J. Appl. Prob. **54,** 873–889 (2017)
*doi:*10.1017/jpr.2017.39

FUNCTIONAL LAWS FOR TRIMMED LÉVY PROCESSES

BORIS BUCHMANN,* ** *Australian National University*
YUGUANG F. IPSEN,* *** *Australian National University and University of Melbourne*
ROSS MALLER,* **** *Australian National University*

Abstract

Two different ways of trimming the sample path of a stochastic process in $\mathbb{D}[0, 1]$: global ('trim as you go') trimming and record time ('lookback') trimming are analysed to find conditions for the corresponding operators to be continuous with respect to the (strong) J_1-topology. A key condition is that there should be no ties among the largest ordered jumps of the limit process. As an application of the theory, via the continuous mapping theorem, we prove limit theorems for trimmed Lévy processes, using the functional convergence of the underlying process to a stable process. The results are applied to a reinsurance ruin time problem.

Keywords: Functional law; Skorokhod J_1-topology; trimming càdlàg function; trimmed Lévy process; reinsurance risk process

2010 Mathematics Subject Classification: Primary 60G51; 60G52; 60G55
Secondary 60G70; 60F17; 91B30

1. Introduction

By 'trimming' a process we mean identifying 'large' jumps of the process, in some sense, and deleting them from it. The term has its origins in the statistical practice of identifying 'outliers' in a sample of independent and identically distributed (i.i.d.) random variables, then removing them from a statistic of interest, typically, the sample sum, which can be considered as a stochastic process in discrete time. More recently, the techniques have been transferred to processes such as extremal processes and Lévy processes indexed by a continuous-time parameter, where asymptotic properties of the trimmed process were solved for a number of interesting cases. The asymptotic studied may be large time ($t \to \infty$), as in the statistical situation, or, for continuous-time processes, small time ($t \downarrow 0$). The small time case extends our understanding of local properties of the process and can have direct application as, for example, in Maller and Fan [8] and Maller and Schmidli [10]; the large time case has the statistical applications alluded to, such as the robustness of statistics, and insurance modelling, and so on, as we discuss later.

This area of research can be regarded as combining studies on properties of extremes of the jumps of a process with those of the process itself; the former, a version of extreme value

Received 23 September 2016; revision received 22 March 2017.
* Postal address: Research School of Finance, Actuarial Studies and Statistics, Australian National University, 26C Kingsley Street, Acton, ACT 2601, Australia.
** Email address: boris.buchmann@anu.edu.au
*** Email address: yuguang.ipsen@anu.edu.au
**** Email address: ross.maller@anu.edu.au

theory; the latter relating, for example, to domains of attraction of the process. A combination of the two fields enriches both, and the trimming idea is a natural way of approaching this. Research in a similar direction has been carried out by Silvestrov and Teugels [11]; see also their references.

In this paper we extend some earlier work of the present authors to consider various ways of trimming the sample path of a stochastic process in the space $\mathbb{D}[0, 1]$ of càdlàg functions. The initial setup is very general. We begin in Section 2 by establishing continuity properties (in the Skorokhod (strong) J_1-topology) of operators which remove extremes. There are a number of intuitively reasonable ways of defining such operators. Not all of them behave in the same way, and Section 2 is devoted to teasing out the differences between them. We take a dynamic *sample path approach* which brings into focus some interesting and distinctive features not previously apparent. Proofs for Section 2 are in Section 4.

An application of the ideas to the functional convergence of a Lévy process in the domain of attraction of a stable law is then given. Statements for these are in Section 3, and their proofs in Section 5. Continuity properties of certain extremal operators are closely related to the occurrence, or otherwise, of tied (equal) values in the large jumps of the limiting process and consequently we need to analyse these too. Finally, in Section 6 we develop a motivating application to a reinsurance ruin time problem.

2. Extremal operators on Skorokhod space

Let $\mathbb{D}([0, 1], \mathbb{R}) =: \mathbb{D}$ be the space of all càdlàg functions: $[0, 1] \to \mathbb{R}$ endowed with the Skorokhod (strong) J_1-topology. Denote the sup norm by $\|\cdot\|$, so that $\|x\| = \sup_{0\le\tau\le1} |x(\tau)|$, where, for each $x \in \mathbb{D}$, $x(\tau)$ is the value of x at time $\tau \in [0, 1]$. Convergence in the J_1-topology is characterised as follows. Let Λ be the set of all continuous and strictly increasing functions $\lambda\colon [0, 1] \to [0, 1]$ with $\lambda(0) = 0$ and $\lambda(1) = 1$. Denote by $I\colon [0, 1] \to [0, 1]$ the identity map. Let $\alpha_n \in \mathbb{D}$. Then $\alpha_n \overset{J_1}{\to} \alpha$ in $\mathbb{D}$ if there exists a sequence $(\lambda_n) \in \Lambda$ such that

$$\|\lambda_n - I\| \vee \|\alpha_n \circ \lambda_n - \alpha\| \to 0 \quad \text{as } n \to \infty.$$

Intuitively, J_1-convergence requires 'matching jumps' at 'matching points' after a deformation of time. We refer the reader to Jacod and Shiryaev [6, Chapter VI] and Billingsley [1, Section 12] for more information on the Skorokhod space. For other topologies on $\mathbb{D}$, we refer the reader to Skorokhod [12] and Whitt [14].

We proceed by setting out some basic methods of trimming extremes.

2.1. Global (pointwise) trimmers ('trim as you go')

Let $x = (x(\tau))_{0\le\tau\le1} \in \mathbb{D}$ with jump process $\Delta = \Delta x(\tau) := x(\tau) - x(\tau-))_{\tau>0}$. Set $\Delta x(0) \equiv 0$. Define the following extremal operators mapping $\mathbb{D}$ into $\mathbb{D}$:

(I) $\mathcal{S}(x)(\cdot) = \sup_{0\le s\le\cdot} x(s)$;

(II) $\widetilde{\mathcal{S}}(x)(\cdot) = \sup_{0\le s\le\cdot} |x(s)|$;

(III) $\mathcal{S}_{\pm\Delta}(x)(\cdot) = \mathcal{S}_{\Delta}(\pm x)(\cdot) := \sup_{0\le s\le\cdot} \Delta(\pm x)(s) \vee 0$; and

(IV) $\widetilde{\mathcal{S}}_{\Delta}(x)(\cdot) = \sup_{0\le s\le\cdot} |\Delta x(s)|$.

In (I), $\mathcal{S}(x)$ is the running supremum process of x, and in (II), $\widetilde{\mathcal{S}}(x)$ is the running supremum process for $|x|$. In (III), $\mathcal{S}_{+\Delta}(x)(\tau)$ represents the magnitude of largest (positive) jump of x

up to time τ, and $\mathcal{S}_{-\Delta}(x)(\tau)$ represents the largest magnitude of the negative jumps of x up to time τ; while in (IV), $\widetilde{\mathcal{S}}_{\Delta}(x)(\tau)$ represents the magnitude of largest jump in modulus of x up to time τ.

With these operators we can define what we call *global, or pointwise, trimming operators*. Let $\mathbb{N} := \{1, 2, \dots\}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. Take $r = 2, 3, \dots$ and define iteratively

(V) the rth extremal positive (negative) trimming operators

$$\mathcal{T}^{(1,\pm)}_{\text{trim}}(x) = x \mp \mathcal{S}_{\pm\Delta}(x) \quad \text{and} \quad \mathcal{T}^{(r,\pm)}_{\text{trim}}(x) = \mathcal{T}^{(1,\pm)}_{\text{trim}} \circ \mathcal{T}^{(r-1,\pm)}_{\text{trim}}(x);$$

(VI) the rth extremal positive (negative) jump operators

$$\mathcal{S}^{(1,\pm)}_{\Delta}(x) = \mathcal{S}_{\pm\Delta}(x) \quad \text{and} \quad \mathcal{S}^{(r,\pm)}_{\Delta}(x) = \mathcal{S}^{(1,\pm)}_{\Delta} \circ \mathcal{T}^{(r-1,\pm)}_{\text{trim}}(x);$$

(VII) the r, s trimming operators (for $s \in \mathbb{N}$)

$$\mathcal{T}^{(r,s)}_{\text{trim}}(x) = \mathcal{T}^{(r,+)}_{\text{trim}} \circ \mathcal{T}^{(s,-)}_{\text{trim}}(x) = \mathcal{T}^{(s,-)}_{\text{trim}} \circ \mathcal{T}^{(r,+)}_{\text{trim}}(x);$$

(VIII) the rth extremal modulus trimming operators

$$\widetilde{\mathcal{T}}_{\text{trim}}(x) = \widetilde{\mathcal{T}}^{(1)}_{\text{trim}}(x) = x - \widetilde{\mathcal{S}}_{\Delta}(x) \quad \text{and} \quad \widetilde{\mathcal{T}}^{(r)}_{\text{trim}}(x) = \widetilde{\mathcal{T}}_{\text{trim}} \circ \widetilde{\mathcal{T}}^{(r-1)}_{\text{trim}}(x);$$

(IX) and the rth extremal modulus jump operators

$$\widetilde{\mathcal{S}}^{(1)}_{\Delta}(x) = \widetilde{\mathcal{S}}_{\Delta}(x) \quad \text{and} \quad \widetilde{\mathcal{S}}^{(r)}_{\Delta}(x) = \widetilde{\mathcal{S}}_{\Delta} \circ \widetilde{\mathcal{T}}^{(r-1)}_{\text{trim}}(x).$$

In (V), $\mathcal{T}^{(r,+)}_{\text{trim}}(x)(\tau)$ is x with the r largest jumps of x up to time τ subtracted, and $\mathcal{T}^{(r,-)}_{\text{trim}}(x)(\tau)$ is similar with the r negative jumps of largest magnitude subtracted. In (VII), $\mathcal{T}^{(r,s)}_{\text{trim}}(x)$ has the r positive and s negative jumps of largest magnitudes subtracted, while in (VIII), $\widetilde{\mathcal{T}}^{(r)}_{\text{trim}}(x)(\tau)$ has the r largest jumps in modulus of x up to time τ removed from x. In (VI) and (IX), $\mathcal{S}^{(r,\pm)}_{\Delta}(x)$ and $\widetilde{\mathcal{S}}^{(r)}_{\Delta}(x)$ are the rth largest values in magnitude for positive (negative), or in modulus, jumps of the corresponding processes.

