_t = \sum_{0<s\le t} (\Delta X_s)^2 \mathbf{1}_{\{|\Delta X_s|>1\}} = \sum_{i=1}^{N_t} J_i^2, \tag{3.11}$$

for $(J_i)_{i=1,2,\dots}$ i.i.d. and distributed as $\Pi(\mathrm{d}x)\mathbf{1}_{\{|x|>1\}}/\overline{\Pi}(1)$, and $(N_t)_{t\ge 0}$ independently distributed as a Poisson process with rate $\overline{\Pi}(1)$.

We decompose X_t as

$$X_t = T_t(\lambda) + R_t(\lambda), \tag{3.12}$$

where

$$T_t(\lambda) := \sum_{i=1}^{N_t} J_i \mathbf{1}_{\{|J_i|\le b_\lambda(t)\}} = \sum_{i=1}^{N_t} J_i \mathbf{1}_{\{1<|J_i|\le b_\lambda(t)\}},$$

and

$$R_t(\lambda) := \sum_{i=1}^{N_t} J_i \mathbf{1}_{\{|J_i|> b_\lambda(t)\}}. \tag{3.13}$$

Then we can calculate

$$E(T_t(\lambda)) = t\overline{\Pi}(1) \int_{1<|x|\le b_\lambda(t)} x\Pi(\mathrm{d}x)/\overline{\Pi}(1) = t\nu(b_\lambda(t)),$$

and

$$\mathrm{Var}(T_t(\lambda)) = t \int_{1<|x|\le b_\lambda(t)} x^2\Pi(\mathrm{d}x) \le tU(b_\lambda(t)).$$

We can thus write, for any $L>0$,

$$\begin{aligned} & P\left(|t\nu(b_\lambda(t)) - a(t)| > 3L^2 b_\lambda(t)\right) \\ & \le P\left(|T_t(\lambda) - ET_t(\lambda)| > L^2 b_\lambda(t)\right) + P\left(|T_t(\lambda) - a(t)| > 2L^2 b_\lambda(t)\right), \end{aligned} \tag{3.14}$$

and we proceed by estimating the quantities on the right-hand side of (3.14).
By Chebyshev's inequality, for any $L>0$, $K>0$,

$$\begin{aligned} P\left(|T_t(\lambda) - ET_t(\lambda)| > LKb_\lambda(t)\right) & \le \frac{\mathrm{Var}(T_t(\lambda))}{L^2K^2b_\lambda^2(t)} \\ & \le \frac{tU(b_\lambda(t))}{L^2K^2b_\lambda^2(t)} = \frac{1}{L^2K^2\lambda}. \end{aligned} \tag{3.15}$$

With $K=L$ this gives a bound for the first term on the right-hand side of (3.14). The second term on the right-hand side of (3.14) does not exceed

$$P\left(|T_t(\lambda) - a(t)| > 2L^2 b_\lambda(t), Lb_\lambda(t) \ge \sqrt{V_t}\right) + P\left(Lb_\lambda(t) < \sqrt{V_t}\right). \tag{3.16}$$

Let

$$U_t(\lambda) := \sum_{i=1}^{N_t} J_i^2 \wedge b_\lambda^2(t). \tag{3.17}$$

It's not hard to check that

$$E\left(U_t(\lambda)\right) = \lambda^{-1} b_\lambda^2(t) - tV(1). \tag{3.18}$$

On the event $\{\max_{1\le i\le N_t}|J_i| \le b_\lambda(t)\}$ we have $V_t = U_t(\lambda)$, so

$$\begin{aligned} P\left(\sqrt{V_t} > Lb_\lambda(t)\right) & \le P\left(V_t > L^2 b_\lambda^2(t),\ \max_{1\le i\le N_t}|J_i| \le b_\lambda(t)\right) \\ & \qquad + 1 - P\left(\max_{1\le i\le N_t}|J_i| \le b_\lambda(t)\right) \\ & \le P\left(U_t(\lambda) > L^2 b_\lambda^2(t)\right) + 1 - \sum_{n\ge 0} P^n(|J_1| \le b_\lambda(t))P(N_t = n) \\ & \le \frac{E(U_t(\lambda))}{L^2 b_\lambda^2(t)} + 1 - \sum_{n\ge 0} e^{-t\overline{\Pi}(1)}(t\overline{\Pi}(1)P(|J_1| \le b_\lambda(t)))^n/n! \\ & = \frac{\left(\lambda^{-1}b_\lambda^2(t) - tV(1)\right)}{L^2 b_\lambda^2(t)} + 1 - e^{-t\overline{\Pi}(1)P(|J_1|>b_\lambda(t))} \\ & \le \lambda^{-1}L^{-2} + 1 - e^{-t\overline{\Pi}(b_\lambda(t))} \le \lambda^{-1}L^{-2} + 1 - e^{-\lambda^{-1}}. \end{aligned} \tag{3.19}$$

(In the last inequality, recall that $x^2\overline{\Pi}(x) \le U(x)$, so $t\overline{\Pi}(b_\lambda(t)) \le 1/\lambda$.) (3.19) gives a bound for the second term in (3.16).

