TAILS OF SOLUTIONS OF CERTAIN NONLINEAR STOCHASTIC DIFFERENTIAL EQUATIONS DRIVEN BY HEAVY TAILED LÉVY MOTIONS

GENNADY SAMORODNITSKY AND MIRCEA GRIGORIU

ABSTRACT. We describe the exact tail behavior of the solutions to certain nonlinear stochastic differential equations driven by Lévy motions with regularly varying tails and establish existence and uniqueness of solutions to these equations.

1. Introduction and preliminaries

In this paper we study solutions to "mean reverting" stochastic differential equations of the form

$$dX(t) = -f(X(t))\,dt + dL(t)\,,$$

where f is a quickly increasing to infinity function, and $(L(t),\, t \geq 0)$ is a Lévy motion with regularly varying tails. Even though establishing existence of solutions of such equations is reasonably straightforward, given the developed theory of semimartingale stochastic calculus, not much is, in general, known about the probabilistic properties of the solutions. The particular equation we are considering comes from applications in mechanics, but similar models have been considered for other applications, for example for storage processes. In these applications the Lévy process L is usually assumed to be a subordinator (i.e. to have nondecreasing sample paths), and the function f is sometimes called "the release rule". See, for example, Brockwell et al. (1982) and Asmussen (1998).

Our main goal in this paper is to establish, under appropriate assumptions on the function f and Lévy process L, the exact rate of decay of the tail probabilities of the random variables $X(t)$, $t > 0$. We establish, furthermore, existence of a unique stationary solution of the stochastic differential equation and obtain its tail behavior as well.

Key words and phrases. stochastic differential equation, Lévy motion, heavy tails, Markov process, stationary distribution, coupling, storage processes .

AMS 2000 subject classification. Primary 60H10. Secondary 60G51, 60E07.

Samorodnitsky's research was partially supported by NSF grants DMS-0071073 and DMI-9713549 at Cornell University. Grigoriu's research was partially supported by NSF grant CMS-9912524 at Cornell University.

First some preliminaries. Throughout this paper $(\Omega, \mathcal{F}, P)$ is a probability space and $(\mathcal{F}_t,\, t \geq 0)$ a filtration on that space. We will always assume, without further special notice, that the filtration is complete and right continuous (the " usual hypothesis"; see Protter (1991)).

We start with a brief discussion of the basic properties of Lévy motions we will need in the sequel. All the facts below are well known, and can be found in many sources. Two thorough recent references are Bertoin (1996) and Sato (1999). Recall that a Lévy motion $(L(t),\, t \geq 0)$ is a continuous in probability stochastic process adapted to the filtration $(\mathcal{F}_t,\, t \geq 0)$ with $L(0) = 0$ and stationary increments, such that for every $0 \leq s < t$ the increment $L(t) - L(s)$ is independent of $\mathcal{F}_s$. A Lévy motion has a version with sample paths in the space $D[0, \infty)$ of right continuous functions with left limits, and we will always assume that we are dealing with such a version. The law of a Lévy process is completely characterized by its one-dimensional distribution at time 1 (say) and

$$Ee^{i\theta L(1)} = \exp\left\{ \int_{-\infty}^{\infty} \left(e^{i\theta x} - 1 - i\theta x \mathbf{1}(|x| \leq 1) \right) \nu(dx) + i\theta a \right\}, \quad \theta \in \mathbb{R}, \tag{1.1}$$

with ν a σ-finite measure that does not charge the origin, such that $\int_{\mathbb{R}} \min(x^2, 1)\, \nu(dx) < \infty$ and $a \in \mathbb{R}$. The measure ν is the Lévy measure of the process. If the Lévy measure is infinite then, on an event of probability 1, the Lévy process has a dense set of discontinuities in every interval of positive length. On the other hand, if the Lévy measure ν is finite, then the jumps of the process L form a homogeneous Poisson process on the positive half line with intensity $\nu(\mathbb{R})$. The jump sizes are iid random variables independent of the jump times with common distribution $\nu / \nu(\mathbb{R})$. Between every two jumps the process is linear with the slope $a - \int_{|x| \leq 1} x\, \nu(dx)$. No matter whether the Lévy measure is finite or not, the Lévy process L is continuous with probability 1 at any fixed point.

A Lévy process is symmetric (i.e. L and $-L$ have the same finite dimensional distributions) if and only if in (1.1) $a = 0$ and the Lévy measure is symmetric. In that case the characteristic function (1.1) can be written in a simpler form

$$Ee^{i\theta L(1)} = \exp\left\{ \int_{-\infty}^{\infty} \left(e^{i\theta x} - 1 \right) \nu(dx) \right\}, \quad \theta \in \mathbb{R}. \tag{1.2}$$

If the Lévy measure ν satisfies a stronger condition $\int_{\mathbb{R}} \min(|x|, 1)\, \nu(dx) < \infty$ (in particular, if ν is finite) then one can write the characteristic function (1.1) in the form

$$Ee^{i\theta L(1)} = \exp\left\{ \int_{-\infty}^{\infty} \left(e^{i\theta x} - 1 \right) \nu(dx) + i\theta b \right\}, \quad \theta \in \mathbb{R} \tag{1.3}$$

with $b \in \mathbb{R}$. If, in addition, ν is concentrated on $(0, \infty)$ and $b \geq 0$ in (1.3), then L is a nondecreasing process (a subordinator). Finally, if $b = 0$ and $\int_{\mathbb{R}} x\, \nu(dx) = 0$ then the Lévy process is zero mean, and then its characteristic function can be written in the form (1.2), except that ν no longer has to be symmetric.

An important fact is that, in certain cases, the tails of the one dimensional distributions of the Lévy process and its Lévy measure are equivalent. Specifically, the equivalence

$$P(L(t) > u) \sim t\nu\Big((u,\infty)\Big) \quad \text{as } u \to \infty \tag{1.4}$$

for any $t > 0$ holds under the assumption of *subexponentiality* on the tail of either $L(1)$ or that of ν; see Embrechts et al. (1979). We will, actually, assume that the tails are regularly varying:

$$\nu\Big((u,\infty)\Big) = u^{-\alpha} l(u) \tag{1.5}$$

for some $\alpha > 0$, where l is a slowly varying at infinity function. Since regularly varying tails are subexponential, (1.4) holds in this case. An important example of Lévy motions satisfying (1.5) is that of α-stable motions in which case

$$\nu(dx) = \begin{cases} c_+ x^{-(\alpha+1)}\, dx & \text{if } x > 0 \\ c_- |x|^{-(\alpha+1)}\, dx & \text{if } x < 0 \end{cases} \tag{1.6}$$

for $0 < \alpha < 2$ and $c_+, c_- \geq 0$. A source of information on α-stable processes, of which α-stable motions is an example, is in Samorodnitsky and Taqqu (1994).

The rest of this paper is organized as follows. In the next section we discuss in details the stochastic differential equation we are considering, and state all the assumptions. The tails behavior of the solution $X(t)$ at a finite time $t > 0$ is establish in Section 3. Existence and uniqueness of the stationary solution to our stochastic differential equation is established in Section 4. An easy corollary here is the tail behavior of the stationary solution. We note that certain results on existence of a stationary solution are also given in Brockwell et al. (1982), and that the paper Asmussen (1998) provides certain information on the tail behavior of the stationary solution. We discuss the relationship between our results and their results in Section 4. Section 5 touches on tail estimation in the context of continuous time processes, like solutions to stochastic differential equations, and warns about the possible pitfalls of applying certain heuristic arguments to solve stochastic differential equations of this kind. Finally, Section 6 contains certain technical results used in the proofs elsewhere in the paper.

2. The stochastic differential equation

We now describe precisely the stochastic differential equation we will study. This is an equation of the form

$$dX(t) = -f(X(t))\, dt + dL(t)\,, \tag{2.1}$$

where $(L(t),\, t \geq 0)$ is a Lévy motion with Lévy measure ν satisfying the regular variation assumption (1.5). In the main result of the next section, that describes the tail behaviour of the solution to the above equation at any fixed positive time t we will also assume that the Lévy motion is symmetric. This assumption is not needed for the arguments used in the present section. The following assumptions are imposed on the function f.

$$f \text{ is Lipschitz on compact intervals} \tag{2.2}$$

$$f(0) = 0, \text{ and } f \text{ is nondecreasing.} \tag{2.3}$$

$$f \text{ is regularly varying at infinity with exponent } \beta > 1, \tag{2.4}$$

and for some constants $A \in (0, \infty)$ and $\beta_1 > 1$,

$$-f(-x) \geq A x^{-\beta_1} \quad \text{for all } x \geq 1. \tag{2.5}$$

Note that a Lévy motion is a semimartingale and, hence, the standard theory of stochastic integration applies to stochastic differential equations with respect to Lévy motions. Our reference on stochastic integration is Protter (1991).

We have to start with the basic analysis of the equation (2.1); it only uses the local Lipschitz assumption (2.2), and the fact that $xf(x) \geq 0$ for all x. The argument is standard. For $b > 0$ let

$$[x]_b = \begin{cases} x & \text{if } -b \leq x \leq b \\ b & \text{if } x > b \\ -b & \text{if } x < -b \end{cases}$$

and consider the stochastic differential equation

$$dX^{(b)}(t) = -f([X^{(b)}(t)]_b)\, dt + dL(t)\,. \tag{2.6}$$

Since the function $f([\cdot]_b)$ is Lipschitz, we conclude by Theorem V.3.6 in Protter (1991) that for any $\mathcal{F}_0$-measurable $X^{(b)}(0)$ the equation (2.6) has a strongly unique solution, which is a semimartingale. Furthermore, this solution is strongly Markov by Theorem V.6.34 in Protter (1991).