We call the operators in (V), (VII), and (VIII), 'trim as you go' operators because at each point in time, the designated number of largest positive (negative) jumps up to that point are removed from the process. See Figure 1 in Section 2.3 for an illustration with a schematic insurance risk process.

To analyse the convergence of these operators in $\mathbb{D}$, we need the following considerations. We say that an operator $\Psi\colon \mathbb{D} \to \mathbb{D}$ is $\|\cdot\|$-*continuous* at $x \in \mathbb{D}$ if $\lim_{n\to\infty} \|x_n - x\| = 0$ implies that $\lim_{n\to\infty} \|\Psi(x_n) - \Psi(x)\| = 0$. We say that Ψ is J_1-*continuous* at $x \in \mathbb{D}$ if $x_n \xrightarrow{J_1} x$ implies that $\Psi(x_n) \xrightarrow{J_1} \Psi(x)$ as $n \to \infty$. In general, Ψ being $\|\cdot\|$-continuous at x does not imply that Ψ is J_1-continuous at x. However, this is true if, in addition, Ψ is Λ-*compatible*, by which we mean $\Psi(x) \circ \lambda = \Psi(x \circ \lambda)$ for all $x \in \mathbb{D}$, $\lambda \in \Lambda$. The operator Ψ is called *jointly* J_1-*continuous at* x if, for any sequence x_n converging to x in the J_1-topology, there exists (λ_n) in Λ such that, simultaneously as $n \to \infty$, $\|\lambda_n - I\| \to 0$, $\|x_n \circ \lambda_n - x\| \to 0$, and $\|\Psi(x_n) \circ \lambda_n - \Psi(x)\| \to 0$. The following simple proposition summarises.

Proposition 1. *Let* $\Psi\colon \mathbb{D} \to \mathbb{D}$ *be* Λ*-compatible and take* $x \in \mathbb{D}$. *Suppose that* Ψ *is* $\|\cdot\|$*-continuous at* x. *Then* Ψ *is jointly* J_1*-continuous at* x.

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

Proof. Assume that Ψ is Λ-compatible and $\|\cdot\|$-continuous at $x \in \mathbb{D}$. Since $\Psi(x_n) \circ \lambda_n = \Psi(x_n \circ \lambda_n)$, $x_n \overset{J_1}{\to} x$, i.e. $\lim_{n\to\infty} \|x_n \circ \lambda_n - x\| = 0$, together with Ψ being $\|\cdot\|$-continuous at x, implies that $\lim_{n\to\infty} \|\Psi(x_n) \circ \lambda_n - \Psi(x)\| = 0$, i.e. $\Psi(x_n) \overset{J_1}{\to} \Psi(x)$, thus proving the proposition. □

Proposition 2. *Each of the operators defined in (I)–(IX) is*

(i) $\|\cdot\|$*-Lipschitz, hence, continuous in* $\|\cdot\|$ *norm;*

(ii) Λ*-compatible; and, consequently, by Proposition 1, jointly* J_1*-continuous.*

Proof. (i) For example, we prove (III). When $x, y \in \mathbb{D}$,

$$\begin{aligned}\|\mathcal{S}_\Delta(x) - \mathcal{S}_\Delta(y)\| &= \sup_{0<t\le 1} \Big\| \sup_{0<s\le t} (x(s) - x(s-)) - \sup_{0<s\le t} (y(s) - y(s-)) \Big\| \\ &\le \sup_{0<s\le 1} \|(x(s) - x(s-)) - (y(s) - y(s-))\| \\ &\le 2\|x - y\|, \end{aligned} \tag{1}$$

using the triangle inequality $|\|\alpha\| - \|\beta\|| \le \|\alpha - \beta\| \le \|\alpha\| + \|\beta\|$.

(ii) We prove this for (IV), for example. Let $x \in \mathbb{D}$, $\lambda \in \Lambda$, and $t \in [0, 1]$. Then

$$\widetilde{\mathcal{S}}_\Delta(x \circ \lambda)(t) = \sup_{0\le s\le t} |\Delta(x \circ \lambda)(s)| = \sup_{0\le s\le t} |\Delta x(\lambda(s))| = \sup_{0\le s\le \lambda(t)} |\Delta x(s)| = \widetilde{\mathcal{S}}_\Delta(x)(\lambda(t)),$$

completing the proof. □

Remark 1. We refer the reader to Jacod and Shiryaev [6, Section IV.2] for other continuity properties of common mappings in the Skorokhod topology.

2.2. Signed modulus trimmers

In Section 3 we will consider a Lévy process $X = (X_t)_{t\ge 0}$ which is to be trimmed. Before this, in the present subsection, we want to draw attention to an issue that arises with modulus trimming when considered pathwise. There may be one or more jumps equal in magnitude to the largest of $|\Delta X_s| = |X_s - X_{s-}|$ for $0 \le s \le t$. We refer to these as 'tied' values (for the modulus, with a similar concept for the positive and negative jumps).

Buchmann *et al.* [2] defined a 'modulus trimmed Lévy process' as follows. Denote the largest modulus jump of X up to time t, i.e. the jump corresponding to the largest of $|\Delta X_s|$, $0 \le s \le t$, by $\widetilde{\Delta X}_t^{(1)}$. When there is no tie for $\sup_{0\le s\le t} |\Delta X_s|$, the sign of $\widetilde{\Delta X}_t^{(1)}$ is uniquely determined. When there is a tie, the procedure in [2] is to nominate a jump chosen at random among the almost surely (a.s.) finite number of tied values according to a discrete uniform distribution on the collection of ties. While appropriate in the context of [2], this definition is problematic when we consider the sample path of the process on [0, 1]. To see why, take a simple example. Suppose that for some ω, the largest modulus jump up to time t is tied at values $0 < s_1 < s_2 < t$ with opposite signs, say

$$\Delta X_{s_1}(\omega) = |\widetilde{\Delta X}_t^{(1)}(\omega)| \quad \text{and} \quad \Delta X_{s_2}(\omega) = -|\widetilde{\Delta X}_t^{(1)}(\omega)|,$$

while $|\widetilde{\Delta X}_s^{(1)}(\omega)| = |\widetilde{\Delta X}_t^{(1)}(\omega)|$ for all $s \in [s_2, t]$. For each $s \in [s_2, t]$, if we were to choose from $\{\Delta X_{s_1}, \Delta X_{s_2}\}$ with equal probability to be trimmed from $X_s(\omega)$, the sample path of the resulting trimmed process would not be in $\mathbb{D}$.

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

Thus, we need to design a way to define signed modulus trimming on the sample path of X so as to stay within $\mathbb{D}$. One way to do this is as follows (we now revert to the general setup). For $x \in \mathbb{D}$, define the *last modulus record time process* on $[0, 1]$ as

$$\widetilde{L}_\tau(x) := \sup\{s \in [0, \tau] : |\Delta x(s)| = \widetilde{\mathscr{S}}_\Delta(x)(\tau)\} \quad \text{for each } \tau \in [0, 1]. \tag{2}$$

Then the *signed largest modulus jump* up to time $\tau \in [0, 1]$ is $\Delta x(\widetilde{L}_\tau(x))$, and the *signed largest trimmer* can be defined as $\mathfrak{T}_{\text{trim}}(x) := x - \Delta x(\widetilde{L}_\tau(x))$. More generally, interpret $\mathfrak{T}^{(0)}_{\text{trim}}(x) = x$, let $\mathfrak{T}^{(1)}_{\text{trim}}(x) := \mathfrak{T}_{\text{trim}}(x)$, and, for $r = 2, 3, \ldots$, set

$$\mathfrak{T}^{(r)}_{\text{trim}}(x) := \mathfrak{T}_{\text{trim}}(\mathfrak{T}^{(r-1)}_{\text{trim}}(x)).$$

Now $\mathfrak{T}^{(1)}_{\text{trim}}$ is not, in general, globally J_1-continuous, that is, it is not J_1-continuous at all $x \in \mathbb{D}$. Take, for example, $x = \mathbf{1}_{[1/3, 2/3)}$ and $x_n = x + (1/n)\,\mathbf{1}_{[1/3, 1]}$. Then $x_n \xrightarrow{J_1} x$, but

$$\mathfrak{T}^{(1)}_{\text{trim}}(x) = \mathbf{1}_{[2/3, 1]} \quad \text{and} \quad \mathfrak{T}^{(1)}_{\text{trim}}(x_n) = -\mathbf{1}_{[2/3, 1]}\,.$$

However, $\mathfrak{T}_{\text{trim}}$ is continuous when there is no sign change of ties in the limit. This is shown in Theorem 1, which uses the following notation. For each $\tau \in [0, 1]$, collect the times of occurrence of the largest values, and the times of occurrence of values having largest modulus, into sets $\mathbb{A}^{\pm}_\tau(x)$ and $\widetilde{\mathbb{A}}_\tau(x)$; thus,

$$\mathbb{A}^{\pm}_\tau(x) := \{0 < s \leq \tau : \Delta x(s) = \mathscr{S}_{\pm\Delta}(x)(\tau)\} \tag{3}$$

and

$$\widetilde{\mathbb{A}}_\tau(x) := \{0 < s \leq \tau : |\Delta x(s)| = \widetilde{\mathscr{S}}_\Delta(x)(\tau)\}. \tag{4}$$

We use the convention that when x is continuous on $[0, \tau]$ then $\mathbb{A}^{\pm}_\tau(x) = \widetilde{\mathbb{A}}_\tau(x) = \varnothing$. Recall that a càdlàg function has only finitely many jumps with magnitude bounded away from 0, so $\mathbb{A}^{\pm}_\tau(x)$ and $\widetilde{\mathbb{A}}_\tau(x)$ are finite sets (we include in this the possibility that one or other of them may be empty) for functions $x \in \mathbb{D}$. Collect the sign changing largest modulus jumps contained in $\widetilde{\mathbb{A}}_\tau(x) = \{s_1, \ldots, s_{\#\widetilde{\mathbb{A}}_\tau(x)}\}$ into the set

$$\mathbb{B}_\tau(x) := \{s_k \in \widetilde{\mathbb{A}}_\tau(x) : \Delta x(s_k) = -\Delta x(s_{k-1}), \text{ where } k = 2, \ldots, \#\widetilde{\mathbb{A}}_\tau(x)\}.$$

Note that $\#\widetilde{\mathbb{A}}_\tau(x) = 1$ implies that $\#\mathbb{B}_\tau(x) = 0$. Conversely, $\#\mathbb{B}_\tau(x) = 0$ implies that $\widetilde{\mathbb{A}}_\tau(x) = \mathbb{A}^{+}_\tau(x)$ or $\widetilde{\mathbb{A}}_\tau(x) = \mathbb{A}^{-}_\tau(x)$.