Next, to estimate $R_t(\lambda)$ in (3.13), put

$$S_t(\lambda) := \sum_{i=1}^{N_t} \mathbf{1}_{\{|J_i|>b_\lambda(t)\}}.$$

Then by Cauchy-Schwarz,

$$|R_t(\lambda)|^2 \le \left(\sum_{i=1}^{N_t} J_i^2\right)\left(\sum_{i=1}^{N_t} \mathbf{1}_{\{|J_i|>b_\lambda(t)\}}\right) = V_t S_t(\lambda).$$

Thus, for $L > 0$,

$$\begin{aligned} P\left(|R_t(\lambda)| > L\sqrt{V_t}\right) &\le P(S_t(\lambda) > L^2) \le L^{-2}E(S_t(\lambda)) \\ &= L^{-2}(t\overline{\Pi}(1))\overline{\Pi}(b_\lambda(t))/\overline{\Pi}(1) \le L^{-2}\lambda^{-1}. \end{aligned} \tag{3.20}$$

So (cf. (3.12)) we see that the first term in (3.16) does not exceed

$$\begin{aligned} &P\left(|T_t(\lambda) - a(t)| > 2L\sqrt{V_t}\right) \\ &\le P\left(|X_t - a(t)| > L\sqrt{V_t}\right) + P(|R_t(\lambda)| > L\sqrt{V_t}) \\ &\le P\left(|X_t - a(t)| > L\sqrt{V_t}\right) + L^{-2}\lambda^{-1}, \end{aligned} \tag{3.21}$$

by (3.20). Going back to (3.14), we put together (3.16) with $K = L$, and the bounds in (3.19) and (3.21), to deduce that

$$\begin{aligned} &P\left(|t\nu(b_\lambda(t)) - a(t)| > 3L^2 b_\lambda(t)\right) = \mathbf{1}_{\{|t\nu(b_\lambda(t))-a(t)|>3L^2b_\lambda(t)\}} \\ &\le \lambda^{-1}L^{-4} + \lambda^{-1}L^{-2} + 1 - e^{-\lambda^{-1}} + P\left(|X_t - a(t)| > L\sqrt{V_t}\right) + L^{-2}\lambda^{-1}. \end{aligned}$$

Choose L so large that $\lambda^{-1}L^{-4} + 2L^{-2}\lambda^{-1} < e^{-\lambda^{-1}}/2$. Then

$$\mathbf{1}_{\{|t\nu(b_\lambda(t))-a(t)|>3L^2b_\lambda(t)\}} \le P(|X_t - a(t)| > L\sqrt{V_t}) + 1 - e^{-\lambda^{-1}}/2.$$

Now assume (3.7), i.e., that $|X_t - a(t)|/\sqrt{V_t}$ is relatively compact as $t \to \infty$. Letting $t \to \infty$ then $L \to \infty$ gives

$$\lim_{L\to\infty} \limsup_{t\to\infty} \mathbf{1}_{\{|t\nu(b_\lambda(t))-a(t)|>3L^2b_\lambda(t)\}} \le 1 - e^{-\lambda^{-1}}/2 < 1.$$

So, for $L \ge$ some $L_0(\lambda) > 0$, and $t \ge$ some $t_0(L, \lambda) > 0$, we have $\mathbf{1}_{\{|t\nu(b_\lambda(t))-a(t)|>3L^2b_\lambda(t)\}} < 1$, hence $|t\nu(b_\lambda(t)) - a(t)| \le 3L^2 b_\lambda(t)$. Thus for $t \ge t_0(L_0, \lambda)$,

$$\frac{|t\nu(b_\lambda(t)) - a(t)|}{b_\lambda(t)} \le 3L_0^2,$$

which implies (3.8).

For the converse, suppose (3.8) holds for all sufficiently small $\lambda > 0$. Fix such a $\lambda \in (0, 1)$. Therefore by (3.8) we can find a $t_\lambda > 0$ such that

$$c(\lambda) := \sup_{t\ge t_\lambda} \frac{|t\nu(b_\lambda(t)) - a(t)|}{b_\lambda(t)} < \infty. \tag{3.22}$$

Then for all $t \geq t_\lambda$ and any $L > 0$

$$
\begin{aligned}
P\left(|X_t - a(t)| > 3L\sqrt{V_t}\right) &\leq P\left(|X_t - t\nu(b_\lambda(t))| > 3L\sqrt{V_t} - c(\lambda)b_\lambda(t)\right) \\
&\leq P\left(|X_t - t\nu(b_\lambda(t))| > 2L\sqrt{V_t}\right) + P\left(L\sqrt{V_t} \leq c(\lambda)b_\lambda(t)\right). \qquad (3.23)
\end{aligned}
$$

To deal with the first term on the right-hand side, take $K \in (0, \lambda^{-1/2})$, and suppose $t > 0$ is so large that $tV(1) < (\lambda^{-1} - K^2)b_\lambda^2(t)$. This is possible since $b_\lambda^2(t)/t = \lambda U(b_\lambda(t)) \to \infty$ as $t \to \infty$. Recall that $X_t = T_t(\lambda) + R_t(\lambda)$ by (3.12), where $ET_t(\lambda) = t\nu(b_\lambda(t))$, and argue as follows:

$$
\begin{aligned}
P\left(|X_t - t\nu(b_\lambda(t))| > 2L\sqrt{V_t}\right) &\leq P\left(|T_t(\lambda) - t\nu(b_\lambda(t))| > L\sqrt{V_t}\right) \\
&\qquad + P\left(|R_t(\lambda)| > L\sqrt{V_t}\right) \\
&\leq P\left(|T_t(\lambda) - t\nu(b_\lambda(t))| > LKb_\lambda(t)\right) + P\left(\sqrt{V_t} \leq Kb_\lambda(t)\right) \\
&\qquad + P\left(|R_t(\lambda)| > L\sqrt{V_t}\right) \\
&\leq \frac{1}{L^2K^2\lambda} + P\left(\sqrt{V_t} \leq Kb_\lambda(t)\right) + \frac{1}{\lambda L^2} \quad \text{(by (3.15) and (3.20))}. \qquad (3.24)
\end{aligned}
$$