Consider the family of stopping times

$$T_b = \inf\Big\{ t \geq 0 : \, |X^{(b)}(t)| \geq b \Big\}, \quad b > 0\,.$$

Then $T_b > 0$ a.s. on the event $\{|X^{(b)}(0)| < b\}$ and for every $0 < b_1 \leq b_2 < \infty$, $T_{b_1} \leq T_{b_2}$ a.s.. We claim that

$$\lim_{\substack{b \to \infty \\ b \text{ rational}}} T_b = \infty \quad \text{a.s..} \tag{2.7}$$

Once (2.7) has been established, it will follow from Theorem V.7.38 in Protter (1991) that for any $\mathcal{F}_0$-measurable $X(0)$ the equation (2.1) has a strongly unique solution $(X(t),\, t \geq 0)$, which is, then, automatically a semimartingale. Moreover, let T be a stopping time, $s > 0$ and h a bounded nonnegative measurable function. For $b > 0$ let $(X^{(b+1)}(t),\, t \geq 0)$ be the solution to (2.6) with $X^{(b+1)}(0) = X(0)$. Since on the event $\{|X(t)| \leq b,\, 0 \leq t \leq T+s\} = \{|X^{(b+1)}(t)| \leq b,\, 0 \leq t \leq T+s\}$ we have $X(t) = X^{(b+1)}(t)$ for $0 \leq t \leq T+s$, we have by the strong Markov property of $(X^{(b+1)}(t),\, t \geq 0)$, for every $A \in \mathcal{F}_T$

$$\int_A \mathbf{1}\Big(|X(t)| \leq b,\, 0 \leq t \leq T+s\Big) f\Big(X(T+s)\Big)\, dP$$

$$= \int_A \mathbf{1}\Big(|X^{(b+1)}(t)| \leq b,\, 0 \leq t \leq T\Big) f\Big(X^{(b+1)}(T+s)\Big) \mathbf{1}\Big(|X^{(b+1)}(T+u)| \leq b,\, 0 \leq u \leq s\Big)\, dP$$

$$= \int_A \mathbf{1}\Big(|X^{(b+1)}(t)| \leq b,\, 0 \leq t \leq T\Big) E_{X^{(b+1)}(T)} \Big(f\Big(X^{(b+1)}(s)\Big) \mathbf{1}\Big(|X^{(b+1)}(u)| \leq b,\, 0 \leq u \leq s\Big)\Big)\, dP$$

$$= \int_A \mathbf{1}\Big(|X(t)| \leq b,\, 0 \leq t \leq T\Big) E_{X(T)} \Big(f\Big(X(s)\Big) \mathbf{1}\Big(|X(u)| \leq b,\, 0 \leq u \leq s\Big)\Big)\, dP\,,$$

and letting $b \to \infty$ and using the monotone convergence theorem we see that

$$\int_A f\Big(X(T+s)\Big)\, dP = \int_A E_{X(T)} f\Big(X(s)\Big)\, dP\,,$$

hence $(X(t),\, t \geq 0)$ is strongly Markov as well.

It remains to check (2.7). Suppose that, to the contrary, there is $a \in (0,\infty)$ such that

$$P(B_1) := P\left(\lim_{\substack{b \to \infty \\ b \text{ rational}}} T_b \leq a \right) = \delta > 0\,.$$

Choose a $b \in (0,\infty)$ such that

$$P(B_2) := P\left(\sup_{0 \leq t \leq a} L(t) \leq b \right) > 1 - \delta\,.$$

Then the intersection of B_1 and B_2 is not empty. Fix $\omega \in B_1 \cap B_2$, and let for $n \geq 0$

$$S_n = \sup\Big\{ t \geq [T_{2^n}, T_{2^{n+1}}) :\, \mathrm{sign}(X^{(2^{n+1})}(t)) \neq \mathrm{sign}(X^{(2^{n+1})}(T_{2^{n+1}})) \Big\}\,.$$

Recall that

$$X^{(2^{n+1})}(T_{2^{n+1}}) - X^{(2^{n+1})}(S_n) = -\int_{S_n}^{T_{2^n+1}} f\Big([X^{(2^{n+1})}(t)]_{2^{n+1}}\Big)\, dt + L(T_{2^{n+1}}) - L(2^{n+1}(S_n))\,.$$

Now, $|X^{(2^{n+1})}(T_{2^{n+1}})| \geq 2^{n+1}$. Further, either $S_n = T_{2^n}$, in which case

$$|X^{(2^{n+1})}(S_n)| = |X^{(2^{n+1})}(T_{2^n})| = |X^{(2^n)}(T_{2^n})| \leq 2^n + 2b\,,$$

since by the choice of ω the Lévy process cannot have a jump of magnitude greater than $2b$ at the time T_{2^n}. For the same reason, if $S_n > T_{2^n}$, it must be that $|X^{(2^{n+1})}(S_n)| \leq 2b$. In any case, for all n large enough $|X^{(2^{n+1})}(T_{2^{n+1}}) - X^{(2^{n+1})}(S_n)| \geq 2^n - 2b \geq 2^{n-1}$. The assumption that $xf(x) \geq 0$ for all x ensures that the sign of $\int_{S_n}^{T_{2^{n+1}}} f\left([X^{(2^{n+1})}(t)]_{2^{n+1}}\right) dt$ is the same as that of $X^{(2^{n+1})}(T_{2^{n+1}}) - X^{(2^{n+1})}(S_n)$. Therefore, for all n large enough $|L(T_{2^{n+1}}) - L(2^{n+1}(S_n))| \geq 2^{n-1} > 2b$ contradicting the choice of ω. This proves (2.7).

The Markov property of the solution to our stochastic differential equation allows us, in particular, to use the usual Markovian notation P_x when we want to emphasize that we are working with a solution to that equation with $X(0) = x$. We will use this notation throughout the paper without further comments. We also note at this point that it is an immediate application of Theorem Theorem 5.4 in Kurtz and Protter (1991) that the Markov process $(X(t), t \geq 0)$ is a *Feller* process. That is, for any bounded and continuous function f on the real line, the function $y \to E_y f(X(t))$ is continuous for every $t \geq 0$.

Even though the equation (2.1) has a “nice” solution, direct understanding of many properties of this solution is not easy. For this reason our approach is to approximate that solution by “throwing away” the small jumps of the Lévy process L. Specifically, given a Lévy process satisfying (1.1) and a number $\sigma > 0$ we consider a Lévy motion L_σ satisfying

$$Ee^{i\theta L_\sigma(1)} = \exp\left\{\int_{|x|>\sigma} \left(e^{i\theta x} - 1 - i\theta x \mathbf{1}(|x| \leq 1)\right) \nu(dx) + i\theta a\right\} \tag{2.8}$$

$$= \exp\left\{\int_{-\infty}^{\infty} \left(e^{i\theta x} - 1 - i\theta x \mathbf{1}(|x| \leq 1)\right) \nu_\sigma(dx) + i\theta a\right\}, \quad \theta \in \mathbb{R},$$

where $\nu_\sigma(A) = \nu(A \cap \{|x| > \sigma\})$ for a Borel set A. Note that ν_σ is a finite measure, and we will consider the corresponding stochastic differential equation driven by L_σ

$$dX_\sigma(t) = -f(X_\sigma(t))\, dt + dL_\sigma(t)\,. \tag{2.9}$$

It is easy to construct the increments $L_\sigma(t) - L_\sigma(s)$ for $0 \leq s < t$ as measurable functions of $L(u) - L(v)$, $s \leq v < u \leq t$. Hence, one may assume that $(L_\sigma(t), t \geq 0)$ is a Lévy motion with respect to the same filtration $(\mathcal{F}_t, t \geq 0)$ as $(L(t), t \geq 0)$ is. For our purposes the specific filtration does not matter. What is important for our purposes that the solution to the equation (2.9) converges weakly to that of (2.1). More precisely, if $(X(t), t \geq 0)$ is the solution to (2.1) and for $\sigma > 0$, $(X_\sigma(t), t \geq 0)$ is the solution to (2.9), then $X_\sigma(0) \Rightarrow X(0)$ as $\sigma \to 0$ implies that $(X_\sigma(t), t \geq 0) \Rightarrow (X(t), t \geq 0)$ weakly in $D[0, \infty)$ as $\sigma \to 0$. See Theorem 5.4 in Kurtz and Protter (1991). Moreover, since the set of discontinuities of $(X(t), t \geq 0)$ coincides with that of $(L(t), t \geq 0)$, this means that the process $(X(t), t \geq 0)$ is a.s. continuous at every fixed t, and so $X_\sigma(t) \Rightarrow X(t)$ as $\sigma \to 0$ for every $t \geq 0$. See Theorem 12.5 in Billingsley (1999).

Of course, the point of switching from the equation (2.1) to the equation (2.9) is that Lévy process $(L_\sigma(t), t \geq 0)$ has finitely many jumps in any interval of a finite length, and is linear between two successive jumps. Thus, between any two successive jumps the process $(X_\sigma(t),\, t\geq 0)$ satisfies a deterministic differential equation, which can be explicitly solved. This allows us to get a good "handle" on the process $(X_\sigma(t),\, t\geq 0)$, and the weak convergence of the latter to $(X(t),\, t\geq 0)$ allows us to derive, thus, conclusions about the solution to the equation (2.1). It is clear that implementation of this approach will require us to obtain "uniform in σ" results for the solution of the easier equation (2.9), so that the results will be preserved under the weak limit.

3. The probability tail at a finite positive time

The main result of this section determines the tail behaviour of the solution to the stochastic differential equation (2.1) at any fixed positive time t. It turns out that, under our assumptions, this tail behaviour is independent of both t and of the initial value $X(0)$. Denote

$$h(u) = \int_u^\infty \frac{\nu\Big((y,\infty)\Big)}{f(y)}\, dy, \quad u\geq 0\,. \tag{3.1}$$

Note that by the assumptions (1.5) and (2.2) – (2.4) the function h is finite for large u and, moreover,

$$h \text{ is regularly varying at infinity with exponent } -(\alpha+\beta)+1\,. \tag{3.2}$$

Theorem 3.1. *Let $(X(t), t\geq 0)$ be the solution to the stochastic differential equation (2.1) with $X(0)=x$. We assume that the the Lévy motion $(L(t),\, t\geq 0)$ is symmetric and that the assumptions (1.5) and (2.2) – (2.5) hold. Then for every $t_0>0$*

$$\lim_{u\to\infty} \frac{P_x(X(t)>u)}{h(u)} = 1 \quad \text{uniformly in } x\in\mathbb{R} \text{ and } t\geq t_0. \tag{3.3}$$

Remark 3.2. Before proving the theorem, we would like to comment on the symmetry assumption on the Lévy motion in Theorem 3.1. It is very easy to see from the proof below that this assumption is redundant if the Lévy measure ν is finite (i.e. if the Lévy process L is a compound Poisson process), or if the Lévy process L is a subordinator. In the later case, in fact, one does not need any assumptions on the behavior of the function f for the negative values of the argument.