In the next theorem, we show that when there is no sign change among ties of large modulus jumps in x and its trimmed versions $\mathfrak{T}^{(j)}_{\text{trim}}(x)$ for all $0 \leq j \leq r-1$, $\mathfrak{T}^{(r)}_{\text{trim}}$ is jointly J_1-continuous at x. The next theorem is proved in Section 4.

Theorem 1. *It holds that $\mathfrak{T}_{\text{trim}}$ is jointly J_1-continuous at x if $\sup_{\tau\in[0,1]} \#\mathbb{B}_\tau(x) = 0$. Consequently, $\mathfrak{T}^{(r)}_{\text{trim}}$ is jointly J_1-continuous at x if $\sup_{\tau\in[0,1]} \#\mathbb{B}_\tau(\mathfrak{T}^{(j)}_{\text{trim}}(x)) = 0$ for all $j = 0, \ldots, r-1$, when $r \in \mathbb{N}$.*

2.3. Record times trimmers ('lookback trimming')

On the function space $\mathbb{D}$, we can extend the idea of trimming by including a random location where trimming starts and, hence, define a second kind of trimming. For $x \in \mathbb{D}$, define *the first (positive) record time* in $[0, 1]$ by

$$R_\tau(x) := \inf\{s \in [0, \tau] : \Delta x(s) = \mathscr{S}_{+\Delta}(x)(\tau)\} \quad \text{for } 0 < \tau \leq 1, \tag{5}$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

and, similarly, we could define the first (negative) record time. Likewise,

$$\widetilde{R}_\tau(x) := \inf\{s \in [0,1] \colon |\Delta x(s)| = \widetilde{\mathcal{S}}_\Delta(x)(\tau)\} \tag{6}$$

yields the *first modulus record time*. The corresponding *record time trimmers* are

$$\mathcal{R}_{\text{trim}}(x) := x - \Delta x(R_1(x))\,\mathbf{1}_{[R_1(x),\,1]}, \qquad \widetilde{\mathcal{R}}_{\text{trim}}(x) := x - \Delta x(\widetilde{R}_1(x))\,\mathbf{1}_{[\widetilde{R}_1(x),\,1]}. \tag{7}$$

Expanding, $\mathcal{R}_{\text{trim}}(x)(\tau)$ can be written for $\tau \in [0,1]$ as

$$\mathcal{R}_{\text{trim}}(x)(\tau) = \begin{cases} x(\tau) - \sup\limits_{0<s\le\tau} \Delta x(s) & \text{if } R_1(x) \le \tau, \\ x(\tau) & \text{otherwise.} \end{cases}$$

Thus, x is trimmed at time τ if the record occurs before τ, otherwise not. In Figure 1 we present an illustration of the two trimming types for a compound Poisson risk process as used in insurance risk modelling (see Section 6).

Set $\mathcal{R}^{(0)}_{\text{trim}}(x) = x$. For $r \in \mathbb{N}$, define the *rth order record times trimmers* as

$$\mathcal{R}^{(r)}_{\text{trim}}(x) := \mathcal{R}_{\text{trim}}(\mathcal{R}^{(r-1)}_{\text{trim}}(x)) = \mathcal{R}^{(r-1)}_{\text{trim}}(x) - \Delta\mathcal{R}^{(r-1)}_{\text{trim}}(R_1 \circ \mathcal{R}^{(r-1)}_{\text{trim}}(x))\,\mathbf{1}_{[R_1\circ\mathcal{R}^{(r-1)}_{\text{trim}}(x),\,1]}, \tag{8}$$

and, for $\widetilde{\mathcal{R}}^{(r)}_{\text{trim}}(x)$, replace each R_1 and $\mathcal{R}_{\text{trim}}$ in (8) by $\widetilde{R}_1$ and $\widetilde{\mathcal{R}}_{\text{trim}}$.

While the record trimming functionals are Λ-compatible, they are not, however, of the type described in Section 2.1. In fact they are not globally J_1-continuous.

Example 1. ($\mathcal{R}_{\text{trim}}$ *is not globally norm or* J_1*-continuous.*) First, $\mathcal{R}_{\text{trim}}$ is not $\|\cdot\|$-continuous. To see this, let $x := \mathbf{1}_{[1/3,\,1]} + \mathbf{1}_{[2/3,\,1]}$. For $n \in \mathbb{N}$, set $x_n := x + (1/n)\,\mathbf{1}_{[2/3,\,1]}$. Observe that $\lim_{n\to\infty} \|x_n - x\| \le \lim_{n\to\infty} 1/n = 0$. In particular, $x_n \xrightarrow{J_1} x$ as $n \to \infty$. However,

$$R_1(x) = \tfrac{1}{3}, \qquad \mathcal{R}_{\text{trim}}(x) = \mathbf{1}_{[2/3,\,1]}, \qquad R_1(x_n) = \tfrac{2}{3}, \qquad n \in \mathbb{N},$$

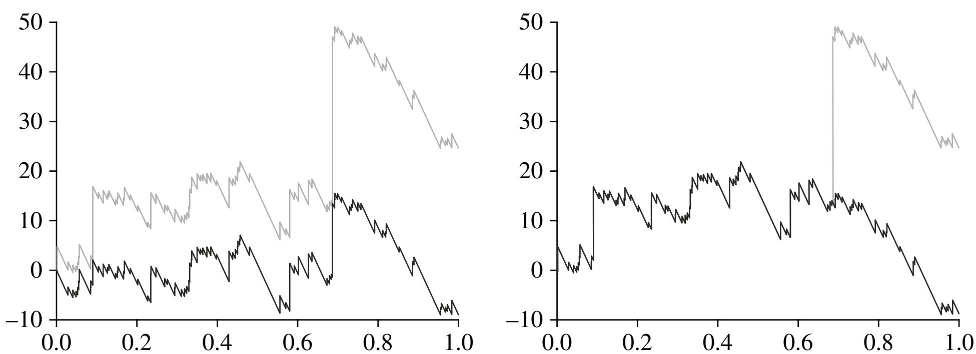

FIGURE 1: A realization of the compound Poisson risk insurance model $X_t = ct + \sum_{i=1}^{N_t} \xi_i$, $t \in [0,1]$, is represented by the solid grey line. The ξ_i are i.i.d. with a Pareto(1, 2) distribution, $c = -110$, and N_t is Poisson with $\mathbb{E}(N_1) = 100$. *Left:* the sample path of $\mathcal{T}^{(1,+)}_{\text{trim}}((X_t)_{t\in[0,1]})$ ('trim as you go') is represented by the solid black line. *Right:* the sample path of $\mathcal{R}_{\text{trim}}((X_t)_{t\in[0,1]})$ ('lookback trimming') is represented by the solid black line. This path coincides with the original process before $R_1((X_t)_{t\in[0,1]})$ and with the $\mathcal{T}^{(1,+)}_{\text{trim}}((X_t)_{t\in[0,1]})$ path after.

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

and

$$\mathcal{R}_{\text{trim}}(x_n) = \mathbf{1}_{[1/3,\,1]}, \qquad \|\mathcal{R}_{\text{trim}}(x_n) - \mathcal{R}_{\text{trim}}(x)\| = 1, \qquad n \in \mathbb{N}.$$

Thus, $\mathcal{R}_{\text{trim}}$ is Λ-compatible, but not $\|\cdot\|$-continuous in x.

Second, $\mathcal{R}_{\text{trim}}$ also is not J_1-continuous. To see this, take $(\lambda_n) \subseteq \Lambda$ with $\lim_{n\to\infty} \|\lambda_n - I\| = 0$. Then $\lim_{n\to\infty} \lambda_n^{-1}(\frac{1}{3}) = \frac{1}{3}$, hence, for $n \in \mathbb{N}$, once $\lambda_n^{-1}(\frac{1}{3}) < \frac{2}{3}$,

$$\begin{aligned}\|\mathcal{R}_{\text{trim}}(x) - \mathcal{R}_{\text{trim}}(x_n \circ \lambda_n)\| &= \|\mathcal{R}_{\text{trim}}(x) - \mathcal{R}_{\text{trim}}(x_n) \circ \lambda_n\| \\ &\geq \big|\mathcal{R}_{\text{trim}}(x)\big(\lambda_n^{-1}\big(\tfrac{1}{3}\big)\big) - \mathcal{R}_{\text{trim}}(x_n)\big(\tfrac{1}{3}\big)\big| \\ &= \mathbf{1}_{[1/3,\,2/3]}\big(\tfrac{1}{3}\big) = 1.\end{aligned}$$

Consequently, we have $x_n \xrightarrow{J_1} x$, but not $\mathcal{R}_{\text{trim}}(x_n) \xrightarrow{J_1} \mathcal{R}_{\text{trim}}(x)$.

Recall the definitions of $\mathbb{A}_\tau^{\pm}$ and $\widetilde{\mathbb{A}}_\tau$ in (3) and (4). Our main result of this section is that the record time trimmer $\mathcal{R}_{\text{trim}}$ is jointly J_1-continuous at x if and only if x does not admit ties. The next theorem is proved in Section 4.

Theorem 2. *Let $x \in \mathbb{D}$ and $r \in \mathbb{N}$.*

(i) *If $\#\mathbb{A}_1^{+}(x) \leq 1$ then $\mathcal{R}_{\text{trim}}$ is jointly J_1-continuous at x. Consequently, if*

$$\#\mathbb{A}_1^{+}(\mathcal{R}_{\text{trim}}^{(j)}(x)) \leq 1 \quad \text{for all } j = 0, \dots, r-1$$

then $\mathcal{R}_{\text{trim}}^{(r)}$ is jointly J_1-continuous at x.

(ii) *If $\mathcal{R}_{\text{trim}}$ is J_1-continuous at x then $\#\mathbb{A}_1^{+}(x) \leq 1$.*

The same holds with $\mathcal{R}_{\text{trim}}$ replaced by $\widetilde{\mathcal{R}}_{\text{trim}}$ and $\#\mathbb{A}_1^{+}(x)$ replaced by $\#\widetilde{\mathbb{A}}_1(x)$.