Recall (3.17) and note that $V_t \geq U_t(\lambda)$, so for $t > 0$

$$
P\left(\sqrt{V_t} > Kb_\lambda(t)\right) \geq P\left(U_t(\lambda) > K^2 b_\lambda^2(t)\right).
$$

Using a second moment version of Wald's lemma we see that

$$
\begin{aligned}
\mathrm{Var}\left(U_t\left(\lambda\right)\right) &\leq t\overline{\Pi}(1)E\left(J_1^4 \wedge b_\lambda^4\left(t\right)\right) \\
&= tb_\lambda^4\left(t\right)\overline{\Pi}\left(b_\lambda\left(t\right)\right) + t\int_{1 \leq |y| \leq b_\lambda(t)} y^4 \Pi\left(dy\right) \\
&\leq tb_\lambda^2\left(t\right)\left(b_\lambda^2\left(t\right)\overline{\Pi}\left(b_\lambda\left(t\right)\right) + \int_{1 \leq |y| \leq b_\lambda(t)} y^2 \Pi\left(dy\right)\right) \\
&\leq tb_\lambda^2\left(t\right)U\left(b_\lambda\left(t\right)\right) = \lambda^{-1}b_\lambda^4\left(t\right).
\end{aligned}
$$

Recall that we keep $K^2 < \lambda^{-1}$ and $tV(1) < (\lambda^{-1} - K^2)b_\lambda^2(t)$. There is a one-sided Chebyshev inequality of the form $P(Y - EY < x) \geq x^2/(x^2 + \mathrm{Var}Y)$, for any rv Y and $x > 0$ (e.g., Beichelt (2006), p.70). Apply this with $Y = -U_t(\lambda)$, recalling that $E(U_t(\lambda)) = \lambda^{-1}b_\lambda^2(t) - tV(1)$, by (3.18), to get

$$
\begin{aligned}
P(\sqrt{V_t} > Kb_\lambda(t)) &\geq P\left(U_t(\lambda) > K^2 b_\lambda^2(t)\right) \\
&= P\left(-U_t(\lambda) + E(U_t(\lambda)) < (\lambda^{-1} - K^2)b_\lambda^2(t) - tV(1)\right) \\
&\geq \frac{\left((\lambda^{-1} - K^2)b_\lambda^2(t) - tV(1)\right)^2}{\left((\lambda^{-1} - K^2)b_\lambda^2(t) - tV(1)\right)^2 + \mathrm{Var}(U_t(\lambda))} \\
&\geq \frac{(1 - K^2\lambda)^2 b_\lambda^4(t) - 2t\lambda(1 - K^2\lambda)b_\lambda^2(t)V(1)}{b_\lambda^4(t)(1 + \lambda) + t^2\lambda^2V^2(1)}. \qquad (3.25)
\end{aligned}
$$

For the second term on the right-hand side of (3.23), use (3.25) with K replaced by $K_\lambda := c(\lambda)L^{-1}$, with $L > c(\lambda)\sqrt{\lambda}$, so $K_\lambda < \lambda^{-1/2}$. Thus, finally,

$$\begin{aligned}
&P(|X_t - a(t)| > 3L\sqrt{V_t}) \\
&\le \frac{1}{\lambda L^2 K^2} + \left[1 - \frac{(1-K^2\lambda)^2 b_\lambda^4(t) - 2t\lambda(1-K^2\lambda)b_\lambda^2(t)V(1)}{b_\lambda^4(t)(1+\lambda) + t^2\lambda^2V^2(1)}\right] \\
&\quad + \left[1 - \frac{(1-K_\lambda^2\lambda)^2 b_\lambda^4(t) - 2t\lambda(1-K_\lambda^2\lambda)b_\lambda^2(t)V(1)}{b_\lambda^4(t)(1+\lambda) + t^2\lambda^2V^2(1)}\right] + \frac{1}{\lambda L^2}.
\end{aligned}$$

Let $t \to \infty$, recalling that $t = o(b_\lambda^2(t))$, then $L \to \infty$, noting that $K_\lambda = c(\lambda)L^{-1} \to 0$, then let $K \downarrow 0$, to see that

$$\lim_{L\to\infty} \limsup_{t\to\infty} P(|X_t - a\,(t)\,| > 3L\sqrt{V_t}) \le \frac{2\lambda}{1+\lambda}.$$

Then let $\lambda \downarrow 0$ to get (3.7). □

Proof of Theorem 3.1: Suppose $X_t/\sqrt{V_t}$ is relatively compact but there is a sequence $x_k \to \infty$ such that

$$\frac{x_k|\nu(x_k)|}{U(x_k)} \to \infty, \text{ as } k \to \infty.$$

Let $t_k = x_k^2/U(x_k)$, so $x_k = b_1(t_k)$, in the notation of (3.4). Then

$$\frac{t_k|\nu(b_1(t_k))|}{b_1(t_k)} = \frac{t_k|\nu(x_k)|}{x_k} = \frac{x_k|\nu(x_k)|}{U(x_k)} \to \infty,$$

which contradicts (3.8) with $a(t) = 0$ and $\lambda = 1$.