We conjecture that the statement of Theorem 3.1 remains valid without the assumption of symmetry of the Lévy process L in general. Technical issues in the proof have, however, forced us to make this assumption in the present paper.

Proof of Theorem 3.1. It is enough to prove the theorem for $t_0=1$. We will use the approximation described at the end of Section 2. The statement (3.3) (for $t_0=1$) will follow once we show that for every $\xi > 0$ there are $u(\xi) > 0$ and $\sigma(\xi) > 0$ such that

$$\left| \frac{P_x(X_\sigma(t) > u)}{h(u)} - 1 \right| \leq \xi \quad \text{for all } u \geq u(\xi),\, 0 < \sigma \leq \sigma(\xi) \text{ and all } x \in \mathbb{R},\, t \geq 1. \tag{3.4}$$

Here $(X_\sigma(t),\, t \geq 0)$ is the solution to the stochastic differential equation (2.9).

We start with introducing some notation. Let $0 < T_1 < \ldots < T_{N_1}$ be the times of the jumps of the Lévy process L_σ in $[0, t]$. Note that N_1 has the Poisson distribution with the mean $t\lambda_\sigma$, where

$$\lambda_\sigma = \nu_\sigma(\mathbb{R}) = \nu\Big(\{|x| > \sigma\}\Big)\,. \tag{3.5}$$

Let, further, W_i be the size of the jump at time T_i, $i = 1, 2, \ldots$. Note that $W_1, W_2, \ldots$ are iid random variables, independent of the jump times $T_1, T_2, \ldots$. Moreover, the common distribution of the jump sizes is

$$F_\sigma(A) = \frac{\nu(A \cap \{|x| > \sigma\})}{\lambda_\sigma}, \quad A \text{ a Borel set.} \tag{3.6}$$

Observe that, by the assumed symmetry of the Lévy process, F_σ is a symmetric distribution. Define also for $a > 0$

$$M(a) = \sup\{i \leq N_1 : \, |W_i| > a\} \quad (= 0 \text{ if empty set}) \tag{3.7}$$

and

$$N(a) = \sup\{i \leq N_1 : \, W_i + W_{i+1} + \ldots + W_{N_1} > a\} \quad (= 0 \text{ if empty set}). \tag{3.8}$$

Let g be the function in (6.1), and let $u_0 > 0$ be such that

$$g(u) \leq 1 \quad \text{for all } u \geq u_0/2. \tag{3.9}$$

Let further k be a positive integer satisfying

$$k > \frac{\alpha + \beta - 1}{\beta - 1}\,. \tag{3.10}$$

Let $0 < \epsilon < 1$, $0 < \gamma < 1$, $0 < \delta < 1$, $\epsilon_1, \ldots, \epsilon_{k-1}$, $\gamma_1, \ldots, \gamma_{k-1}$ and $\delta_1, \ldots, \delta_{k-1}$ be parameters satisfying the following restrictions.

$$\gamma < \frac{\delta}{1 + \delta}\,, \tag{3.11}$$

$$\epsilon < \frac{\gamma}{4}\,, \tag{3.12}$$

$$\epsilon + \gamma + \delta_1 < 1\,, \tag{3.13}$$

$$\epsilon_j + \gamma_j < \delta_j \quad \text{for } j = 1, \ldots, k-1\,, \tag{3.14}$$

$$\epsilon < \frac{\gamma_j}{4} \quad \text{for } j = 1, \ldots, k-1\,, \tag{3.15}$$

$$\epsilon + \gamma_j + \delta_j < \delta_{j-1} \quad \text{for } j = 2, \dots, k-1 . \tag{3.16}$$

Observe that there is a solution to the system of conditions (3.11) - (3.16) with arbitrarily small γ and δ_1, and with ϵ arbitrarily small in relation to $\gamma_* = \min(\gamma, \gamma_1, \dots, \gamma_{k-1})$. Indeed, let ζ be an arbitrary positive number. Let $\delta = 1$ (say), select sequentially $0 < \gamma < \min(\zeta, 1/3)$, $0 < \delta_1 < \min(\zeta, 1/3)$, $\delta_j = \delta_{j-1}/2$ for $j = 2, \dots, k-1$, $0 < \gamma_j < \delta_j$ and $0 < \epsilon_j < \delta_j - \gamma_j$ for $j = 1, \dots, k-1$, and, finally, $0 < \epsilon < \min\big(\zeta\gamma_*, \gamma/4, 1 - \gamma - \delta_1, \min(\gamma_1/4, \dots, \gamma_{k-1}/4, \delta_2 - \gamma_2, \dots, \delta_{k-1} - \gamma_{k-1})\big)$.

We have for $u \geq u_0$

$$P_x(X_\sigma(t) > u) = P_x\left(X_\sigma(t) > u,\, |W_i| \leq \epsilon u \ \text{ for all } T_i \in \left(t - g\left(\frac{u}{1+\delta}\right), t\right]\right) \tag{3.17}$$

$$+P_x\left(X_\sigma(t) > u,\, |W_i| > \epsilon u \ \text{ for some } T_i \in \left(t - g\left(\frac{u}{1+\delta}\right), t\right]\right) := p_1(u) + p_2(u)\,;$$

further we write

$$A_{\epsilon,\gamma,\delta;u} = \left\{ \sup_{\substack{1 \leq i \leq N_1 \\ T_i > t - g(u/(1+\delta))}} \sum_{j=i}^{N_1} W_j \mathbf{1}\big(|W_j| \leq \epsilon u\big) > \gamma u \right\} . \tag{3.18}$$

We have

$$p_1(u) \leq P\left(A_{\epsilon,\gamma,\delta;u}\right) + P_x\left(A^c_{\epsilon,\gamma,\delta;u} \cap \left\{X_\sigma(t) > u,\, T_{M(\epsilon u)} < t - g\left(\frac{u}{1+\delta}\right)\right\}\right) \tag{3.19}$$

$$:= P\left(A_{\epsilon,\gamma,\delta;u}\right) + p_{11}(u)\,.$$

Consider the event above whose probability is $p_{11}(u)$. Note that by (6.6), the system without any jumps in the interval $\big(t - g\big(u/(1+\delta)\big), t\big]$ will end up, at time t, not higher than at $u/(1+\delta)$. Hence, because of the restriction (3.11) we have by part (ii) of Lemma 6.1

$$p_{11}(u) = 0\,. \tag{3.20}$$

Let, further, I_u be the number of jumps of the process L_σ in the interval $\big(t - g\big(u/(1+\delta)\big), t\big]$. Note that I_u has the Poisson distribution with the mean $\lambda_\sigma g\big(u/(1+\delta)\big)$. By the symmetry of the random variables $W_1, W_2, \dots$ and by the independence of the jump sizes and the jump times,

$$P\left(A_{\epsilon,\gamma,\delta;u}\right) \leq 2P\left(\sum_{j=1}^{J_u} W_j \mathbf{1}\big(|W_j| \leq \epsilon u\big) > \gamma u\right) \tag{3.21}$$

$$= 2e^{-\lambda_\sigma g\left(\frac{u}{1+\delta}\right)} \sum_{n=1}^{\infty} \frac{\left(\lambda_\sigma g\left(\frac{u}{1+\delta}\right)\right)^n}{n!} P\left(\sum_{j=1}^{n} W_j \mathbf{1}\big(|W_j| \leq \epsilon u\big) > \gamma u\right) .$$

By Lemma 6.2 for any $n \geq 1$

$$P\left(\sum_{j=1}^{n} W_j \mathbf{1}\big(|W_j| \leq \epsilon u\big) > \gamma u\right) \leq \exp\left\{-\frac{\gamma}{2\epsilon} \operatorname{arsinh} \frac{\epsilon\gamma u^2}{2 \operatorname{var}\left(\sum_{j=1}^{n} W_j \mathbf{1}\big(|W_j| \leq \epsilon u\big)\right)}\right\} , \tag{3.22}$$

and so the next step is to estimate the variance in the right hand side above. By the symmetry and independence

$$\text{var}\left(\sum_{j=1}^{n} W_j \mathbf{1}\Big(|W_j| \le \epsilon u\Big)\right) = nE\left(W_1^2 \mathbf{1}\Big(|W_1| \le \epsilon u\Big)\right) \tag{3.23}$$

$$= \frac{2n}{\lambda_\sigma}\int_\sigma^{\epsilon u} x^2\,\nu(dx) \le \frac{2n}{\lambda_\sigma}\int_0^{\epsilon u} x^2\,\nu(dx)\,.$$

Suppose first that the tail index in (1.5) satisfies $0 < \alpha < 2$. Then by the Karamata theorem (see e.g. Resnick (1987), p.17),

$$\int_0^{\epsilon u} x^2\,\nu(dx) \sim \frac{\alpha}{2-\alpha}\epsilon^{2-\alpha}u^2\nu\Big((u,\infty)\Big) \quad \text{as } u \to \infty.$$

Therefore, for all $u > u_0(\epsilon)$ and all $n \ge 1$

$$P\left(\sum_{j=1}^{n} W_j \mathbf{1}\Big(|W_j| \le \epsilon u\Big) > \gamma u\right) \le \exp\left\{-\frac{\gamma}{2\epsilon}\,\text{arsinh}\,\frac{\epsilon^{\alpha-1}\gamma\lambda_\sigma}{Cn\nu\Big((u,\infty)\Big)}\right\}. \tag{3.24}$$

Here $u_0(\epsilon)$ is a finite positive constant depending only on ϵ and C is a finite positive constant. In general, we will use in the sequel the generic notation $C(\cdot)$ and $u_0(\cdot)$ to denote a finite positive constant that may depend on the parameters in $(\cdot)$. Moreover, the constants may change as we go along. By (6.10) there is $a > 0$ be such that $\text{arsinh}(x) \ge (\log x)/2$ for $x \ge a$. Then by (3.21) and (3.24)

$$P\left(A_{\epsilon,\gamma,\delta;u}\right) \le 2P\left(I_u > \frac{\epsilon^{\alpha-1}\gamma\lambda_\sigma}{C_1 a\nu\Big((u,\infty)\Big)}\right) \tag{3.25}$$

$$+C_2(\epsilon,\gamma)\lambda_\sigma^{-\gamma/4\epsilon}\left(\nu\Big((u,\infty)\Big)\right)^{\gamma/4\epsilon} E\left(I_u^{\gamma/4\epsilon}\right)\,,$$

where $C(\epsilon,\gamma)$ is a finite positive constant that may depend on ϵ and γ.