3. Functional laws for Lévy processes

A number of interesting processes can be derived by applying the operators in Section 2 to Lévy processes. In the present section, $X = (X_t)_{t\geq 0}$, $X_0 = X_{0-} = 0$, will be a real-valued càdlàg Lévy process with canonical triplet $(\gamma, \sigma^2 = 0, \Pi_X)$. The positive, negative, and two-sided tails of the Lévy measure Π_X are, for $x > 0$,

$$\overline{\Pi}_X^{+}(x) := \Pi_X\{(x,\infty)\}, \qquad \overline{\Pi}_X^{-}(x) := \Pi_X\{(-\infty,-x)\}, \qquad \overline{\Pi}_X(x) := \overline{\Pi}_X^{+}(x) + \overline{\Pi}_X^{-}(x).$$

The jump process of X is $(\Delta X_t = X_t - X_{t-})_{t\geq 0}$, the positive jumps are $\Delta X_t^{+} = \Delta X_t \vee 0$, and the magnitudes of the negative jumps are $\Delta X_t^{-} = (-\Delta X_t) \vee 0$. The processes $(\Delta X_t^{+})_{t\geq 0}$ and $(\Delta X_t^{-})_{t\geq 0}$, when present, are nonnegative independent processes. For any integers $r, s > 0$, let $\Delta X_t^{(r)}$ be the rth largest positive jump, and let $\Delta X_t^{(s),-}$ be the sth largest jump in $\{\Delta X_s^{-},\ 0 < s \leq t\}$, i.e. the negative of the sth smallest jump. These types of ordered jumps were carefully defined in [2], allowing for the possibility of tied values. (Recall the discussion in Subsection 2.2.) We can similarly define $\Delta X_{t-}^{(r)}$ and $\Delta X_{t-}^{(s),-}$ for the ordered jumps in $\{\Delta X_s,\ 0 < s < t\}$.

Throughout, for small time convergence ($t \downarrow 0$), we assume that $\overline{\Pi}_X(0+) = \infty$ when dealing with modulus trimming and $\overline{\Pi}_X^{+}(0+) = \infty$ or $\overline{\Pi}_X^{-}(0+) = \infty$ (or both when appropriate) when dealing with one-sided trimming. In particular, these ensure there are infinitely many jumps ΔX_t, or $\Delta X_t^{\pm}$, a.s. in any bounded interval of time.

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

As demonstrated in Section 2, the largest modulus trimming as defined in [2] is not natural for the functional setting, so here we adopt a modified definition. Write $\widetilde{\Delta X}_t^{(r)}$ to denote the rth largest jump in modulus up to time t, taking the sign of the *latest* rth largest modulus jump. Then define the trimmed Lévy processes

$$^{(r,s)}X_t := X_t - \sum_{i=1}^{r} \Delta X_t^{(i)} + \sum_{j=1}^{s} \Delta X_t^{(j),-} \quad \text{and} \quad ^{(r)}\widetilde{X}_t := X_t - \sum_{i=1}^{r} \widetilde{\Delta X}_t^{(i)},$$

which we call the *asymmetrically trimmed* and *modulus trimmed* processes, respectively. With the convention $\sum_1^0 \equiv 0$, taking $r = 0$ or $s = 0$ in asymmetrical trimming yields *one-sided trimmed processes* $^{(r)}X_t := {}^{(r,0)}X_t$ and $^{(s,-)}X_t := {}^{(0,s)}X_t$.

In this section we apply a functional law for Lévy processes attracted to a nonnormal stable law to obtain two theorems for trimmed Lévy processes. We say that X_t is in a *nonnormal domain of attraction* at small (large) times if there exist nonstochastic functions $a_t \in \mathbb{R}$ and $b_t > 0$ such that

$$\frac{X_t - a_t}{b_t} \xrightarrow{\text{D}} Y \quad \text{as } t \downarrow 0 \, (t \to \infty), \tag{9}$$

where Y is an a.s. finite, nondegenerate, nonnormal random variable. Then (9) implies that the two-sided tail $\overline{\Pi}_X$ of X is regularly varying (at 0 or ∞, as appropriate), with index $\alpha \in (0, 2)$. The limit random variable Y has the distribution of Y_1, where $(Y_\tau)_{0 \le \tau \le 1} \equiv \mathcal{Y}$ is a stable(α) Lévy process. The canonical triplet for $\mathcal{Y}$ will be taken as $(0, 0, \Pi_Y)$, where Π_Y has tail function $\overline{\Pi}_Y(x) = cx^{-\alpha}$, $x > 0$, for some $c > 0$.

In the small time case, conditions on the Lévy measure for (9) to hold can be deduced from Maller and Mason [9, Theorem 2.2], whose result can also be used to show that (9) can be extended to convergence in $\mathbb{D}$; that is,

$$\mathcal{I}_t = \{\mathcal{I}_t(\tau)\}_{0 \le \tau \le 1} := \left(\frac{X_{\tau t} - \tau a_t}{b_t}\right)_{0 \le \tau \le 1} \to (Y_\tau)_{0 \le \tau \le 1} = \mathcal{Y}, \tag{10}$$

weakly as $t \downarrow 0$ with respect to the J_1-topology. Large time ($t \to \infty$) convergence in (10) also follows from (9) as is well known.

Assuming the convergence in (10), we can prove a variety of interesting functional limit theorems for X by applying the operators in Section 2. We list some examples in Theorem 3 and prove them in Section 5.

In Theorem 3(i)–(iii) we consider lookback trimming, two-sided (or one-sided, with r or s taken as 0) trimming, and signed modulus trimming defined as in Subsection 2.2, respectively. To specify the lookback trimming in this situation, recall the definition of the record time trimming functionals in (7) and (8). Using them, we define, for X, *lookback trimmed paths of order* r, based on *positive* jumps, being processes on $\tau \in [0, 1]$, indexed by $t > 0$, as

$$(^{(1)}X_{t\tau}^R)_{\tau \in [0,1]} = \mathcal{R}_{\text{trim}}((X_{t\tau})_{\tau \in [0,1]}),$$

and, for $r = 2, 3, \ldots,$

$$(^{(r)}X_{t\tau}^R)_{\tau \in [0,1]} = \mathcal{R}_{\text{trim}}^{(r)}((X_{t\tau})_{\tau \in [0,1]}) =: \mathcal{R}_{\text{trim}}(\mathcal{R}_{\text{trim}}^{(r-1)}((X_{t\tau})_{\tau \in [0,1]})).$$

Theorem 3. *Assume that* $(X_t)_{t \ge 0}$ *is in the domain of attraction of a stable law at* 0 *with nonstochastic centering and norming functions* $a_t \in \mathbb{R}$, $b_t > 0$, *so that (9) and (10) hold. In the following, convergences are with respect to the* J_1*-topology in* $\mathbb{D}$.

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

(i) *Suppose that* $\overline{\Pi}_X^+(0+) = \infty$ *and* $r \in \mathbb{N}$. *Then, for the same* a_t *and* b_t,

$$\left(\frac{{}^{(r)}X_{\tau t}^R - \tau a_t}{b_t}\right)_{0\le\tau\le1} \xrightarrow{\text{D}} ({}^{(r)}Y_\tau^R)_{0\le\tau\le1} =: {}^{(r)}\mathcal{Y}^R \text{ in } \mathbb{D} \quad \text{as } t \downarrow 0. \tag{11}$$

(ii) *Assume that* $\overline{\Pi}_X^+(0+) = \overline{\Pi}_X^-(0+) = \infty$ *and* $r, s \in \mathbb{N}_0$. *Then, for the same* a_t *and* b_t,

$$\left(\frac{{}^{(r,s)}X_{\tau t} - \tau a_t}{b_t}\right)_{0\le\tau\le1} \xrightarrow{\text{D}} ({}^{(r,s)}Y_\tau)_{0\le\tau\le1} =: {}^{(r,s)}\mathcal{Y} \text{ in } \mathbb{D} \quad \text{as } t \downarrow 0. \tag{12}$$

(iii) *Suppose that only* $\overline{\Pi}_X(0+) = \infty$ *and* $r \in \mathbb{N}$. *Then, for the same* a_t *and* b_t,

$$\left(\frac{{}^{(r)}\widetilde{X}_{\tau t} - \tau a_t}{b_t}\right)_{0\le\tau\le1} \xrightarrow{\text{D}} ({}^{(r)}\widetilde{Y}_\tau)_{0\le\tau\le1} =: {}^{(r)}\widetilde{\mathcal{Y}} \text{ in } \mathbb{D} \quad \text{as } t \downarrow 0. \tag{13}$$

Remark 2. (i) Buchmann *et al.* [2] and also Maller and Mason [9] included convergence of the quadratic variation of X in their expositions. Using these as basic convergences (i.e. together with (9) and (10)) would lead to functional convergences of the jointly trimmed process together with its trimmed quadratic variation process, and we could then consider self-normalised versions. But we omit the details of these.

(ii) Fan [3] proved the converses of Theorems 3(ii) and 3(iii) for $t \downarrow 0$, i.e. if the convergence in (12) or (13) holds for a fixed $\tau > 0$, then X is in the domain of attraction of a stable law with index $\alpha \in (0, 2)$ at small times. When Y is $N(0, 1)$, a standard normal random variable, the large jumps are asymptotically negligible with respect to b_t and (11) and (12) remain true with ${}^{(r,s)}\mathcal{Y}$ and ${}^{(r)}\widetilde{\mathcal{Y}}$ a standard Brownian motion; see [4].

We conclude this section by mentioning that the same methods can be used to obtain functional convergence for jumps of an extremal process together with trimmed versions. Again, we omit further details.

4. Proofs for Section 2

For the proof of Theorem 1 we need a preliminary lemma. Let $\mathfrak{C}(r_n)$ denote the set of accumulation points of a sequence $(r_n) \subseteq \mathbb{R}$ as $n \to \infty$. Recall that $\widetilde{L}_\tau$ and $\widetilde{R}_\tau$, the last and first modulus record time processes, are defined in (2) and (6); R_τ, the first positive record time process, is defined in (5).