Conversely, suppose $\limsup_{x\to\infty} x|\nu(x)|/U(x) < c < \infty$. Then with $a(t) \equiv 0$ we have

$$\limsup_{t\to\infty} \frac{|a(t) - t\nu(b_\lambda(t))|}{b_\lambda(t)} = \limsup_{t\to\infty} \frac{b_\lambda(t)|\nu(b_\lambda(t))|}{\lambda U(b_\lambda(t))} \le \frac{c}{\lambda},$$

so (3.8) holds with $a(t) = 0$, and $X_t/\sqrt{V_t}$ is relatively compact as $t \to \infty$, by Proposition 3.1. □

Proposition 3.2 *Suppose $T_t := X_t/\sqrt{V_t}$ is relatively compact as $t \to \infty$, and also that*

$$\limsup_{x\to\infty} x^2\overline{\Pi}(x)/V(x) = \infty.$$

Then there is a sequence $t_k \to \infty$ such that

$$\lim_{\delta\downarrow 0} \limsup_{t_k\to\infty} P(||T_{t_k}| - 1| \le \delta) > 0. \tag{3.26}$$

Proof of Proposition 3.2: Assume that T_t is relatively compact as $t \to \infty$ and let

$$R(x) := x^2\overline{\Pi}(x)/V(x). \tag{3.27}$$

Suppose $\limsup_{x\to\infty} R(x) = \infty$. This implies that $EX_1^2 = \infty$, thus by (3.9) again we need only deal with the big jump process. So we take X and V as in (3.10) and (3.11), and set $\gamma = \sigma^2 = 0$ in (1.12).

We first show the existence of a sequence $\zeta_k \to \infty$ such that

$$\lim_{k\to\infty} \inf_{0<\lambda_1\le\lambda\le\lambda_2} R(\lambda\zeta_k) = \infty, \text{ for each } 0 < \lambda_1 < 1 < \lambda_2 < \infty. \tag{3.28}$$

To this end, fix $0 < \lambda_1 < 1 < \lambda_2 < \infty$, choose $c_n \uparrow \infty$ such that $R(c_n) \uparrow \infty$, let $n_1 = \min\{m \ge 1 : c_m > 2, \text{ and } R(c_m) > 2^3\}$, and then for $k = 1, 2, \ldots$, set

$$n_{k+1} = \min\{m > n_k : c_m/2^{k+1} > 2^{-k}c_{n_k} + k + 1, \text{ and } R(c_m) > 2^{3(k+1)}\}.$$

Then put $\zeta_k = 2^{-k}c_{n_k}$, so that $\zeta_k \to \infty$ as $k \uparrow \infty$. Note that $R(x)/x^2$ is nonincreasing on $(0,\infty)$. Choose $\lambda \in [1, \lambda_2]$ and k such that $2^k \ge \lambda_2$. Then $\lambda \le 2^k$ and

$$\frac{R(\lambda\zeta_k)}{\lambda^2} = \frac{\zeta_k^2 R(\lambda\zeta_k)}{(\lambda\zeta_k)^2} \ge \frac{\zeta_k^2 R(2^k\zeta_k)}{(2^k\zeta_k)^2} = 2^{-2k}R(2^k\zeta_k) = 2^{-2k}R(c_{n_k}) \ge 2^k,$$

so $\inf_{1\le\lambda\le\lambda_2} R(\lambda\zeta_k) \ge 2^k \to \infty$. Thus $R(\zeta_k) \ge 2^k \to \infty$, and for $\lambda \in [\lambda_1, 1]$, $\lambda_1^2 R(\zeta_k) \le \lambda^2 R(\zeta_k) \le R(\lambda\zeta_k)$, so $\inf_{\lambda_1\le\lambda\le 1} R(\lambda\zeta_k) \to \infty$. Hence (3.28) holds.

Recall that $U(x) \ge x^2\overline{\Pi}(x)$ for all $x > 0$, so for $\lambda > 0$,

$$0 \le 1 - \frac{(\lambda\zeta_k)^2\overline{\Pi}(\lambda\zeta_k)}{U(\lambda\zeta_k)} = \frac{V(\lambda\zeta_k)}{U(\lambda\zeta_k)} \le \frac{V(\lambda\zeta_k)}{(\lambda\zeta_k)^2\overline{\Pi}(\lambda\zeta_k)} = \frac{1}{R(\lambda\zeta_k)} \to 0, \tag{3.29}$$

uniformly in $\lambda \in [\lambda_1, \lambda_2]$, where $0 < \lambda_1 < 1 < \lambda_2 < \infty$. Now

$$\int_{\lambda_1}^{\lambda_2} \frac{(s\zeta_k)^2\overline{\Pi}(s\zeta_k)}{U(s\zeta_k)}\frac{\mathrm{d}s}{s} = \frac{1}{2}\int_{\lambda_1\zeta_k}^{\lambda_2\zeta_k} \frac{\mathrm{d}U(s)}{U(s)} = \frac{1}{2}\log\left(\frac{U(\lambda_2\zeta_k)}{U(\lambda_1\zeta_k)}\right) \tag{3.30}$$