Using the assumption (3.12) and the bound (6.11) with $p = \gamma/4\epsilon$ we see that the second term in the right hand side of (3.25) can be bounded from above by

$$C(\epsilon,\gamma)\left(\nu\Big((u,\infty)\Big)\right)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right)$$

for $u \ge u_0(\epsilon)$ and $0 < \sigma \le \sigma_0$, some $\sigma_0 > 0$. In general, σ_0 will stand, in the sequel, for a (small) positive constant, that may change from time to time.

An exponential Markov inequality is enough to estimate the first term in the right hand side of (3.25). Letting

$$\theta = \log\left(\frac{\epsilon^{\alpha-1}\gamma}{g\left(\frac{u}{1+\delta}\right)\nu\Big((u,\infty)\Big)}\right) > 0$$

if $u \geq u_0(\epsilon, \gamma)$, we get

$$P\left(I_u > \frac{\epsilon^{\alpha-1}\gamma\lambda_\sigma}{Ca\nu\Big((u,\infty)\Big)}\right) \leq \exp\left\{-\theta\frac{\epsilon^{\alpha-1}\gamma\lambda_\sigma}{Ca\nu\Big((u,\infty)\Big)} + \left(e^\theta - 1\right)\lambda_\sigma g\left(\frac{u}{1+\delta}\right)\right\}$$

$$\leq \exp\left\{-C(\epsilon,\gamma)\theta\frac{\lambda_\sigma}{\nu\Big((u,\infty)\Big)}\right\} \leq C(\epsilon,\gamma)\left(\nu\Big((u,\infty)\Big)g\left(\frac{u}{1+\delta}\right)\right)^{C(\epsilon,\gamma)\frac{\lambda_\sigma}{\nu((u,\infty))}}$$

for $u \geq u_0(\epsilon, \gamma)$ and $0 < \sigma \leq \sigma_0$. Summarizing, we conclude that

$$P\left(A_{\epsilon,\gamma,\delta;u}\right) \leq C(\epsilon,\gamma)\left(\nu\Big((u,\infty)\Big)\right)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) \quad \text{if } 0<\alpha<2 \tag{3.26}$$

for all $u \geq u_0(\epsilon, \gamma)$ and $0 < \sigma \leq \sigma_0$.

On the other hand, if $\alpha > 2$, then $\int_0^{\epsilon u} x^2\,\nu(dx)$ converges, as $u \to \infty$, to a finite positive limit. Therefore, for all $u \geq u_0(\epsilon)$ and $n \geq 1$ we have

$$P\left(\sum_{j=1}^n W_j \mathbf{1}\Big(|W_j| \leq \epsilon u\Big) > \gamma u\right) \leq \exp\left\{-\frac{\gamma}{2\epsilon}\operatorname{arsinh}\frac{\epsilon\gamma u^2\lambda_\sigma}{Cn}\right\},$$

and the same argument as that leading to (3.26) shows that

$$P\left(A_{\epsilon,\gamma,\delta;u}\right) \leq C(\epsilon,\gamma)u^{-\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) \quad \text{if } \alpha>2 \tag{3.27}$$

for all $u \geq u_0(\epsilon, \gamma)$ and $0 < \sigma \leq \sigma_0$.

Finally, if $\alpha = 2$, then

$$\int_0^{\epsilon u} x^2\,\nu(dx) \sim L(u) \quad \text{for some slowly varying } L$$

as $u \to \infty$. This implies, as above, that for all $u \geq u_0(\epsilon)$ and $n \geq 1$ we have

$$P\left(\sum_{j=1}^n W_j \mathbf{1}\Big(|W_j| \leq \epsilon u\Big) > \gamma u\right) \leq \exp\left\{-\frac{\gamma}{2\epsilon}\operatorname{arsinh}\frac{\epsilon\gamma u^2\lambda_\sigma}{CnL(u)}\right\},$$

and

$$P\left(A_{\epsilon,\gamma,\delta;u}\right) \leq C(\epsilon,\gamma)u^{-\gamma/4\epsilon}L(u)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) \quad \text{if } \alpha=2 \tag{3.28}$$

for all $u \geq u_0(\epsilon, \gamma)$ and $0 < \sigma \leq \sigma_0$. Overall,

$$p_1(u) \leq \begin{cases} C(\epsilon,\gamma)\left(\nu\Big((u,\infty)\Big)\right)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } 0<\alpha<2 \\ C(\epsilon,\gamma)\,u^{-\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } \alpha>2 \\ C(\epsilon,\gamma)\,u^{-\gamma/4\epsilon}L(u)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } \alpha=2 \end{cases} \tag{3.29}$$

for all $u \geq u_0(\epsilon, \gamma)$ and $0 < \sigma \leq \sigma_0$.

We now switch to estimating the probability $p_2(u)$ in the right hand side of (3.17). Let k be a positive integer satisfying (3.10). Define recursively for $a>0$, $M_1(a)=M(a)$ (given by (3.7)) and for $j\geq 2$

$$M_j(a)=\sup\{i\leq M_{j-1}:\ |W_i|>a\}\quad (=0 \text{ if empty set})\,.$$

Write

$$p_2(u)\leq P\left(t-T_{M_1(\epsilon u)}\leq g\left(\frac{u}{1+\delta}\right),\, T_{M_{j-1}(\epsilon u)}-T_{M_j(\epsilon u)}\leq g(\epsilon_j u),\, j=2,\dots,k\right) \tag{3.30}$$
$$+P_x\left(X_\sigma(t)>u,\, t-T_{M_1(\epsilon u)}\leq g\left(\frac{u}{1+\delta}\right),\, T_{M_{j-1}(\epsilon u)}-T_{M_j(\epsilon u)}> g(\epsilon_j u),\quad \text{some } j=2,\dots,k\right)$$
$$:= p_{21}(u)+p_{22}(u)\,.$$

Here (ϵ_j) is a sequence of positive numbers that keep the system (3.11) - (3.16) feasible, and $T_0=0$. Obviously,

$$p_{21}(u)=\left(1-e^{-\nu\left((\epsilon u,\infty)\right)g\left(u/(1+\delta)\right)}\right)\prod_{j=2}^{k}\left(1-e^{-\nu\left((\epsilon u,\infty)\right)g\left(\epsilon_j u\right)}\right) \tag{3.31}$$
$$\leq\left(\nu\Big((\epsilon u,\infty)\Big)\right)^k g\left(\frac{u}{1+\delta}\right)\prod_{j=2}^{k} g\Big(\epsilon_j u\Big)\,.$$

Furthermore,

$$p_{22}(u)\leq P_x\left(X_\sigma(t)>u,\, t-T_{M_1(\epsilon u)}\leq g\left(\frac{u}{1+\delta}\right),\, X_\sigma\left(T_{M_1(\epsilon u)}-\right)\leq\delta_1 u\right)$$
$$+\sum_{j=2}^{k}P_x\Bigg(X_\sigma(t)>u,\, t-T_{M_1(\epsilon u)}\leq g\left(\frac{u}{1+\delta}\right),\, X_\sigma\left(T_{M_1(\epsilon u)}-\right)>\delta_1 u, \tag{3.32}$$
$$T_{M_{i-1}(\epsilon u)}-T_{M_i(\epsilon u)}\leq g(\epsilon_{j-1}u),\, i=2,\dots,j-1,\, T_{M_{j-1}(\epsilon u)}-T_{M_j(\epsilon u)}>g(\epsilon_{j-1}u)\Bigg)$$
$$:= p_{221}(u)+\sum_{j=2}^{k}p_{222}^{(j)}(u)\,.$$

Here (δ_j) is a sequence of positive numbers that keep the system (3.11) - (3.16) feasible. Let $A_{\epsilon,\gamma,\delta;u}$ be the event in (3.18). Clearly,

$$p_{221}(u)\leq P\Big(A_{\epsilon,\gamma,\delta;u}\Big)+P_x\left(A^c_{\epsilon,\gamma,\delta;u}\cap\left\{X_\sigma(t)>u,\, t-T_{M_1(\epsilon u)}\leq g\left(\frac{u}{1+\delta}\right),\right.\right. \tag{3.33}$$
$$\left.\left. X_\sigma\left(T_{M_1(\epsilon u)}-\right)\leq\delta_1 u\right\}\right):=P\Big(A_{\epsilon,\gamma,\delta;u}\Big)+p_{\text{main}}(u)\,.$$

Recall that an upper bound on $P\big(A_{\epsilon,\gamma,\delta;u}\big)$ is given by (3.26), (3.27) and (3.28). We proceed now to estimate $p_{\text{main}}(u)$ in (3.33). As implied by the notation, this is the main term. It follows from (6.6) and Lemma 6.1, part (ii), that for all $u \geq u_0$ on the event whose probability $p_{\text{main}}(u)$ measures, the magnitude $W_{M_1(\epsilon u)}$ of the jump at time $T_{M_1(\epsilon u)}$ has to satisfy

$$t - T_{M_1(\epsilon u)} + g\Big(\delta_1 u + W_{M_1(\epsilon u)}\Big) \leq g\Big((1-\gamma)u\Big),$$

hence

$$t - T_{M_1(\epsilon u)} \leq (1-\gamma)u \tag{3.34}$$

and

$$W_{M_1(\epsilon u)} \geq g^{-1}\left(g\Big((1-\gamma)u\Big) - \Big(t - T_{M_1(\epsilon u)}\Big)\right) - \delta_1 u\,. \tag{3.35}$$