Lemma 1. *Take* $x \in \mathbb{D}$ *and suppose that* $(x_n) \subseteq \mathbb{D}$ *with* $x_n \xrightarrow{J_1} x$. *Then, for each* $\tau \in [0, 1]$,

(i) *if* $\widetilde{\mathbb{A}}_\tau(x) \neq \varnothing$ *then* $\mathfrak{C}(\widetilde{R}_\tau(x_n)) \subseteq \widetilde{\mathbb{A}}_\tau(x)$ *and* $\mathfrak{C}(\widetilde{L}_\tau(x_n)) \subseteq \widetilde{\mathbb{A}}_\tau(x)$;

(ii) *if* $\mathbb{A}_\tau^+(x) \neq \varnothing$ *then* $\mathfrak{C}(R_\tau(x_n)) \subseteq \mathbb{A}_\tau^+(x)$;

(iii) *if* $\|x_n - x\| \to 0$ *and* $\#\widetilde{\mathbb{A}}_\tau(x) = 1$ *then, for all sufficiently large* n, $\widetilde{R}_\tau(x_n) = \widetilde{L}_\tau(x_n) = \widetilde{R}_\tau(x) = \widetilde{L}_\tau(x)$ *for each* $\tau \in [0, 1]$.

Proof. (i) We consider the $\mathbb{A}^+$ case only; $\widetilde{\mathbb{A}}$ can be argued similarly.

(ii) Take $x \in \mathbb{D}$ and let $(x_n) \subseteq \mathbb{D}$ with $x_n \xrightarrow{J_1} x$. Then there are $\lambda_n \in \Lambda$ such that $\|\lambda_n - I\| \vee \|y_n - x\| \to 0$ for $y_n := x_n \circ \lambda_n$. This also means that $\|\Delta y_n - \Delta x\| \to 0$.

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

Fix a time $\tau \in [0, 1]$ and assume that $\mathbb{A}_\tau^+(x) \neq \varnothing$. Recall that $\mathbb{A}_\tau^+(x)$ is a finite set. Let $\mathbb{A}_\tau^+ = \{s_1, \ldots, s_N\} \neq \varnothing$ with $s_1 = R_\tau(x) = \min \mathbb{A}_\tau^+(x)$. Observe that $R_\tau(y_n) = \lambda_n^{-1}(R_\tau(x_n))$ and, thus, since $\|y_n - x\| \to 0$, $\mathfrak{C}(R_\tau(x_n)) = \mathfrak{C}(R_\tau(y_n))$. To complete the proof, it thus suffices to show that $\mathfrak{C}(R_\tau(y_n)) \subseteq \mathbb{A}_\tau^+(x)$.

To see this, note that there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, $8(1 + \#\mathbb{A}_\tau^+(x))\|y_n - x\| \leq \delta$ (because $\|y_n - x\| \to 0$ and $\mathbb{A}_\tau^+(x)$ is finite), and, also,

$$\mathcal{S}_\Delta\biggl(x - \sum_{s \in \mathbb{A}_\tau^+(x)} \Delta x(s)\, \mathbf{1}_{[s,1]}\biggr)(\tau) < \Delta x(R_\tau(x)) - \delta. \tag{14}$$

To explain (14): the quantity

$$\biggl(x - \sum_{s \in \mathbb{A}_\tau^+(x)} \Delta x(s)\, \mathbf{1}_{[s,1]}\biggr)(\tau)$$

is x with all positive jumps equal in magnitude to the largest jump up to time τ subtracted. Applying the operator $\mathcal{S}_\Delta$ to this produces the largest of the remaining jumps, hence, the second largest jump in magnitude of x up to time τ. This is strictly smaller than the magnitude of the largest jump, which is $\Delta x(R_\tau(x))$. So indeed there is a $\delta > 0$ such that (14) holds.

Let $\alpha \in \mathfrak{C}(R_\tau(y_n))$ be the limit along a subsequence $(n') \subseteq (n)$. Contrary to the hypothesis, suppose that $\alpha \notin \mathbb{A}_\tau^+(x)$ and, thus, $\alpha_{n'} := R_\tau(y_{n'}) \notin \mathbb{A}_\tau^+(x)$ for all sufficiently large n'. For those n' also being larger than n_0, observe that

$$\begin{aligned}
\Delta x(R_\tau(x)) &= \mathcal{S}_\Delta(x)(\tau) \\
&\leq \mathcal{S}_\Delta(y_{n'})(\tau) + 2\|y_{n'} - x\| \\
&= \mathcal{S}_\Delta\biggl(y_{n'} - \sum_{s \in \mathbb{A}_\tau^+(x)} \Delta y_{n'}(s)\, \mathbf{1}_{[s,1]}\biggr)(\tau) + 2\|y_{n'} - x\|.
\end{aligned}$$

Here the second equality holds because $\alpha_{n'} \notin \mathbb{A}_\tau^+(x)$ implies that $\Delta y_{n'}(s) \leq \mathcal{S}_\Delta(y_{n'})(\tau)$ for any $s \in \mathbb{A}_\tau^+(x)$, for large n', and, thus, subtracting any such jumps from $y_{n'}$ does not affect the value of $\mathcal{S}_\Delta(y_{n'})$. Using (1) again now yields

$$\begin{aligned}
\Delta x(R_\tau(x)) &\leq \mathcal{S}_\Delta\biggl(x - \sum_{s \in \mathbb{A}_\tau^+(x)} \Delta x(s)\, \mathbf{1}_{[s,1]}\biggr)(\tau) + 4(1 + \#\mathbb{A}_\tau^+(x))\|y_{n'} - x\| \\
&\leq \Delta x(R_\tau(x)) - \delta + \tfrac{1}{2}\delta \\
&= \Delta x(R_\tau(x)) - \tfrac{1}{2}\delta,
\end{aligned}$$

where the last inequality holds by (14). This contradiction yields $\alpha \in \mathbb{A}_\tau^+(x)$, completing the proof that $\mathfrak{C}(R_\tau(y_n)) \subset \mathbb{A}_\tau^+(x)$.

(iii) Again, suppose that $\|\lambda_n - I\| \vee \|y_n - x\| \to 0$ and, in addition, $\#\mathbb{A}_\tau^+(x) = 1$. We can take $16\|y_n - x\| \leq \delta$ and (14) now takes the form

$$\mathcal{S}_\Delta(x - \Delta x(R_\tau(x))\, \mathbf{1}_{[R_\tau(x),\,1]})(\tau) < \Delta x(R_\tau(x)) - \delta \quad \text{for some } \delta > 0 \text{ and all } n \geq n_0.$$

Suppose that $R_\tau(y_n) \neq R_\tau(x)$ for some n. Then we must have $n < n_0$, as otherwise

$$\begin{aligned}
\Delta x(R_\tau(x)) &\leq |\Delta y_n(R_\tau(x))| + 2\|y_n - x\| \\
&\leq |\Delta y_n(R_\tau(y_n))| + 2\|y_n - x\| \\
&\leq |\Delta x(R_\tau(y_n))| + 4\|y_n - x\|
\end{aligned}$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

$$
\begin{aligned}
&< \Delta x(R_\tau(x)) - \delta + \tfrac{1}{3}\delta \\
&= \Delta x(R_\tau(x)) - \tfrac{2}{3}\delta.
\end{aligned}
$$

This contradiction proves the result. □

Proof of Theorem 1. Assume that $\sup_{\tau\in[0,1]} \#\mathbb{B}_\tau(x) = 0$. We first show that $\mathfrak{T}_{\text{trim}}$ is Λ-compatible. Let $\lambda \in \Lambda$ and recall from Proposition 2 that $\widetilde{\mathcal{S}}_\Delta$ is Λ-compatible. Since

$$\widetilde{L}_\tau(x \circ \lambda) = \sup\{0 \le s \le \tau \colon \Delta x(\lambda(s)) = \widetilde{\mathcal{S}}_\Delta(x)(\lambda(\tau))\}$$

and

$$\widetilde{L}_{\lambda(\tau)}(x) = \sup\{0 \le s \le \lambda(\tau) \colon \Delta x(s) = \widetilde{\mathcal{S}}_\Delta(x)(\lambda(\tau))\},$$

we have $\lambda^{-1} \circ \widetilde{L}_{\lambda(\tau)}(x) = \widetilde{L}_\tau(x \circ \lambda)$. Thus,

$$\Delta(x \circ \lambda)(\widetilde{L}_\tau(x \circ \lambda)) = \Delta(x \circ \lambda)(\lambda^{-1} \circ \widetilde{L}_{\lambda(\tau)}(x)) = \Delta x(\widetilde{L}_{\lambda(\tau)}(x)).$$

Then $\mathfrak{T}_{\text{trim}}$ is Λ-compatible, since

$$
\begin{aligned}
\mathfrak{T}_{\text{trim}}(x \circ \lambda) &= x \circ \lambda - \Delta(x \circ \lambda)(\widetilde{L}_\tau(x \circ \lambda)) \\
&= x \circ \lambda - \Delta x(\widetilde{L}_\tau(x)) \circ \lambda \\
&= (x - \Delta x(\widetilde{L}_\tau(x)) \circ \lambda \\
&= \mathfrak{T}_{\text{trim}}(x) \circ \lambda.
\end{aligned}
$$

It remains to show that $\mathfrak{T}_{\text{trim}}$ is $\|\cdot\|$-continuous at x. Suppose that $\|x_n - x\| \to 0$. If $\widetilde{\mathbb{A}}_\tau(x) = \varnothing$ then $\mathfrak{T}_{\text{trim}}(x) = x$, hence, $\mathfrak{T}_{\text{trim}}$ is trivially $\|\cdot\|$-continuous. Alternatively, suppose that $\widetilde{\mathbb{A}}_\tau \neq \varnothing$. Then

$$\|\mathfrak{T}_{\text{trim}}(x_n) - \mathfrak{T}_{\text{trim}}(x)\| \le \|x_n - x\| + \sup_{0\le\tau\le1} |\Delta x_n(\widetilde{L}_\tau(x_n)) - \Delta x(\widetilde{L}_\tau(x))|.$$

The first term on the right-hand side tends to 0 and the second term on the right-hand side does not exceed

$$
\begin{aligned}
&\sup_{0\le\tau\le1} |\Delta x_n(\widetilde{L}_\tau(x_n)) - \Delta x(\widetilde{L}_\tau(x_n))| + \sup_{0\le\tau\le1} |\Delta x(\widetilde{L}_\tau(x_n)) - \Delta x(\widetilde{L}_\tau(x))| \\
&\qquad \le 2\|x_n - x\| + \sup_{0\le\tau\le1} |\Delta x(\widetilde{L}_\tau(x_n)) - \Delta x(\widetilde{L}_\tau(x))|. \qquad (15)
\end{aligned}
$$

By Lemma 1, $\mathfrak{C}_\infty(\widetilde{L}_\tau(x_n)) \subseteq \widetilde{\mathbb{A}}_\tau(x)$ for each $\tau \in [0, 1]$. If $\widetilde{L}_\tau(x_n) \to \widetilde{L}_\tau(x)$ then the second term on the right-hand side of (15) tends to 0. Suppose that $\widetilde{L}_\tau(x_n) \to s_1 \neq \widetilde{L}_\tau(x) = s_2$, where $s_1, s_2 \in \widetilde{\mathbb{A}}_\tau(x)$. Then $|\Delta x(s_1)| = |\Delta x(s_2)|$. But since $\#\mathbb{B}_\tau(x) = 0$ for all $\tau \in [0, 1]$, we also have $\Delta x(s_1) = \Delta x(s_2)$. Thus, $\sup_{0\le\tau\le1} |\Delta x(\widetilde{L}_\tau(x_n)) - \Delta x(\widetilde{L}_\tau(x))| \to 0$. This completes the proof. □

Proof of Theorem 2. Again, we only consider the $\mathbb{A}_1^+$ case.