(recall that $U(x)$ is continuous at each $x > 0$). The left-hand side of the last expression tends to $\int_{\lambda_1}^{\lambda_2} \mathrm{d}s/s = \log(\lambda_2/\lambda_1)$, so we have $U(\lambda_2\zeta_k)/U(\lambda_1\zeta_k) \to (\lambda_2/\lambda_1)^2$. Then by (3.29), $\overline{\Pi}(\lambda_2\zeta_k)/\overline{\Pi}(\lambda_1\zeta_k) \to 1$, and so we deduce

$$\lim_{k\to\infty} \sup_{0<\lambda_1\le\lambda\le\lambda_2} \left|\frac{\overline{\Pi}(\lambda\zeta_k)}{\overline{\Pi}(\zeta_k)} - 1\right| = 0, \text{ for each } 0 < \lambda_1 < \lambda_2. \tag{3.31}$$

Recall the representation of X_t in (3.10) and the definition of the J_i. Let $J_{N_t}^{(1)}$ be any J_i which is largest in modulus among $J_1, \ldots, J_{N_t}$, and let $J_{N_t}^{(2)}$ denote any term among $J_1, \ldots, J_{N_t}$ of second largest modulus.

Define

$$^{(1)}\widetilde{X}_t := X_t - J_{N_t}^{(1)} = \sum_{i=1}^{N_t} J_i - J_{N_t}^{(1)}$$

and

$$^{(1)}V_t := V_t - |J_{N_t}^{(1)}|^2.$$

Put $t_k := 1/\overline{\Pi}(\zeta_k)$, so that $t_k \to \infty$. For $\delta > 0$ and $0 < \lambda_1 < \lambda_2$, define the events

$$A_k := \left\{ |J^{(2)}_{N_{t_k}}| \le \lambda_1\zeta_k < \lambda_2\zeta_k < |J^{(1)}_{N_{t_k}}| \right\},$$

$$B_k(\delta) := \left\{ |{}^{(1)}\widetilde{X}_{t_k}| > \delta |J^{(1)}_{N_{t_k}}| \right\},$$

and

$$C_k(\delta) := \left\{ {}^{(1)}V_{t_k} > \delta^2 |J^{(1)}_{N_{t_k}}|^2 \right\}.$$

In the following, we will keep $\lambda_1\zeta_k > 1$. A straightforward calculation gives

$$P(A_k) = t_k\overline{\Pi}(\lambda_2\zeta_k)e^{-t_k\overline{\Pi}(\lambda_1\zeta_k)} =: \rho_k(\lambda_1, \lambda_2), \text{ say.} \tag{3.32}$$

Also, on A_k, we have

$$B_k(\delta) \subseteq \left\{ \left| \sum_{i=1}^{N_{t_k}} J_i \mathbf{1}_{\{|J_i| \le \lambda_1\zeta_k\}} \right| > \delta\lambda_2\zeta_k \right\} = \left\{ \left| \sum_{i=1}^{N_{t_k}} J_i^k \right| > \delta\lambda_2\zeta_k \right\}, \tag{3.33}$$

where $J_i^k := J_i \mathbf{1}_{\{|J_i| \le \lambda_1\zeta_k\}}$. Note that

$$E(J_1^k) = \int_{1<|x|\le\lambda_1\zeta_k} x\Pi(\mathrm{d}x)/\overline{\Pi}(1),$$

and $E(N_{t_k}) = t_k\overline{\Pi}(1)$, so we can write

$$\sum_{i=1}^{N_{t_k}} J_i^k = \sum_{i=1}^{N_{t_k}} \left(J_i^k - E(J_1^k) \right) + (N_{t_k} - E(N_{t_k}))\, E(J_1^k) + t_k \int_{1<|x|\le\lambda_1\zeta_k} x\Pi(\mathrm{d}x). \tag{3.34}$$

Now since T_t is assumed relatively compact, we have by (3.6)

$$x|\nu(x)| \le M(x^2\overline{\Pi}(x) + V(x)),$$

for $x \ge$ some x_0, for some $M \in (0, \infty)$. Further, by (3.29), we have $V(\lambda\zeta_k) \le (\lambda\zeta_k)^2\overline{\Pi}(\lambda\zeta_k)$, for k large, uniformly in $\lambda \in [\lambda_1, \lambda_2]$. Thus for k large, firstly,

$$\begin{aligned} \left| t_k \int_{1<|x|\le\lambda_1\zeta_k} x\Pi(\mathrm{d}x) \right| &= t_k|\nu(\lambda_1\zeta_k)| \text{ (recall } \gamma = 0 \text{ in (1.12))} \\ &\le 2Mt_k(\lambda_1\zeta_k)\overline{\Pi}(\lambda_1\zeta_k) \\ &\le 4M\lambda_1\zeta_k \quad (\text{by (3.31), and } t_k = 1/\overline{\Pi}(\zeta_k)) \\ &\le (\delta/2)\lambda_2\zeta_k, \end{aligned}$$

for $\lambda_2 > \lambda_1$ large enough. Secondly,

$$\begin{aligned} \mathrm{Var}\left(\sum_{i=1}^{N_{t_k}} \left(J_i^k - E(J_1^k) \right) \right) &= E(N_{t_k})\mathrm{Var}(J_1^k) \\ &\le t_k\overline{\Pi}(1) \int_{1<|x|\le\lambda_1\zeta_k} x^2\Pi(\mathrm{d}x)/\overline{\Pi}(1) \\ &\le t_k V(\lambda_1\zeta_k). \end{aligned}$$