In particular, g^{-1} is well defined on the range in question. Using the restriction (3.13) we conclude that

$$\begin{aligned} p_{\text{main}}(u) &\leq \int_0^{g\left((1-\gamma)u\right)} \nu\Big((\epsilon u,\infty)\Big) e^{-\nu\left((\epsilon u,\infty)\right)s} \frac{\nu\left(\left(g^{-1}\left(g\Big((1-\gamma)u\Big) - s\right) - \delta_1 u, \infty\right)\right)}{\nu\Big((\epsilon u,\infty)\Big)}\, ds \\ &\leq \int_0^{g\left((1-\gamma)u\right)} \nu\left(\left(g^{-1}\left(g\Big((1-\gamma)u\Big) - s\right) - \delta_1 u, \infty\right)\right) ds \\ &= \int_{(1-\gamma)u}^{\infty} \frac{\nu\Big((y-\delta_1 u,\infty)\Big)}{f(y)}\, dy \quad \leq \int_{(1-\gamma)u}^{\infty} \frac{\nu\left(\left(y\left(1-\frac{\delta_1}{1-\gamma}\right),\infty\right)\right),\infty)\Big)}{f(y)}\, dy \\ &\leq \left(1+\frac{\xi}{4}\right)\left(1-\frac{\delta_1}{1-\gamma}\right)^{-\alpha} \int_{(1-\gamma)u}^{\infty} \frac{\nu\left((y,\infty)\right),\infty)\Big)}{f(y)}\, dy \\ &= \left(1+\frac{\xi}{4}\right)\left(1-\frac{\delta_1}{1-\gamma}\right)^{-\alpha} h\Big((1-\gamma)u\Big) \end{aligned} \tag{3.36}$$

for all $u \geq u_0(\gamma,\delta_1,\xi)$, where $\xi > 0$ is a given number. Using (3.26), (3.27), (3.28) and (3.36) we conclude that

$$\begin{aligned} p_{221}(u) &\leq \left(1+\frac{\xi}{4}\right)\left(1-\frac{\delta_1}{1-\gamma}\right)^{-\alpha} h\Big((1-\gamma)u\Big) \\ &+ \begin{cases} C(\epsilon,\gamma)\left(\nu\Big((u,\infty)\Big)\right)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } 0<\alpha<2 \\ C(\epsilon,\gamma)\; u^{-\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } \alpha>2 \\ C(\epsilon,\gamma)\; u^{-\gamma/4\epsilon} L(u)^{\gamma/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } \alpha=2 \end{cases} \end{aligned} \tag{3.37}$$

for all $u \geq u_0(\gamma,\delta_1,\xi)$ and $0 < \sigma \leq \sigma_0$.

Next we consider the probabilities $p_{222}^{(j)}(u)$ in the right hand side of (3.32). Note that

$$p_{222}^{(2)}(u) \leq P\left(t - T_{M_1(\epsilon u)} \leq g\left(\frac{u}{1+\delta}\right)\right)$$

$$P_x\left(T_{M_1(\epsilon u)} - T_{M_2(\epsilon u)} > g(\epsilon_1 u),\, X_\sigma\left(T_{M_1(\epsilon u)}-\right) > \delta_1 u\right)$$

and now the same argument as that leading to (3.29) shows that for any $\gamma_1 > 0$ satisfying (3.14) and (3.15) (with $j = 1$) we have

$$P_x\left(T_{M_1(\epsilon u)} - T_{M_2(\epsilon u)} > g(\epsilon_1 u),\, X_\sigma\left(T_{M_1(\epsilon u)}-\right) > \delta_1 u\right)$$
$$\le \begin{cases} C(\epsilon,\epsilon_1,\gamma_1)\left(\nu\Big((\delta_1 u,\infty)\Big)\right)^{\gamma_1/4\epsilon} g(\epsilon_1 u) & \text{if } 0<\alpha<2 \\ C(\epsilon,\epsilon_1,\gamma_1,\delta_1)\, u^{-\gamma_1/4\epsilon} g(\epsilon_1 u) & \text{if } \alpha>2 \\ C(\epsilon,\epsilon_1,\gamma_1,\delta_1)\, u^{-\gamma_1/4\epsilon}(L(u))^{\gamma_1/4\epsilon} g(\epsilon_1 u) & \text{if } \alpha=2 \end{cases}$$

for $u \ge u_0(\epsilon,\epsilon_1,\gamma_1)$ and $0<\sigma\le\sigma_0$ and, hence,

$$p_{222}^{(2)}(u) \le \begin{cases} C(\epsilon,\epsilon_1,\gamma_1,\delta_1)\left(\nu\Big((u,\infty)\Big)\right)^{1+\gamma_1/4\epsilon} (g(u))^2 & \text{if } 0<\alpha<2 \\ C(\epsilon,\epsilon_1,\gamma_1,\delta_1)\, u^{-\gamma_1/4\epsilon}(g(u))^2 \nu\Big((u,\infty)\Big) & \text{if } \alpha>2 \\ C(\epsilon,\epsilon_1,\gamma_1,\delta_1)\, u^{-\gamma_1/4\epsilon}(L(u))^{\gamma_1/4\epsilon}(g(u))^2 \nu\Big((u,\infty)\Big) & \text{if } \alpha=2 \end{cases} \tag{3.38}$$

for $u \ge u_0(\epsilon,\epsilon_1,\gamma_1,\delta_1)$ and $0<\sigma\le\sigma_0$. Furthermore, for $j = 3,\dots,k$ we have

$$p_{222}^{(j)}(u) \le P_x\Bigg(t - T_{M_1(\epsilon u)} \le g\left(\frac{u}{1+\delta}\right),$$
$$T_{M_{j-1}(\epsilon u)} - T_{M_j(\epsilon u)} > g(\epsilon_{j-1}u),\, X_\sigma\left(T_{M_{j-1}(\epsilon u)}-\right) > \delta_{j-1}u\Bigg)$$
$$+\sum_{m=2}^{j-1} P_x\Bigg(X_\sigma(t)>u,\, t - T_{M_1(\epsilon u)} \le g\left(\frac{u}{1+\delta}\right),\, X_\sigma\left(T_{M_1(\epsilon u)}-\right) > \delta_1 u, \tag{3.39}$$
$$T_{M_{i-1}(\epsilon u)} - T_{M_i(\epsilon u)} \le g(\epsilon_{j-1}u),\, i=2,\dots,j-1,\, T_{M_{j-1}(\epsilon u)} - T_{M_j(\epsilon u)} > g(\epsilon_{j-1}u),$$
$$X_\sigma\left(T_{M_m(\epsilon u)}-\right) \le \delta_m u,\, X_\sigma\left(T_{M_{m-1}(\epsilon u)}-\right) > \delta_{m-1}u\Bigg) := p_{222}^{(j,1)}(u) + \sum_{m=2}^{j-1} p_{222}^{(j,m)}(u)\,.$$

Using, once again, the argument leading to (3.38) we see that for any $\gamma_{j-1} > 0$ satisfying the restrictions (3.14) and (3.15) we have

(3.40)
$$p_{222}^{(j,1)}(u) \le \begin{cases} C(\epsilon,\epsilon_{j-1},\gamma_{j-1},\delta_{j-1})\left(\nu\Big((u,\infty)\Big)\right)^{1+\gamma_{j-1}/4\epsilon} (g(u))^2 & \text{if } 0<\alpha<2 \\ C(\epsilon,\epsilon_{j-1},\gamma_{j-1},\delta_{j-1})\, u^{-\gamma_{j-1}/4\epsilon}(g(u))^2 \nu\Big((u,\infty)\Big) & \text{if } \alpha>2 \\ C(\epsilon,\epsilon_{j-1},\gamma_{j-1},\delta_{j-1})\, u^{-\gamma_{j-1}/4\epsilon}(L(u))^{\gamma_{j-1}/4\epsilon}(g(u))^2 \nu\Big((u,\infty)\Big) & \text{if } \alpha=2 \end{cases}$$

for $u \ge u_0(\epsilon,\epsilon_{j-1},\gamma_{j-1},\delta_{j-1})$, $0<\sigma\le\sigma_0$ and $j=3,\dots,k$.

On the other hand, for $j = 3,\dots,k$ and $m = 2,\dots,j-1$ we have

$$p_{222}^{(j,m)}(u) \le P\left(t - T_{M_1(\epsilon u)} \le g\left(\frac{u}{1+\delta}\right)\right)$$

$$P_x\left(T_{M_{m-1}(\epsilon u)} - T_{M_m(\epsilon u)} > g(\epsilon_{m-1}u),\, X_\sigma\left(T_{M_m(\epsilon u)}-\right) \le \delta_m u,\, X_\sigma\left(T_{M_{m-1}(\epsilon u)}-\right) > \delta_{m-1}u\right)$$

If $\gamma_m > 0$ satisfies (3.16), then the same argument as that leading to (3.37) shows that

$$p_{222}^{(j,m)}(u) \le C(\epsilon, \gamma_m, \delta_{m-1}, \delta_m)\, g\left(\frac{u}{1+\delta}\right) h(u) \tag{3.41}$$

$$+ \begin{cases} C(\epsilon, \epsilon_{m-1}, \gamma_m, \delta_{m-1}, \delta_m)\ \Big(\nu\Big((u,\infty)\Big)\Big)^{\gamma_m/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } 0 < \alpha < 2 \\ C(\epsilon, \epsilon_{m-1}, \gamma_m, \delta_{m-1}, \delta_m)\ u^{-\gamma_m/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } \alpha > 2 \\ C(\epsilon, \epsilon_{m-1}, \gamma_m, \delta_{m-1}, \delta_m)\ u^{-\gamma_m/4\epsilon} L(u)^{\gamma_m/4\epsilon} g\left(\frac{u}{1+\delta}\right) & \text{if } \alpha = 2 \end{cases}$$

for $u \ge u_0(\epsilon, \epsilon_{m-1}, \gamma_m, \delta_{m-1}, \delta_m)$, $0 < \sigma \le \sigma_0$ and $j = 3, \ldots, k$, $m = 2, \ldots, j-1$.