(i) Let $x_n \xrightarrow{J_1} x$ in $\mathbb{D}$ and $\#\mathbb{A}_1^+(x) \le 1$. Then there exists a sequence $\lambda_n \in \Lambda$ such that $\|\lambda_n - I\| \vee \|x \circ \lambda_n - x\| \to 0$. If $\mathbb{A}_1^+(x) = \varnothing$ then $\mathcal{S}_\Delta(x) \equiv 0$ and $\mathcal{R}_{\text{trim}}(x) = x$, so

$$
\begin{aligned}
\|\mathcal{R}_{\text{trim}}(x_n) \circ \lambda_n - \mathcal{R}_{\text{trim}}(x)\| &= \|\mathcal{R}_{\text{trim}}(x_n \circ \lambda_n) - x\| \\
&\le \|x_n \circ \lambda_n - x\| + |\Delta(x_n \circ \lambda_n)(R_1(x_n \circ \lambda_n))| \\
&= \|x_n \circ \lambda_n - x\| + \|\mathcal{S}_\Delta(x_n \circ \lambda_n) - \mathcal{S}_\Delta(x)\| \\
&\le 3\|x_n \circ \lambda_n - x\| \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}
$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

Alternatively, if $\#\mathbb{A}_1^+ = 1$ then by Lemma 1 there is an $n_0 \in \mathbb{N}$ such that $R_1(x_n \circ \lambda_n) = R_1(x)$ for all $n \geq n_0$ and, for those n, we also have

$$\mathcal{R}_{\text{trim}}(x_n) \circ \lambda_n = \mathcal{R}_{\text{trim}}(x_n \circ \lambda_n) = x_n \circ \lambda_n - \Delta x_n(R_1(x))\, \mathbf{1}_{[R_1(x),\,1]}\,.$$

As $n \to \infty$, the right-hand side converges uniformly (in the supremum norm) to

$$x - \Delta x(R_1(x))\, \mathbf{1}_{[R_1(x),\,1]} = \mathcal{R}_{\text{trim}}(x),$$

which shows that $\mathcal{R}_{\text{trim}}$ is jointly J_1-continuous at x. For $r = 2$, recall the definition of $\mathcal{R}^{(2)}_{\text{trim}}$ from (8). Since $\mathcal{R}_{\text{trim}}$ is J_1-continuous at x and $\mathcal{R}_{\text{trim}}$ is assumed J_1-continuous at $\mathcal{R}_{\text{trim}}(x)$, then the composition $\mathcal{R}^{(2)}_{\text{trim}}(x) = \mathcal{R}_{\text{trim}}(\mathcal{R}_{\text{trim}}(x))$ is J_1-continuous at x. An analogous argument holds for $r > 2$.

(ii) Contrary to the hypothesis, assume that $\{s_1, s_2\} \subseteq \mathbb{A}_1^+(x)$ for some $0 < s_1 < s_2 \leq 1$. Noting that $s_1 = R_1(x)$ and $\mathcal{R}_{\text{trim}}(x) = x - \Delta x(s_1)\, \mathbf{1}_{[s_1,\,1]}$, we introduce

$$x_n := x + \frac{1}{n}\, \mathbf{1}_{[s_2,\,1]}, \qquad n \in \mathbb{N}.$$

As $n \to \infty$, we have $\|x_n - x\| = 1/n \to 0$ and, in particular, $x_n \overset{J_1}{\to} x$. Observe that $R_1(x_n) = s_2$ and $\mathcal{R}_{\text{trim}}(x_n) = x - \Delta x(s_2)\, \mathbf{1}_{[s_2,\,1]}$, $n \in \mathbb{N}$. Hence, for all n,

$$\|\mathcal{R}_{\text{trim}}(x_n) - \mathcal{R}_{\text{trim}}(x)\| \geq |\mathcal{R}_{\text{trim}}(x_n)(s_1) - \mathcal{R}_{\text{trim}}(x)(s_1)| = |\Delta x(s_1)| > 0.$$

Finally, let $(\lambda_n) \subseteq \Lambda$ be such that $\lim_{n\to\infty} \|\lambda_n - I\| = 0$. Then $\lim_{n\to\infty} \lambda_n^{-1}(s_1) = s_1$. As $\mathcal{R}_{\text{trim}}(x)$ is continuous at $s_1 = R_1(x)$,

$$\delta_n := \mathcal{R}_{\text{trim}}(x)(s_1) - \mathcal{R}_{\text{trim}}(x)(\lambda_n^{-1}(s_1)) \to 0 \quad \text{as } n \to \infty,$$

and, thus,

$$\begin{aligned}\|\mathcal{R}_{\text{trim}}(x_n \circ \lambda_n) - \mathcal{R}_{\text{trim}}(x)\| &= \|\mathcal{R}_{\text{trim}}(x_n) - \mathcal{R}_{\text{trim}}(x) \circ \lambda_n^{-1}\| \\ &\geq |\mathcal{R}_{\text{trim}}(x_n)(s_1) - \mathcal{R}_{\text{trim}}(x)(\lambda_n^{-1}(s_1))| \\ &= |\Delta x(s_1) + \delta_n| \\ &\to |\Delta x(s_1)| > 0 \quad \text{as } n \to \infty.\end{aligned}$$

To summarise, we showed that $x_n \overset{J_1}{\to} x$, but not $\mathcal{R}_{\text{trim}}(x_n) \overset{J_1}{\to} \mathcal{R}_{\text{trim}}(x)$, contradicting the J_1-continuity of $\mathcal{R}_{\text{trim}}$ at x. □

5. Proof of Theorem 3

Let $(\Delta X_s)_{0<s\leq t}$ be the jumps of a Lévy process $(X_s)_{0<s\leq t}$ having Lévy measure Π_X, with ordered jumps $(\Delta X_t^{(i)})_{i\geq 1}$ and $(\Delta X_t^{(j),-})_{j\geq 1}$, as specified in Section 3. In what follows, we will assume that $\overline{\Pi}_X^+(0+) = \overline{\Pi}_X^-(0+) = \infty$ throughout, so there are always infinitely many positive and negative jumps of X, a.s., in any interval of time.

Let $(\mathcal{E}_i)_{i\geq 1}$ be an i.i.d. sequence of exponentially distributed random variables with common parameter $E\mathcal{E}_i = 1$ and let $\Gamma_r := \sum_{i=1}^r \mathcal{E}_i$ with $r \in \mathbb{N}$. Write

$$\overline{\Pi}_X^{+,\leftarrow}(x) = \inf\{y > 0 \colon \overline{\Pi}_X^+(y) \leq x\}, \qquad x > 0,$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

for the right-continuous inverse of the right tail $\overline{\Pi}_X^+$ (with similar notation for the left tail $\overline{\Pi}_X^-$ and the two-sided tail $\overline{\Pi}_X$). The following distributional equivalence can be deduced from [2, Lemma 1.1]:

$$(\Delta X_t^{(i)})_{1\le i\le r} \stackrel{\mathrm{D}}{=} \left(\overline{\Pi}_X^{+,\leftarrow}\left(\frac{\Gamma_i}{t}\right)\right)_{1\le i\le r}, \qquad t>0,\ r\in\mathbb{N}. \tag{16}$$

We refer the reader to [2] for more background information on the properties of the extremal processes $(\Delta X_t^{(r)})_{t\ge 0}$ and the trimmed Lévy processes.

Proof of Theorem 3. (i) We state proofs just for $t\downarrow 0$; $t\to\infty$ is very similar. Recall the definition of $\{\mathcal{I}_t(\tau)\}_{\tau\in[0,1]}$ from (10) and assume the convergence of $\mathcal{I}_t$ to a stable process $\mathcal{Y}$ as in (10). The process $\mathcal{Y}$ has Lévy measure Π_Y which is diffuse (continuous at each $x\in\mathbb{R}\setminus\{0\}$).