Third, using Cauchy-Schwarz,

$$\begin{aligned}\mathrm{Var}\left[(N_{t_k}-E(N_{t_k}))\,E(J_1^k)\right] &= t_k\overline{\Pi}(1)\left(\int_{1<|x|\le\lambda_1\zeta_k} x\Pi(\mathrm{d}x)/\overline{\Pi}(1)\right)^2\\ &\le t_k\int_{1<|x|\le\lambda_1\zeta_k} x^2\Pi(\mathrm{d}x)\\ &\le t_kV(\lambda_1\zeta_k).\end{aligned}$$

Putting the three estimates into (3.34) and using Chebyshev's inequality, we find that for $\delta_1>0$ and $\lambda_1>\lambda_2$ large enough

$$\begin{aligned}&P(B_k(\delta_1)\cap A_k)\\ &\le P\left(\left|\sum_{i=1}^{N_{t_k}}\left(J_i^k-E(J_1^k)\right)+(N_{t_k}-E(N_{t_k}))\,E(J_1^k)\right|>(\delta_1/2)\lambda_2\zeta_k\right)\\ &\le\frac{8t_kV(\lambda_1\zeta_k)}{(\delta_1\lambda_2\zeta_k)^2}.\end{aligned}$$

By a similar argument as in (3.33) and Markov's inequality we get for $\delta_2>0$

$$\begin{aligned}P(C_k(\delta_2)\cap A_k) &\le P\left(\sum_{i=1}^{N_{t_k}}|J_i^k|^2>(\delta_2\lambda_2\zeta_k)^2\right)\\ &\le\frac{E(N_{t_k})E(J_1^k)^2}{(\delta_2\lambda_2\zeta_k)^2}\le\frac{t_kV(\lambda_1\zeta_k)}{(\delta_2\lambda_2\zeta_k)^2}.\end{aligned}$$

Putting these together gives

$$\begin{aligned}P\left(\{B_k(\delta_2)\cap A_k\}\cup\{C_k(\delta_1)\cap A_k\}\right) &\le\left(\frac{1}{\delta_1^2}+\frac{1}{\delta_2^2}\right)\frac{8t_kV(\lambda_1\zeta_k)}{(\lambda_2\zeta_k)^2}\\ &=:\eta_k(\delta_1,\delta_2),\ \text{say}.\end{aligned}\tag{3.35}$$

Now, since $V_t={}^{(1)}V_t+|J_{N_t}^{(1)}|^2\ge|J_{N_t}^{(1)}|^2$, we can write, for $\delta>0$,

$$\begin{aligned}&P\left(\left|\frac{|X_{t_k}|}{\sqrt{V_{t_k}}}-1\right|>\delta\right)\\ &=P\left(\left||X_{t_k}|-\sqrt{V_{t_k}}\right|>\delta\sqrt{V_{t_k}}\right)\\ &\le P\Big(\left\{\left||X_{t_k}|-\sqrt{V_{t_k}}\right|>\delta|J_{N_{t_k}}^{(1)}|,\ {}^{(1)}V_{t_k}\le(\delta/2)^2|J_{N_{t_k}}^{(1)}|^2\right\}\\ &\qquad\cup\left\{{}^{(1)}V_{t_k}>(\delta/2)^2|J_{N_{t_k}}^{(1)}|^2\right\}\Big).\end{aligned}$$

The latter does not exceed

$$P\left(\left\{\left|X_{t_k}-J_{N_{t_k}}^{(1)}\right|>\delta|J_{N_{t_k}}^{(1)}|/2\right\}\cup\left\{{}^{(1)}V_{t_k}>(\delta/2)^2|J_{N_t}^{(1)}|^2\right\}\right);\tag{3.36}$$

because, $\sqrt{{}^{(1)}V_{t_k}} \le (\delta/2)|J^{(1)}_{N_{t_k}}|$, thus $|\sqrt{V_{t_k}} - |J^{(1)}_{N_{t_k}}|| \le (\delta/2)|J^{(1)}_{N_{t_k}}|$, together with

$$\left||X_{t_k}| - \sqrt{V_{t_k}}\right| > \delta|J^{(1)}_{N_{t_k}}| > \left|\sqrt{V_{t_k}} - |J^{(1)}_{N_{t_k}}|\right|,$$

imply

$$\begin{aligned}
&\left|X_{t_k} - J^{(1)}_{N_{t_k}}\right| \\
&\ge \left||X_{t_k}| - |J^{(1)}_{N_{t_k}}|\right| \ge \left|\left||X_{t_k}| - \sqrt{V_{t_k}}\right| - \left|\sqrt{V_{t_k}} - |J^{(1)}_{N_{t_k}}|\right|\right| \\
&= \left||X_{t_k}| - \sqrt{V_{t_k}}\right| - \left|\sqrt{V_{t_k}} - |J^{(1)}_{N_{t_k}}|\right| \\
&\ge (\delta - \delta/2)|J^{(1)}_{N_{t_k}}| = \delta|J^{(1)}_{N_t}|/2.
\end{aligned}$$

Observe that (3.36) does not exceed $P(B_k(\delta/2) \cup C_k(\delta/2))$. Argue that, by (3.35) and (3.32),

$$\begin{aligned}
P(B_k(\delta_1) \cup C_k(\delta_2)) &\le P\left(\{B_k(\delta_1) \cap A_k\} \cup \{C_k(\delta_2)) \cap A_k\}\right) \\
&\quad + 1 - P(A_k) \\
&\le \eta_k(\delta_1, \delta_2) + 1 - \rho_k(\lambda_1, \lambda_2).
\end{aligned}$$