Summarizing our findings, for any $\xi > 0$ and for any choice of $0 < \epsilon < 1$, $0 < \gamma < 1$, $0 < \delta < 1$, $\epsilon_1, \ldots, \epsilon_{k-1}$, $\gamma_1, \ldots, \gamma_{k-1}$ and $\delta_1, \ldots, \delta_{k-1}$ satisfying the restrictions (3.11) - (3.16) we have, by (3.17), (3.29), (3.30), (3.31), (3.32), (3.37), (3.38), (3.39), (3.40) and (3.41) that

$$P_x(X_\sigma(t) > u) \le \left(1 + \frac{\xi}{4}\right)\left(1 - \frac{\delta_1}{1-\gamma}\right)^{-\alpha} h\Big((1-\gamma)u\Big)$$

$$+ \Big(\nu\Big((\epsilon u, \infty)\Big)\Big)^k g\left(\frac{u}{1+\delta}\right) \prod_{j=2}^{k} g\Big(\epsilon_j u\Big) \tag{3.42}$$

$$+ \begin{cases} C(\text{parameters})\ \Big(\nu\Big((u,\infty)\Big)\Big)^{\gamma_*/4\epsilon} g(u) & \text{if } 0 < \alpha < 2 \\ C(\text{parameters})\ u^{-\gamma_*/4\epsilon} g(u) & \text{if } \alpha > 2 \\ C(\text{parameters})\ u^{-\gamma_*/4\epsilon} L(u)^{\gamma_*/4\epsilon} g(u) & \text{if } \alpha = 2 \end{cases}$$

for all $u \ge u_0(\xi, \text{parameters})$ and all $0 \le \sigma \le \sigma_0$. Here $\gamma_* = \min(\gamma, \gamma_1, \ldots, \gamma_{k-1})$, and $C(\text{parameters})$ may depend on all parameters $0 < \epsilon < 1$, $0 < \gamma < 1$, $0 < \delta < 1$, $\epsilon_1, \ldots, \epsilon_{k-1}$, $\gamma_1, \ldots, \gamma_{k-1}$ and $\delta_1, \ldots, \delta_{k-1}$, and $u_0(\xi, \text{parameters})$ may depend, in addition, on ξ.

Recall that the function h is regularly varying at infinity with exponent $-(\alpha+\beta)+1$ and that the function g is regularly varying at infinity with exponent $-\beta+1$. Hence by the choice (3.10) of k, we can select first δ_1 and γ small enough, and then ϵ small enough (we know from our previous discussion that this is feasible) to conclude that for every $\xi > 0$ there are $u(\xi) > 0$ and $\sigma(\xi) > 0$ such that

$$\frac{P_x(X_\sigma(t) > u)}{h(u)} \le 1 + \xi \quad \text{for all } u \ge u(\xi),\ 0 < \sigma \le \sigma(\xi) \text{ and all } x \in \mathbb{R},\ t \ge 1, \tag{3.43}$$

hence establishing the upper bound in (3.4).

We remark at this point that, if instead of a regular variation of the function f at infinity we only knew that $f(x) \ge \hat{f}(x)$ for all $x \ge 0$ for some function $\hat{f}$ which is regularly varying at infinity with exponent greater than one, then we would still have the upper bound (3.43) with the function h defined now by

$$h(u)=\int_u^\infty \frac{\nu\Big((y,\infty)\Big)}{\hat f(y)}\,dy,\quad u\geq 0\,.$$

We now proceed to establish a corresponding lower bound. For every $0<\epsilon<1$, $0<\delta<1$ and $0<\tau<1$ we have by the strong Markov property and by Lemma 6.1 (i)

$$P_x(X_\sigma(t)>u)\geq P_x\left(X_\sigma(t)>u,\, T_{M(\epsilon u)}\geq t-g\Big(u(1+\delta)\Big),\, X_\sigma\Big(T_{M(\epsilon u)}-\Big)\geq -\tau u\right)$$

$$\geq \int_0^{g\left(u(1+\delta)\right)} \nu\Big((\epsilon,\infty)\Big)e^{-\nu\left((\epsilon,\infty)\right)s}\,P_x\left(X_\sigma(t-s-)\geq -\tau u\right)\,ds \tag{3.44}$$

$$\int_{\epsilon u}^\infty P_{y-\tau u}\left(X_{\epsilon u}(s)>u\right)\frac{\nu(dy)}{\nu\Big((\epsilon,\infty)\Big)}\,.$$

Observe that X_σ has under the measure P_x the same law as $-X_\sigma$ under the measure P_{-x}. Hence we conclude by the remark following (3.43) that for all $0<\sigma\leq\sigma_0$ and $u\geq u_0$ we have

$$P_x\left(X_\sigma(t-s-)\geq -\tau u\right)\geq 1-r(u) \tag{3.45}$$

for all $0<s<g\big(u(1+\delta)\big)$, where $r=r(u)$ is function converging to zero. In the sequel such a function may be different every time it appears.

Observe, further, that by (6.6) and Lemma 6.1, part (ii), if

$$y\geq \tau u+g^{-1}\Big(g\big(u(1+\delta)\big)-s\Big)\,, \tag{3.46}$$

then, under the law $P_{y-\tau u}$, the event $\{X_{\epsilon u}(s)>u\}$ contains, modulo a set of measure zero, the event

$$\left\{ sup_{\substack{1\leq i\leq N_1\\ T_i>t-g(u(1+\delta))}} \sum_{j=i}^{N_1} W_j\mathbf{1}\Big(|W_j|\leq \epsilon u\Big)\leq \delta u\right\}$$

(compare with (3.18), and, hence, for all s and y satisfying (3.46) we have

$$P_{y-\tau u}\left(X_{\epsilon u}(s)>u\right)\geq 1-P\left(sup_{\substack{1\leq i\leq N_1\\ T_i>t-g(u(1+\delta))}} \sum_{j=i}^{N_1} W_j\mathbf{1}\Big(|W_j|\leq \epsilon u\Big)\leq \delta u\right)\,.$$

Therefore, we conclude by (3.26), (3.27) and (3.28) that $0<\sigma\leq\sigma_0$ and $u\geq u_0(\epsilon,\delta)$

$$P_{y-\tau u}\left(X_{\epsilon u}(s)>u\right)\geq 1-r(u) \tag{3.47}$$

for all s and y satisfying (3.46), as long as ϵ is small relatively to δ. We conclude by (3.44), (3.45) and (3.47) that for every $0<\epsilon<1$, $0<\delta<1$ and $0<\tau<1$ such that ϵ is small relatively to δ, and all $0 < \sigma \le \sigma_0$ and $u \ge u_0(\epsilon, \delta)$ we have

$$P_x(X_\sigma(t) > u) \ge \Big(1 - r(u)\Big)^2 \int_0^{g\big(u(1+\delta)\big)} e^{-\nu\big((\epsilon,\infty)\big)s}\, \nu\bigg(\Big(\tau u + g^{-1}\Big(g\big(u(1+\delta)\big) - s\Big), \infty\Big)\bigg)\, ds$$

$$\ge \Big(1 - r(u)\Big)^2 e^{-\nu\big((\epsilon,\infty)\big) g\big(u(1+\delta)\big)} \int_0^{g\big(u(1+\delta)\big)} \nu\Big(\big(\tau u + g^{-1}(s), \infty\big)\Big)\, ds$$

$$= \Big(1 - r(u)\Big)^2 e^{-\nu\big((\epsilon,\infty)\big) g\big(u(1+\delta)\big)} \int_{u(1+\delta)}^{\infty} \frac{\nu\Big(\big(\tau u + y, \infty\big)\Big)}{f(y)}\, dy \tag{3.48}$$

$$\ge \Big(1 - r(u)\Big)^3 \int_{u(1+\delta)}^{\infty} \frac{\nu\Big(\Big(y\left(1 + \frac{\tau}{1+\delta}\right), \infty\Big)\Big)}{f(y)}\, dy$$

$$\ge \Big(1 - r(u)\Big)^4 \left(1 + \frac{\tau}{1+\delta}\right)^{-\alpha} \int_{u(1+\delta)}^{\infty} \frac{\nu\Big(\big(y, \infty\big)\Big)}{f(y)}\, dy = \Big(1 - r(u)\Big)^4 \left(1 + \frac{\tau}{1+\delta}\right)^{-\alpha} h\Big(u(1+\delta)\Big)\,.$$

Choosing $\tau > 0$ small, $\epsilon > 0$ small and, then, $\delta > 0$ small, and using regular variation of h, we conclude that for every $\xi > 0$ there are $u(\xi) > 0$ and $\sigma(\xi) > 0$ such that

$$\frac{P_x(X_\sigma(t) > u)}{h(u)} \ge 1 - \xi \quad \text{for all } u \ge u(\xi),\ 0 < \sigma \le \sigma(\xi) \text{ and all } x \in \mathbb{R},\ t \ge 1. \tag{3.49}$$

Clearly the upper bound (3.43) and the lower bound (3.49) together establish the claim (3.4) and, hence, complete the proof of the theorem. □

4. The stationary solution

In this section we show that the Markov process $(X(t),\, t \ge 0)$ solving the stochastic differential equation (2.1) has a unique stationary distribution.

To this end let, once again, $\sigma > 0$ and let $(X_\sigma(t),\, t \ge 0)$ be the solution to the approximating equation (2.9). We start with showing that, at least for small $\sigma > 0$, this process has a (unique) stationary distribution. The idea is to exploit the regenerative structure of the latter process. It follows from (3.4) that there is $\sigma_0 > 0$ such that the family of laws of $X_\sigma(t)$ under P_x for all $0 < \sigma \le \sigma_0$, $x \in \mathbb{R}$ and $t \ge 1$ is tight. From now on we choose $0 < \sigma \le \sigma_0$.

Define recursively $V_0 = 0$ and for $i \ge 1$

$$V_i = \inf\Big\{t \ge V_{i-1} + 1 :\ X_\sigma(t) = 1\Big\}$$

if $V_{i-1} < \infty$. If $V_{i-1} = \infty$ we also let $V_i = \infty$. By the strong Markov property we see that $V_1, V_2, \ldots$ are (delayed) regeneration times of the process $(X_\sigma(t),\, t \ge 0)$. In particular, $\big((V_{i+1} - V_i),\, i \ge 1\big)$ are iid random variables. We claim that

$$V_i < \infty \ \text{ a.s., } i = 0, 1, 2, \ldots, \text{ and } \ E\Big(V_i - V_{i-1}\Big) < \infty \ \text{ for } i = 1, 2, \ldots. \tag{4.1}$$

Notice that, in that case, it will follow by the Smith theorem (see e.g. Corollary 3.12.3 in Resnick (1992)) that the process $(X_\sigma(t),\, t\geq 0)$ has a limiting distribution. Since this Markov process is Feller, this limiting distribution is, automatically, a stationary distribution and it is, obviously, unique.

We proceed, therefore, to prove (4.1). Let $i\geq 1$. On the event $\{V_{i-1}<\infty\}$ we define recursively

$$A_1^{(i)}=\inf\Big\{t\geq V_{i-1}+1:\, X_\sigma(t)\geq 1\Big\}$$

and for $j\geq 2$ we let

$$A_j^{(i)}=\inf\Big\{t\geq A_{j-1}^{(i)}+g(1)+1:\, X_\sigma(t)\geq 1\Big\}$$

if $A_{j-1}^{(i)}<\infty$ and, as usually, we let $A_j^{(i)}=\infty$ if $A_{j-1}^{(i)}=\infty$. Here, once again, g is given by (6.1) below.