For each $\tau\in(0,1]$, the jump of $\mathcal{I}_t$ at τ is

$$\Delta\mathcal{I}_t(\tau) := \mathcal{I}_t(\tau)-\mathcal{I}_t(\tau-) = \frac{X_{t\tau}-\tau a_t}{b_t}-\frac{X_{t\tau-}-\tau a_t}{b_t}=\frac{\Delta X_{t\tau}}{b_t}. \tag{17}$$

Hence, $\mathcal{S}_\Delta(\mathcal{I}_t)(\tau)=\mathcal{S}_\Delta(X_{t\tau}/b_t)$ for each $t>0$, and we can write

$$\left(\frac{{}^{(r)}X_{\tau t}^R-\tau a_t}{b_t}\right)_{\tau\in[0,1]} = \mathcal{R}_{\mathrm{trim}}^{(r)}(\mathcal{I}_t).$$

We want to apply the continuous mapping theorem and deduce the convergence in (11) from this. By Theorem 2, to apply the continuous mapping theorem it is enough to verify that there are no ties a.s. among the first r largest positive jumps in the limit process $\mathcal{Y}$. Let $\mathcal{C}:=\{x\in\mathbb{D}\colon \#\mathbb{A}_1^+(\mathcal{R}_{\mathrm{trim}}^{(j)}(x))\le 1 \text{ for all } j=0,\dots,r-1\}$. We wish to show that $\mathbb{P}(\mathcal{Y}\in\mathcal{C})=1$. Denote by $\Delta Y_1^{(j)}$ the jth largest jump of $\mathcal{Y}$ up to time 1. Note that

$$\mathbb{P}(\mathcal{Y}\in\mathcal{C})\ge 1-\sum_{j=1}^r\mathbb{P}(\Delta Y_1^{(j)}=\Delta Y_1^{(j+1)}).$$

Since Π_Y is diffuse, we have $\overline{\Pi}_Y^+(\overline{\Pi}_Y^{+,\leftarrow}(v))=v=\overline{\Pi}_Y^+(\overline{\Pi}_Y^{+,\leftarrow}(v)-)$ for all $v>0$. Thus, by (16) (with X and Π_X replaced by Y and Π_Y),

$$\begin{aligned}\mathbb{P}(\Delta Y_1^{(j)}=\Delta Y_1^{(j+1)}) &= \mathbb{P}\{\overline{\Pi}_Y^{+,\leftarrow}(\Gamma_j)=\overline{\Pi}_Y^{+,\leftarrow}(\Gamma_j+\mathcal{E}_{j+1})\}\\ &= \int_0^\infty \mathbb{P}\{\overline{\Pi}_Y^{+,\leftarrow}(v)=\overline{\Pi}_Y^{+,\leftarrow}(v+\mathcal{E}_{j+1})\}\mathrm{e}^{-v}\frac{v^{j-1}}{(j-1)!}\,\mathrm{d}v\\ &= \int_0^\infty \mathbb{P}\{0\le\mathcal{E}_{j+1}\le\overline{\Pi}_Y^+(\overline{\Pi}_Y^{+,\leftarrow}(v)-)-v\}\mathrm{e}^{-v}\frac{v^{j-1}}{(j-1)!}\,\mathrm{d}v\\ &= 0.\end{aligned}$$

So we can apply Theorem 2 as forecast and complete the proof.

(ii) We first prove (12) and consider only the trimming operator $\mathcal{T}_{\mathrm{trim}}^{(1,+)}$ ($\mathcal{T}_{\mathrm{trim}}^{(r,s)}$ is treated analogously). By (17) we can write, for each $\tau\in(0,1]$ and $t>0$,

$$\mathcal{T}_{\mathrm{trim}}^{(1,+)}(\mathcal{I}_t)(\tau)=\frac{X_{\tau t}-\tau a_t}{b_t}-\sup_{0<s\le\tau}\Delta\mathcal{I}_t(s)=\frac{{}^{(1)}X_{\tau t}-\tau a_t}{b_t}.$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

Since $\mathcal{T}^{(1,+)}_{\text{trim}}$ is globally J_1-continuous on $\mathbb{D}$ by Proposition 2, we can apply the continuous mapping theorem to obtain

$$\left(\frac{{}^{(1)}X_{\tau t} - \tau a_t}{b_t}\right)_{0\le\tau\le1} = \mathcal{T}^{(1,+)}_{\text{trim}}(\mathcal{I}_t) \xrightarrow{\text{D}} \mathcal{T}^{(1,+)}_{\text{trim}}(\mathcal{Y}) = ({}^{(1)}Y_\tau)_{0\le\tau\le1}$$

in J_1 as $t \downarrow 0$ or $t \to \infty$. This completes the proof of (12).

(iii) Again, by (17) we have

$$\left(\frac{{}^{(r)}\widetilde{X}_{\tau t} - \tau a_t}{b_t}\right)_{0\le\tau\le1} = \mathfrak{T}^{(r)}_{\text{trim}}(\mathcal{I}_t).$$

From Theorem 1, recall that $\mathfrak{T}^{(r)}_{\text{trim}}$ is J_1-continuous on $\widetilde{\mathcal{C}}$, where

$$\widetilde{\mathcal{C}} := \{x \in \mathbb{D}\colon \#\mathbb{B}_\tau(\mathfrak{T}^{(j)}_{\text{trim}}(x)) = 0 \text{ for all } \tau \in [0,1],\ j = 0,\dots,r-1\}.$$

Thus, in order to apply the continuous mapping theorem, we need to show that $\mathbb{P}(\mathcal{Y} \in \widetilde{\mathcal{C}}) = 1$. Note that $\widetilde{\mathcal{C}} \supseteq \widetilde{\mathcal{V}}$, where

$$\widetilde{\mathcal{V}} := \{x \in \mathbb{D}\colon \#\widetilde{\mathbb{A}}_\tau(\mathfrak{T}^{(j)}_{\text{trim}}(x)) \le 1 \text{ for all } \tau \in [0,1],\ j = 0,\dots,r-1\}.$$

Hence, it is enough to show that $\mathbb{P}(\mathcal{Y} \in \widetilde{\mathcal{V}}) = 1$, or, equivalently,

$$\mathbb{P}\left(\bigcup_{1\le j\le r}\bigcup_{0<\tau<1}\{|\widetilde{\Delta Y}^{(j+1)}_\tau| = |\widetilde{\Delta Y}^{(j)}_\tau|\}\right) = 0, \tag{18}$$

where $\widetilde{\Delta Y}^{(j)}_\tau$ denotes the jth largest modulus jump of $\mathcal{Y}$ up to time τ.

To simplify notation, for the remainder of this proof, write Δ_t for the modulus jumps $|\Delta Y_t|$, and for their ordered values in the intervals $[0,t]$ or $[0,t)$, write $\Delta^{(j)}_t = |\widetilde{\Delta Y}^{(j)}_t|$ or $\Delta^{(j)}_{t-} = |\widetilde{\Delta Y}^{(j)}_{t-}|$, $t > 0$, $j = 1, 2, \dots$. We aim to show that

$$\mathbb{P}\left(\bigcup_{0<\tau<1}\{\Delta^{(j+1)}_\tau = \Delta^{(j)}_\tau\}\right) = 0, \qquad j = 1, 2, \dots, r, \tag{19}$$

from which (18) will follow immediately.

We consider first the first case $j = 1$. Define a sequence of random times $(\tau_k)_{k\ge0}$ by

$$\tau_0 = 1 \quad\text{and}\quad \tau_{k+1} := \inf\{0 < t < \tau_k\colon \Delta_t = \Delta^{(1)}_{\tau_k-}\}, \qquad k = 0, 1, \dots. \tag{20}$$

Since $\lim_{t\downarrow0}\Delta^{(1)}_t = 0$ a.s., we have $0 < \tau_{k+1} < \tau_k \le 1$ and $\lim_{k\to\infty}\tau_k = 0$ a.s. On $\{\tau_{k+1} \le t < \tau_k\}$, we have $\Delta^{(1)}_t = \Delta^{(1)}_{\tau_k-}$, hence, on the event $\{\Delta^{(2)}_t = \Delta^{(1)}_t\}$,

$$1 = \frac{\Delta^{(2)}_t}{\Delta^{(1)}_t} \le \frac{\Delta^{(2)}_{\tau_k-}}{\Delta^{(1)}_{\tau_k-}} \le 1.$$

This implies that

$$\begin{aligned}\bigcup_{0<\tau<1}\{\Delta^{(2)}_\tau = \Delta^{(1)}_\tau\} &= \bigcup_{k\ge0}\bigcup_{t\in[\tau_{k+1},\tau_k)}\{\Delta^{(2)}_t = \Delta^{(1)}_t\}\\ &= \bigcup_{k\ge0}\{\Delta^{(2)}_{\tau_k-} = \Delta^{(1)}_{\tau_k-}\}\\ &= \{\Delta^{(2)}_{t-} = \Delta^{(1)}_{t-} \text{ for some } t \le 1 \text{ with } \Delta_t > \Delta^{(1)}_{t-}\}\\ &=: E.\end{aligned}$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

Define $\mathbb{S} = \sum_{0<t<1} \delta_{(t,\Delta_t)}$, where $\delta_{(t,\Delta_t)}$ denotes a point mass at (t, Δ_t). Also, $\mathbb{S}$ is a Poisson random measure on $(0, 1) \times (0, \infty)$ with intensity $\mathrm{d}t \times \Pi_Y(\mathrm{d}x)$. Let

$$N = \int_{(0,1)\times(0,\infty)} \mathbf{1}_{\{\Delta_{t-}^{(2)}=\Delta_{t-}^{(1)}<x\}} \mathbb{S}(\mathrm{d}t \times \mathrm{d}x)$$

be the number of points (t, Δ_t) which satisfy $\Delta_{t-}^{(2)} = \Delta_{t-}^{(1)} < \Delta_t$ with $t < 1$. Then, recalling that $\Delta_t^{(j)} = |\widetilde{\Delta Y}_t^{(j)}|$, event E has probability

$$\begin{aligned}
\mathbb{P}(E) &= \mathbb{P}(N \geq 1) \\
&\leq \mathbb{E}(N) \quad \text{(by Markov's inequality)} \\
&= \int_0^1 \mathrm{d}t \int_{x>0} \mathbb{E}\,\mathbf{1}_{\{\Delta_{t-}^{(2)}=\Delta_{t-}^{(1)}<x\}} \Pi_Y(\mathrm{d}x) \\
&= \int_0^1 \mathrm{d}t \int_{x>0} \int_{\overline{\Pi}_Y^{\leftarrow}(y/t)<x} \mathbb{P}\Big(y + \mathscr{E}_2 \leq t\overline{\Pi}_Y\Big(\overline{\Pi}_Y^{\leftarrow}\Big(\frac{y}{t}\Big)-\Big)\Big)\mathbb{P}(\mathscr{E}_1 \in \mathrm{d}y)\Pi_Y(\mathrm{d}x) \\
&= \int_0^1 \mathrm{d}t \int_{x>0} \int_{\overline{\Pi}_Y^{\leftarrow}(y)<x} \mathbb{P}(\mathscr{E}_2 \leq t(\overline{\Pi}_Y(\overline{\Pi}_Y^{\leftarrow}(y)-) - y))\mathrm{e}^{-ty}\,\mathrm{d}y\Pi_Y(\mathrm{d}x) \\
&= 0.
\end{aligned} \tag{21}$$

In the second equality we used the compensation formula, and in the third we used a version of (16) appropriate to the $|\widetilde{\Delta Y}_t^{(j)}|$. The last expression in (21) is 0 because $\overline{\Pi}_Y$ is diffuse, so $\overline{\Pi}_Y(\overline{\Pi}_Y^{\leftarrow}(v)) = v = \overline{\Pi}_Y(\overline{\Pi}_Y^{\leftarrow}(v)-)$ for all $v > 0$. This means, with probability 1, that there are no tied values among the largest jumps in $(\Delta_\tau)_{0<\tau<t}$ for all $t \in (0, 1)$. (Note that this is ostensibly a much stronger statement than requiring there be no tied values among the largest jumps up until a fixed time t.)