Thus by (3.36)

$$P\left(\left|\frac{|X_{t_k}|}{\sqrt{V_{t_k}}} - 1\right| > \delta\right) \le \eta_k(\delta/2, \delta/2) + 1 - \rho_k(\lambda_1, \lambda_2). \tag{3.37}$$

Now by (3.29) and (3.31)

$$t_k V(\lambda_1 \zeta_k) = o\left(t_k \zeta_k^2 \overline{\Pi}(\lambda_1 \zeta_k)\right) = o(\zeta_k^2),$$

so $\eta_k(\delta_1, \delta_2) \to 0$ as $k \to \infty$, while, by (3.31), $\rho_k(\lambda_1, \lambda_2) \to e^{-1}$ as $k \to \infty$. Letting $k \to \infty$ in (3.37) gives

$$\limsup_{t_k \to \infty} P\left(\left|\frac{|X_{t_k}|}{\sqrt{V_{t_k}}} - 1\right| > \delta\right) \le 1 - e^{-1} < 1,$$

so (3.26) holds. □

We are now ready to complete the proof of Proposition 1.2. Suppose $X_t/\sqrt{V_t}$ is relatively compact as $t \to \infty$ and no subsequential limit has positive mass at ± 1. Then by Theorem 3.1 we have

$$\limsup_{x \to \infty} \frac{x|\nu(x)|}{U(x)} = \limsup_{x \to \infty} \frac{x|\nu(x)|}{x^2\overline{\Pi}(x) + V(x)} < \infty, \tag{3.38}$$

while by Proposition 3.2 we have

$$\limsup_{x \to \infty} \frac{x^2\overline{\Pi}(x)}{V(x)} < \infty, \tag{3.39}$$

since if (3.39) fails then there is a subsequential limit rv of $X_t/\sqrt{V_t}$ with positive mass at ± 1. Now (3.38) and (3.39) imply

$$\limsup_{x \to \infty} \frac{x|\nu(x)|}{V(x)} < \infty, \tag{3.40}$$

which together with (3.39) gives (1.16). □

3.2.1 Comments on proofs of Propositions 3.1 and 3.2

The proofs of Propositions 3.1 and 3.2 parallel very closely, after notational changes, those of Propositions 5.1 and 5.5 of Maller and Mason (2010), which are their small time versions. Therefore for the sake of brevity, but at the sacrifice of readability, we could have replaced the foregoing proofs by the following road maps:

For the proof of Proposition 3.1 proceed exactly as it is given above until right before equation (3.14), and then continue on as in the proof of Proposition 5.1 of Maller and Mason (2010) starting at its equation (5.8) with the role of ε suppressed, i.e. $T_t(\varepsilon,\lambda)$, $R_t(\varepsilon,\lambda), S_t(\varepsilon,\lambda)$, $N_t(\varepsilon,\lambda)$, $U_t(\varepsilon,\lambda)$, $V_t(\varepsilon), V(\varepsilon)$ and $X_t(\varepsilon)$, are replaced by $T_t(\lambda)$, $R_t(\lambda)$, $S_t(\lambda)$, $N_t(\lambda)$, $U_t(\lambda)$, $V_t, V(1)$ and X_t, respectively. Also replace $t\nu(\varepsilon)$ by 0 and $\alpha_t(\varepsilon,\lambda)$ by $T_t(\lambda)$, and use the definition of $c(\lambda)$ given in (3.22).

From equation (3.29) the proof of Proposition 3.2 goes exactly as that of Proposition 5.5 of Maller and Mason (2010) beginning from its equation (5.38) with the role of ε suppressed in the notation analogously as it was done in the proof of Proposition 1.1 and with $t_k \searrow 0$ changed to $t_k \to \infty$, $t_k\nu(t_k)$ to 0 and $\varepsilon < \lambda_1\xi_k$ to $1 < \lambda_1\xi_k$

3.3 Appendix: Strictly stable bivariate laws

Theorem 1.1 of Chistyakov and Götze (2004) says that $\mathbb{T}_n \xrightarrow{\mathrm{D}} Y$, where $P\{|Y|=1\}=0$, if and only there exists a sequence of norming constants b_n such that either $b_n^{-1}S_n \xrightarrow{\mathrm{D}} Z$ or ξ is in the domain of attraction of a stable law of index $0<\alpha<2$. Moreover, in the normal case $E\xi=0$, in the case $1<\alpha<2$, $E\xi=0$ and in the case $\alpha=1$, ξ is in the domain of attraction of Cauchy's law and Feller's condition holds, that is,

$$\lim_{n\to\infty} E\sin(\xi/b_n) \text{ exists and is finite.}$$

This means that in the stable law of index $0<\alpha<2$ case, one can use the fact that necessarily, for some function L slowly varying at infinity,

$$1-F(x) := P\{|\xi|>x\} = x^{-\alpha}L(x)\,, \text{ for } x>0,$$

and some $0\le p\le 1$, as $x\to\infty$,

$$P\{\xi>x\}/P\{|\xi|>x\}\to p \text{ and } P\{\xi<-x\}/P\{|\xi|>x\}\to q.$$

By applying Theorem 15.14 of Kallenberg (2001) much as we did in the proofs of Lemma 2.1 and 2.2, we can show that for a suitable $c>0$ with a norming sequence b_n of the form