Let $i\geq 2$. By the strong Markov property for every $k\geq 1$ there is a non-random number p_k such that for any $x\in\mathbb{R}$

$$p_k=P_x\Big(A_1^{(i)}-V_{i-1}>2k+1\Big|\mathcal{F}_{V_{i-1}}\Big)$$

on the event $\{V_{i-1}<\infty\}$. Using once again the strong Markov property we have for $k\geq 2$

$$p_k=E_x\Bigg(\mathbf{1}\Big(A_1^{(i)}-V_{i-1}>2k-1\Big)P_{X_\sigma(V_{i-1}+2k-1)}\Big(X_\sigma(t)\leq 1 \text{ for all } 0\leq t\leq 2\Big)\Bigg|\mathcal{F}_{V_{i-1}}\Bigg).$$

By the tightness guaranteed by the choice of σ above, there is $M>\sigma_0$ such that

$$P_x\Big(|X_\sigma(1)|\leq M\Big)\geq \frac{1}{2} \text{ for all } x\in\mathbb{R}.$$

Therefore, for all $x\in\mathbb{R}$ we have

$$P_x\Big(X_\sigma(t)\leq 1 \text{ for all } 0\leq t\leq 2\Big)\leq 1-P_x\Big(X_\sigma(t)>1 \text{ for some } 1\leq t\leq 2\Big)$$

$$\leq 1-P_x\Big(X_\sigma(1)\geq -M\Big)P\Big(L_\sigma \text{ has one jump in the interval } (1,2] \text{ and its size exceeds } M+1\Big)$$

$$\leq 1-\frac{e^{-\lambda_\sigma}}{2}\nu\Big((M+1,\infty)\Big):=\rho_M<1\,.$$

We conclude that the numbers p_k satisfy $p_k\leq\rho_M p_{k-1}$ for all $k\geq 2$ and, hence, for some $C>0$

$$p_k\leq C\rho_M^{k-1} \text{ for } k\geq 1. \tag{4.2}$$

An identical argument shows both that (4.2) holds also for $i=1$ (perhaps with some other constant $C>0$) and also that that for $i\geq 1$ and $j\geq 2$

$$P\Bigg(A_j^{(i)}-\Big(A_{j-1}^{(i)}+g(1)\Big)>2k+1\Bigg|\mathcal{F}_{A_{j-1}^{(i)}+g(1)}\Bigg)\leq C\rho_M^{k-1} \text{ for } k\geq 1 \tag{4.3}$$

on the event $\{A_{j-1}^{(i)}<\infty\}$. Here $C>0$ is a finite positive constant (in particular, it does not depend on i and j). In particular, we conclude that $A_j^{(i)}<\infty$ a.s. on the event $\{V_{i-1}<\infty\}$.

The next step is to observe that for every $x \geq 1$

$$P_x\Big(X_\sigma(t) = 1 \text{ for some } 0 \leq t \leq g(1)\Big) \tag{4.4}$$

$$\geq P\big(L_\sigma \text{ has no jumps in the interval } [0, g(1)]\big) = e^{-\lambda_\sigma g(1)}\,.$$

Keeping $i \geq 1$ fixed we define

$$B_j = \Big\{A_j^{(i)} < \infty \text{ and } X_\sigma(t) = 1 \text{ for some } A_j^{(i)} \leq t \leq A_j^{(i)} + g(1)\Big\}, \quad j = 1, 2, \ldots .$$

It follows from (4.4) and the strong Markov property that for every $x \in \mathbb{R}$

$$P_x\Big(B_j \Big| \mathcal{F}_{A_j^{(i)}}\Big) \geq e^{-\lambda_\sigma g(1)} \tag{4.5}$$

a.s. on the event $\{A_j^{(i)} < \infty\}$. Therefore, a random variable

$$N(\omega) = \begin{cases} n & \text{if } \omega \in B_n \cap \big(\cap_{j=1}^{n-1} B_j^c\big),\ n = 1, 2, \ldots \\ \infty & \text{if } \omega \in \cap_{j=1}^{\infty} B_j^c \end{cases} \tag{4.6}$$

defined on the event $\{V_{i-1} < \infty\}$ is a.s. finite and, furthermore, on that event

$$V_i \leq A_N^{(i)} + g(1)\,. \tag{4.7}$$

We conclude that for every $i \geq 1$, V_i is a.s. finite on the event $\{V_{i-1} < \infty\}$, and an inductive argument shows that $V_i < \infty$ a.s. for all $i = 0, 1, 2, \ldots$. Furthermore, it follows from (4.7) that, with $A_0^{(i)} = V_{i-1}$,

$$E\Big(V_i - V_{i-1}\Big) \leq E\left(\sum_{j=1}^{N}\Big(A_j^{(i)} - A_{j-1}^{(i)}\Big) + g(1)\right) = g(1) + \sum_{j=1}^{\infty} E\Big(\big(A_j^{(i)} - A_{j-1}^{(i)}\big)\mathbf{1}\big(N > j-1\big)\Big)\,.$$

However, we have by (4.3) that for some finite constant $C > 0$, for all $j \geq 1$

$$E\Big(\big(A_j^{(i)} - A_{j-1}^{(i)}\big)\mathbf{1}(N > j-1)\Big) = E\Big(\mathbf{1}(N > j-1) E\Big(\big(A_j^{(i)} - A_{j-1}^{(i)}\big)\Big|\mathcal{F}_{A_{j-1}^{(i)} + g(1)}\Big)\Big)$$

$$\leq CP(N > j-1)\,.$$

Hence

$$E\Big(V_i - V_{i-1}\Big) \leq g(1) + CEN < \infty$$

by (4.6). This proves (4.1) and, hence, the stochastic differential equation (2.9) has a (unique) stationary distribution μ_σ, at least for $0 < \sigma \leq \sigma_0$.

We are now in a position to show that the process $(X(t),\, t \geq 0)$ solving the stochastic differential equation (2.1) has a stationary distribution. To this end, let μ_σ be the stationary distribution for $(X_\sigma(t),\, t \geq 0)$, $0 < \sigma \leq \sigma_0$. It follows from (3.4) that the family $\big(\mu_\sigma,\, 0 < \sigma \leq \sigma_0\big)$ is tight, at least if we reduce σ_0. Let $\big(\mu_{\sigma_n},\, n \geq 1\big)$ be a weakly convergent sequence, for some $\sigma_n \downarrow 0$, $n \to \infty$. Then $\mu_{\sigma_n} \Rightarrow \mu$ as $n \to \infty$ for some probability measure μ.

In the stochastic differential equation (2.9) we choose the initial value distributed according to μ_{σ_n} for $n \geq 1$. We now apply, once again, Theorem 5.4 in Kurtz and Protter (1991) to conclude that the resulting sequence $\big((X_{\sigma_n}(t),\, t \geq 0),\, n \geq 1\big)$ of stationary processes converges weakly in $D[0,\infty)$ to the solution $(X(t),\, t \geq 0)$ of the equation (2.1) for which the initial has the distribution μ. Since the process $(X(t),\, t \geq 0)$ is a.s. continuous at each $t \geq 0$ we conclude that $X(t)$ has the law μ for each $t \geq 0$ and, hence, μ is a stationary distribution for the Markov process $(X(t),\, t \geq 0)$.

In order to show that a stationary distribution of the Markov process $(X(t),\, t \geq 0)$ solving the stochastic differential equation (2.1) is unique, we use a coupling argument. Coupling is a powerful technique of treating stationary distributions of Markov processes as well as other limit phenomena. One indication of its successes are the two recent books, Lindvall (1992) and Thorisson (2000). We will use a simple version of approximate coupling, described in the lemma below. It is similar to the ϵ-coupling in Lindvall (1992), page 74. Note, however, that Thorisson (2000) uses the term ϵ-coupling in a different sense.

Lemma 4.1. *Let $\big(X(t),\, t \geq 0\big)$ and $\big(Y(t),\, t \geq 0\big)$ be two stochastic processes on a probability space $\big(\Omega, \mathcal{F}, P\big)$ such that $X(t) \overset{d}{=} X(0)$ and $Y(t) \overset{d}{=} Y(0)$ for all $t \geq 0$. Let $\epsilon > 0$, and suppose that there is an event $\Omega_+ \in \mathcal{F}$ with $P\big(\Omega_+\big) = 1$ and a random variable $T_\epsilon = T_\epsilon(\omega) \in [0,\infty)$ such that for all $\omega \in \Omega_+$ and $t \geq T_\epsilon(\omega)$ we have $\big|X(t,\omega) - Y(t,\omega)\big| \leq \epsilon$. Then*

$$P\Big(X(0) > x + \epsilon\Big) \leq P\Big(Y(0) > x\Big) \leq P\Big(X(0) > x - \epsilon\Big) \tag{4.8}$$

for all $x \in \mathbb{R}$.

Proof. Notice that

$$\limsup_{n\to\infty} P\Big(\Big|X(n) - Y(n)\Big| > \epsilon\Big) \leq P\Big(\limsup_{n\to\infty}\Big\{\Big|X(n) - Y(n)\Big| > \epsilon\Big\}\Big) \leq P\Big(\Omega_+^c\Big) = 0\,.$$

Therefore, $P\big(|X(n) - Y(n)| > \epsilon\big) \to 0$ as $n \to \infty$. Now (4.8) follows from the obvious observation that for every $x \in \mathbb{R}$

$$P\Big(\Big|X(n) - Y(n)\Big| > \epsilon\Big)$$

$$\geq \max\Big(P\Big(X(n) > x + \epsilon\Big) - P\Big(Y(n) > x\Big),\; P\Big(Y(n) > x\Big) - P\Big(X(n) > x - \epsilon\Big)\Big)$$

$$= \max\Big(P\Big(X(0) > x + \epsilon\Big) - P\Big(Y(0) > x\Big),\; P\Big(Y(0) > x\Big) - P\Big(X(0) > x - \epsilon\Big)\Big)\,.$$

□

Consider now two stochastic processes solving the equation (2.1) with two different initial points,

$$dX^{(i)}(t) = -f(X^{(i)}(t))\,dt + dL(t)\,, \quad X^{(i)}(0) = x_i\,,\ i = 1,2\,,$$

driven by the same Lévy motion L. We claim that

(4.9) if for some $\epsilon > 0$, $|x_1 - x_2| \le \epsilon$ then $\left|X^{(1)}(t) - X^{(2)}(t)\right| \le \epsilon$ a.s. for all $t \ge 0$.