Next we consider $j = 2$. It is enough to show that $\mathbb{P}(\bigcup_{0<t<1}\{\Delta_t^{(3)} = \Delta_t^{(2)}\}) = 0$. We restrict ourselves to the event $\mathscr{F} := \bigcap_{0<t<1}\{\Delta_t^{(2)} \neq \Delta_t^{(1)}\}$, which we have proved has probability 1. On this event, there are no ties for the largest value among $(\Delta_t)_{0<t\leq 1}$. Recall the definition of the sequence (τ_k) in (20). The largest jump $\Delta_t^{(1)}$ remains constant on the interval $\tau_{k+1} \leq t < \tau_k$. We aim to subdivide the interval $[\tau_{k+1}, \tau_k)$ so that the second largest jump up to time t, which is strictly less than $\Delta_{\tau_{k+1}}$, is constant within that subinterval. First, we consider the case when $\Delta_t^{(2)} = \Delta_{\tau_{k+1}}$. Define, for each $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$,

$$s_k := \sup\{0 < t < \tau_k : \Delta_{\tau_{k+2}} = \Delta_t^{(2)}\}.$$

Note that $s_k \geq \tau_{k+1}$ as $\Delta_{\tau_{k+2}} = \Delta_{\tau_{k+1}}^{(2)}$. Next, define a further sequence $(\sigma_m(k))_{m\geq 1}$ in $[s_k, \tau_k)$ such that $\sigma_0(k) = \tau_k$ and, for $m = 1, 2, \ldots,$

$$\sigma_m(k) = \inf\{0 < t < \sigma_{m-1}(k) : \Delta_t = \Delta_{\sigma_{m-1}(k)-}^{(2)}\} \vee s_k.$$

Then we can decompose

$$\bigcup_{0<t<1} \{\Delta_t^{(3)} = \Delta_t^{(2)}\} = \bigcup_{k\geq 0}\Big(\bigcup_{\tau_{k+1}\leq t<s_k} \cup \bigcup_{m\geq 1} \bigcup_{\sigma_m(k)\leq t<\sigma_{m-1}(k)}\Big)\{\Delta_t^{(3)} = \Delta_t^{(2)}\}. \tag{22}$$

https://doi.org/10.1017/jpr.2017.39 Published online by Cambridge University Press

When $\Delta_t^{(3)} = \Delta_t^{(2)}$ and $\{\tau_{k+1} \le t < s_k\}$, we have $\Delta_t^{(2)} = \Delta_{s_k-}^{(2)} = \Delta_{\tau_{k+2}}$, hence,

$$1 = \frac{\Delta_t^{(3)}}{\Delta_t^{(2)}} \le \frac{\Delta_{s_k-}^{(3)}}{\Delta_{s_k-}^{(2)}} \le 1.$$

When $\Delta_t^{(3)} = \Delta_t^{(2)}$ and $\{\sigma_{m+1}(k) \le t < \sigma_m(k)\}$, $m \ge 1$, $k \in \mathbb{N}_0$, we have $\Delta_t^{(2)} = \Delta_{\sigma_m(k)-}^{(2)}$, hence,

$$1 = \frac{\Delta_t^{(3)}}{\Delta_t^{(2)}} \le \frac{\Delta_{\sigma_m(k)-}^{(3)}}{\Delta_{\sigma_m(k)-}^{(2)}} \le 1.$$

So the events on the right-hand side of (22) are subsets of

$$\begin{aligned} &\bigcup_{k\ge 0}\{\Delta_{t-}^{(3)} = \Delta_{t-}^{(2)} \text{ for some } t \in [\tau_{k+1}, \tau_k) \text{ with } \Delta_t > \Delta_{t-}^{(2)}\} \\ &\quad = \bigcup_{k\ge 0}\ \bigcup_{t\in[\tau_{k+1},\tau_k)} \{\Delta_{t-}^{(3)} = \Delta_{t-}^{(2)},\ \Delta_t > \Delta_{t-}^{(2)}\} \\ &\quad = \{\Delta_{t-}^{(3)} = \Delta_{t-}^{(2)} \text{ for some } t < 1 \text{ with } \Delta_t > \Delta_{t-}^{(2)}\}. \end{aligned} \tag{23}$$

The probability of the event on the right-hand side of (23) can be computed to be 0 in a similar way as in (21). Hence, reverting to the original notation, we have

$$\mathbb{P}\bigg(\bigcup_{0<t<1}\{|\widetilde{\Delta Y}_t^{(3)}| = |\widetilde{\Delta Y}_t^{(2)}|\}\bigg) \le \mathbb{P}\bigg(\bigcup_{0<t<1}\{|\widetilde{\Delta Y}_t^{(3)}| = |\widetilde{\Delta Y}_t^{(2)}|\}, \mathcal{F}\bigg) + \mathbb{P}(\mathcal{F}^c) = 0.$$

For $j \ge 3$, we can proceed iteratively with similar arguments to arrive at (19), hence, (18). This completes the proof of (13). □

6. Applications to reinsurance

Many examples can be generated from the convergences in (11)–(13) using the continuous mapping theorem. Here we mention one that is of particular interest in reinsurance. The idea is that the largest claim up to a specified time incurred by an insurance company (the 'cedant') is referred to a higher level insurer (the 'reinsurer'). See Fan *et al.* [5] for details and references to the applications literature. This is known as the largest claim reinsurance (LCR) treaty: having set a fixed follow-up time t, we delete from the process the largest claim occurring up to and including that time. We refer the reader to [7] and [13] for more detailed expositions.

The LCR procedure can be made prospective by implementing it as a forward looking dynamic procedure in real time, from the cedant's point of view. Designate as time 0 the time at which the reinsurance is taken out. The first claim arriving after time 0 is referred to the reinsurer and not debited to the cedant. Subsequent claims smaller than the initial claim are paid by the cedant until a claim larger than the first (the previous largest) arrives. The difference between these two claims is referred to the reinsurer and not debited to the cedant. The process continues in this way so that at time t, the accumulated amount referred to the reinsurer is equal to the largest claim up to that time. This procedure has the same effect as applying the 'trim as you go' operator to the risk process. (It is also possible to apply 'lookback' trimming to a reinsurance model in a natural way.)

A primary quantity of interest is the ruin time, at which the process $(X_t)_{t\ge 0}$ describing the claims incoming to the company reaches a high level, $u > 0$. After reinsurance of the r highest claims, the process is reduced to $({}^{(r)}X_t)_{t\geq 0}$, with ruin time $T^{(r)}(u) = \inf\{t > 0 \colon {}^{(r)}X_t > u\}$. Supposing that X is Lévy with heavy-tailed canonical measure Π_X, not uncommon assumptions in the modern insurance literature, we assume that (9) holds with no centering necessary, and from the continuous mapping theorem immediately deduce an asymptotic distribution for $\sup_{0<\tau\leq 1} {}^{(r)}X_{t\tau}/b_t$ as $t \to \infty$, and, hence, for $T^{(r)}(\cdot)$, for high levels. Specifically, if (12) holds with $t \to \infty$ and $s = 0$ then

$$\lim_{t\to\infty} \mathbb{P}(T^{(r)}(ub_t) > t) = \mathbb{P}(T_Y^{(r)} > u), \qquad u > 0,$$

where $T_Y^{(r)} = \inf\{t > 0 \colon {}^{(r)}Y_t > 1\}$.

Acknowledgements

We are grateful to Professor Jean Jacod whose suggestion to Y. F. Ipsen at the Building Bridges conference in Braunschweig 2013 started us on this pathwise analysis. B. Buchmann and R. Maller's research was partially funded by the Australian Research Council (ARC) (grant numbers DP1092502 and DP160104737). Y. Ipsen was formerly at the ARC Centre of Excellence for Mathematical and Statistical Frontiers, School of Mathematics and Statistics, University of Melbourne. She acknowledges support from the ARC. We are grateful also to two anonymous referees whose careful readings resulted in significant improvements to the paper.

References

[1] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd edn. John Wiley, New York.
[2] Buchmann, B., Fan, Y. and Maller, R. A. (2016). Distributional representations and dominance of a Lévy process over its maximal jump processes. *Bernoulli* **22,** 2325–2371.
[3] Fan, Y. (2015). Convergence of trimmed Lévy process to trimmed stable random variables at 0. *Stoch. Process. Appl.* **125,** 3691–3724.
[4] Fan, Y. (2017). Tightness and convergence of trimmed Lévy processes to normality at small times. *J. Theoret. Prob.* **30,** 675–699.
[5] Fan, Y. *et al.* (2017). The effects of largest claim and excess of loss reinsurance on a company's ruin time and valuation. *Risks* **5,** 27pp. Available at http://www.mdpi.com/2227-9091/5/1/3.
[6] Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin.
[7] Ladoucette, S. A. and Teugels, J. L. (2006). Reinsurance of large claims. *J. Comput. Appl. Math.* **186,** 163–190.
[8] Maller, R. A. and Fan, Y. (2015). Thin and thick strip passage times for Lévy flights and Lévy processes. Preprint. Available at https://arxiv.org/abs/1504.06400.
[9] Maller, R. 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.
[10] Maller, R. A. and Schmidli, P. C. (2017). Small-time almost-sure behaviour of extremal processes. *Adv. Appl. Prob.* **49,** 411–429.
[11] Silvestrov, D. S. and Teugels, J. L. (2004). Limit theorems for mixed max-sum processes with renewal stopping. *Ann. Appl. Prob.* **14,** 1838–1868.
[12] Skorokhod, A. V. (1956). Limit theorems for stochastic processes. *Theory Prob. Appl.* **1,** 261–290.
[13] Teugels, J. L. (2003). Reinsurance actuarial aspects. Res. Rep. 2003-006, Eindhoven University of Technology.
[14] Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer, New York.