$$b_n \sim cF^{-1}(1/n) = cn^{1/\alpha}L^*(1/n)\,,$$

with L^* slowly varying at zero, one has $\left(b_n^{-1}S_n, b_n^{-2}V_n\right) \xrightarrow{\mathrm{D}} (U^\alpha, V^\alpha)$, where (U^α, V^α) has joint characteristic function

$$\varphi_\alpha(s,t) = E\exp(isU^\alpha + itV^\alpha)\,,$$

of the following form: For $r\ge 0$ and $l\ge 0$ with at least one strictly positive define

$$K_{r,l}(y) = \begin{cases} r, & y>0 \\ 0, & y=0 \\ l, & y<0 \end{cases}$$

where $p = r/(r+l)$.
Case 1: $0 < \alpha < 1$

$$\varphi_\alpha(s,t) = \exp\left(\int_{-\infty}^{\infty} \left(\exp\left(isy + ity^2\right) - 1\right) \frac{K_{r,l}(y)}{|y|^{1+\alpha}} dy\right).$$

Case 2: $\alpha = 1$

$$\varphi_1(s,t) = \exp\left(\int_{-\infty}^{\infty} \left(\exp\left(ity^2\right)\cos(sy) - 1\right) \frac{1}{y^2} dy\right).$$

Case 3: $1 < \alpha < 2$

$$\varphi_\alpha(s,t) = \exp\left(\int_{-\infty}^{\infty} \left(\exp\left(isy + ity^2\right) - 1 - isy\right) \frac{K_{r,l}(y)}{|y|^{1+\alpha}} dy\right).$$

An easy calculation verifies that for all $n \geq 1$ and $0 < \alpha < 2$

$$\varphi_\alpha^n\left(s/n^{1/\alpha}, t/n^{2/\alpha}\right) = \varphi_\alpha(s,t).$$

Thus if $(U_1^\alpha, V_1^\alpha), \ldots, (U_n^\alpha, V_n^\alpha)$ are i.i.d. (U^α, V^α) then for all $n \geq 1$

$$\left(n^{-1/\alpha}\sum_{i=1}^n U_i^\alpha, n^{-2/\alpha}\sum_{i=1}^n V_i^\alpha\right) \stackrel{\mathrm{D}}{=} (U^\alpha, V^\alpha).$$

This says that (U^α, V^α) is a *strictly bivariate stable random vector* and in the stable $0 < \alpha < 2$ case $S_n/\sqrt{V_n} \xrightarrow{\mathrm{D}} Y$, where $Y \stackrel{\mathrm{D}}{=} U^\alpha/\sqrt{V^\alpha}$.

References

[1] Beichelt, F. (2006) *Stochastic Processes in Science, Engineering and Finance* CRC Press LLC, USA.

[2] Bertoin, J. (1996) *Lévy Processes.* Cambridge University Press, Cambridge.

[3] Chistyakov, G.P. and Götze, F. (2004) Limit distributions of studentized sums. *Ann. Probab.,* **32**, 28–77.

[4] Doney, R.A. and Maller, R.A. (2002) Stability and attraction to Normality for Lévy processes at zero and infinity. *J. Theoretical Probab.*, **15**, 751-792.

[5] Giné, E., Götze, F., and Mason, D.M. (1997) When is the student t-statistic asymptotically standard normal? *Ann. Prob.* **25**, 1514-1531.

[6] Giné, E. and Mason, D.M. (1998) On the LIL for Self-Normalized Sums of IID Random Variables, *J. Theoretical Probab.*, **11**, 351-370.

[7] Griffin, P.S. (2002) Tightness of the Student T-statistic, *Elect. Comm. in Probab.*, **7**, 181-190.

[8] Kallenberg, O. (2001) *Foundations of Modern Probability*, 2nd Ed., Springer.

[9] Logan, B.F., Mallows, C.L., Rice, S.O. and Shepp, L. (1973) Limit distributions of self-normalized sums. *Ann. Probab.* **1**, 788–809.

[10] Maller, R.A. and Mason, D.M. (2008) Convergence in distribution of Lévy processes at small times with self-normalisation, *Acta Sci. Math. (Szeged),* **74,** 315–347.

[11] Maller, R.A. and Mason, D.M. (2009) Stochastic compactness of Lévy processes, Proceedings of High Dimensional Probability V, Luminy, France, 2008 (C. Houdré, V. Kolthchinskii, M. Peligrad and D. Mason, eds.) *I.M.S. Collections, High Dimensional Probability V: The Luminy Volume,* **Vol. 5** , 239-257.

[12] Maller, R.A. and Mason, D.M. (2010) Small-time compactness and convergence behavior of deterministically and self-normalised Lévy processes, *Trans. Amer. Math. Soc.,* **362**, 2205-2248.

[13] Mason, D.M. (2005) The asymptotic distribution of self-normalized triangular arrays. J. *Theor. Probab.*, **18**, 853-870.

[14] Mason, D.M. and Zinn, J. (2005) When does a randomly weighted self-normalized sum converge in distribution? *Electronic Comm. Prob.*, **10**, 70–81.

[15] Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

Centre for Mathematical Analysis
& School of Finance and Applied Statistics
Australian National University
PO Box
Canberra ACT
Australia
Ross.Maller@anu.edu.au

Statistics Program
University of Delaware
206 Townsend Hall
Newark, DE 19717
USA
davidm@Udel.Edu