Indeed, let for $\sigma > 0$

$$dX_\sigma^{(i)}(t) = -f(X_\sigma^{(i)}(t))\,dt + dL(t)\,, \quad X_\sigma^{(i)}(0) = x_i\,,\ i = 1,2\,,$$

keeping the same driving process L and the same initial values x_1 and x_2. It follows from Theorem 5.4 of Kurtz and Protter (1991) that for each $t \ge 0$ $\left(X_\sigma^{(1)}(t), X_\sigma^{(2)}(t)\right) \Rightarrow \left(X^{(1)}(t), X^{(2)}(t)\right)$ as $\sigma \to 0$. Hence

$$P\left(\left|X^{(1)}(t) - X^{(2)}(t)\right| \le \epsilon\right) \ge \limsup_{\sigma\to 0} P\left(\left|X_\sigma^{(1)}(t) - X_\sigma^{(2)}(t)\right| \le \epsilon\right),$$

and the fact that the latter limit is equal to one is an immediate conclusion from part (ii) of Lemma 6.1.

We have now all the ingredient needed to prove the uniqueness of a stationary distribution of the Markov process $(X(t),\, t \ge 0)$ solving the stochastic differential equation (2.1). Let μ and ν be two such stationary distributions. Define

$$dX(t) = -f(X(t))\,dt + dL^{(1)}(t)\,, \quad X(0) \sim \mu\,,$$

$$dY(t) = -f(Y(t))\,dt + dL^{(2)}(t)\,, \quad Y(0) \sim \nu\,,$$

with $L^{(1)}$ and $L^{(2)}$ being independent copies of the Lévy motion L. Let

$$T_\epsilon = \inf\left\{t \ge 0 :\ \left|X(t)\right| \le \frac{\epsilon}{2},\ \left|Y(t)\right| \le \frac{\epsilon}{2}\right\}.$$

It follows from Lemma 6.3 that $T_\epsilon < \infty$ a.s. By the strong Markov property

$$L_*^{(i)}(t) = \begin{cases} L^{(i)}(t) & \text{if } 0 \le t \le T_\epsilon \\ L^{(i)}(T_\epsilon) + \left(L^{(1)}(t) - L^{(1)}(T_\epsilon)\right) & \text{if } t > T_\epsilon \end{cases}$$

for $i = 1,2$ form a pair of (non-independent) copies of the Lévy motion L. Let

$$dX_*(t) = -f(X_*(t))\,dt + dL_*^{(1)}(t)\,, \quad X_*(0) = X(0) \sim \mu\,,$$

$$dY_*(t) = -f(Y_*(t))\,dt + dL_*^{(2)}(t)\,, \quad Y_*(0) = Y(0) \sim \nu\,.$$

Note that $(X_*(t),\, t \ge 0)$ coincides with $(X(t),\, t \ge 0)$ on the interval $\left[0, T_\epsilon\right]$ and $(Y_*(t),\, t \ge 0)$ coincides with $(Y(t),\, t \ge 0)$ on the same interval. Using (4.9) and the strong Markov property we conclude that on an event of probability 1 we have $\left|X_*(t)-Y_*(t)\right|\leq\epsilon$ for all $t\geq T_\epsilon$. Applying Lemma 4.1 we conclude that

$$\mu\Big\{\Big(x+\epsilon,\infty\Big)\Big\}\leq\nu\Big\{\Big(x,\infty\Big)\Big\}\leq\mu\Big\{\Big(x-\epsilon,\infty\Big)\Big\} \tag{4.10}$$

for all $x\in\mathbb{R}$, and since (4.10) holds for all $\epsilon>0$ we let $\epsilon\to 0$ to conclude that $\mu=\nu$. We have, therefore,

Theorem 4.1. *Assume that the the Lévy motion* $(L(t),\, t\geq 0)$ *is symmetric and that the assumptions (1.5) and (2.2) – (2.5) hold. Then the Markov process* $(X(t),\, t\geq 0)$ *solving the stochastic differential equation (2.1) has a unique stationary distribution* μ. *This stationary distribution satisfies*

$$\lim_{u\to\infty}\frac{\mu\{(u,\infty)\}}{h(u)}=1\,, \tag{4.11}$$

where the function h *is defined in (3.1).*

Proof. Existence and uniqueness of the stationary distribution has been established above, and (4.11) follows from Theorem 3.1. □

Various sufficient conditions for existence of the stationary distribution have been established in Brockwell et al. (1982), under the assumption that the L'evy process is a subordinator, and, hence, the state space of the solution to the stochastic differential equation is a subset of $[0,\infty)$. The above paper allows a very general "release rate" function f. Results on tail behavior of the stationary distribution are established in Asmussen (1998), under the assumption that the L'evy process L is both a subordinator and compound Poisson, with a finite mean. These results allow general subexponential tails of the Lévy measure ν, and a very general function f (which is only required to be at least at a positive distance above the mean for large values of the argument to guarantee existence of a stationary distribution). If a certain level crossings identity holds, then the above paper establishes a *density version* of (4.11).

The following example illustrates what may happen if the function f in the equation (2.1) does not increase as fast as required in Theorems 3.1 and 4.1.

Example 4.2. Let

$$dX(t)=-aX(t)\,dt+dL(t)\,,\ \text{some }a>0\,,$$

$X(0)=x$. The solution is the Ornstein-Uhlenbeck process

$$X(t)=xe^{-at}+\int_0^t e^{-a(t-u)}\,dL(u)\,,\quad t\geq 0\,,$$

which is itself an infinitely divisible process. An easy computation of the Lévy measure ν_t of $X(t)$ gives us

$$\nu_t\Big((u,\infty)\Big) = \int_0^t \nu\Big((ue^{ax},\infty)\Big) \sim \frac{1}{\alpha a}\left(1-e^{-\alpha a t}\right)\nu\Big((u,\infty)\Big)$$

as $u \to \infty$, and hence by, say, Embrechts et al. (1979),

$$P_x(X(t) > u) \sim \nu_t\Big((u,\infty)\Big) \sim \frac{1}{\alpha a}\left(1-e^{-\alpha a t}\right)\nu\Big((u,\infty)\Big) \tag{4.12}$$

as $u \to \infty$. On the other hand,

$$h(u) \sim \frac{1}{\alpha a}\nu\Big((u,\infty)\Big) \tag{4.13}$$

as $u \to \infty$. In this example the difference between (4.12) and (4.13) disappears as $t \to \infty$.

5. Tail estimates and characteristic function

The process in (2.1) can be viewed as the state of a nonlinear mechanical system with restoring function f that responds quasi–statically to a Lévy noise $\big(L(t),\, t \geq 0\big)$. Such applications were the main reason for this paper. Correspondingly, in this section we treat certain more applied issues related to such nonlinear systems.

In the cases when one does not know the tail behavior of the distribution of $X(t)$, but suspects a power–like tail, a common approach to verification of that suspicion and estimation of the tail exponent is to use a semiparametric tail estimator, of which Hill estimator (Hill (1975)). It is well known that tail estimators are not very robust, depend heavily on the fraction of upper order statistics used in the estimation, and are not likely to provide reliable results unless the sample size is really large, see e.g. Drees et al. (2000). In the case of data produced by a mechanism that allows sampling at high frequencies the temptation is to increase the sample size by sampling at a high frequency. In the case of a stochastic differential equation which is normally simulated by running a recusrion with very small time steps this is especially natural. There is, however, a danger that sampling the points too frequently in a relatively short time interval may introduce bias in tail estimation caused by dependence in the data. While the general issue of tail estimation in continuous time models remains under investigation, we provide here some simulation experiments that allow us to observe some of the issues we have just discussed.

Given a data set X_i, $i = 1, 2, \ldots, n$, let $X_{(1)} \geq X_{(2)} \geq \cdots \geq X_{(n)}$ denote the corresponding order statistics. We define

$$H_{k,n} = \frac{1}{k}\sum_{i=1}^{k} \log\left(\frac{X_{(i)}}{X_{(k+1)}}\right)$$

to be the Hill estimator based on the $k+1$ upper order statistics of the sample of X. If $k, n \to \infty$ such that $k/n \to 0$, then $1/H_{k,n} \to \rho$ in probability, if the observations are independent and identically distributed and the right tail of the common distribution of the observations is regularly varying with exponent ρ. The same conclusion holds if the assumption of independence is replaced by the assumption of certain stationary models; see e.g. Resnick and Stărică (1995).

Let X be defined by (2.1) with $f(x) = x + x^3$, $x \in \mathbb{R}$, with L being a Cauchy motion, i.e. a symmetric α–stable motion with $\alpha = 1$. In this case, $\nu((y,\infty)) = Cy^{-1}$, $y > 0$, for some $C > 0$. Therefore, for the function h in (3.1) we have $h(u) \sim Cu^{-3}$ (recall we are allowing finite positive constants to change) as $u \to \infty$. By Theorem 3.1, $P_x(X(t) > u) \sim Cu^{-3}$ as $u \to \infty$, that is, the the marginal distribution of X has a power–like right tail with exponent $\rho = 3$.

We have generated a sample from a solution to the equation (2.1) by a finite difference approximation with a time step $\Delta t = 0.001$. Call the resulting data set is $X_1, X_2, \ldots$. We then select an equally spaced subset of observations $X_1, X_{q+1}, X_{2\,q+1}, \ldots$, to which Hill estimator is applied. We view $q \geq 1$ as a parameter in the estimation scheme.

To obtain the he plots in Fig. 1 we we generated a sample $X_1, X_2, \ldots$ as above of the $n = 1{,}000{,}000$. The four collections of plots in this figure correspond to different choices of q; for each choice of q the horizontal axis shows the percentage of the resulting effective sample (of the size n/q) used in the tail estimation, and the different plots are for different numbers of the tail order statistics used in the Hill estimator. If m is equal to the fraction on the horizontal axis times n/q, then m is the number of the observations used in the tail estimation, and if k is the number of tail order statistics used in the Hill estimator, then the different plots are for $k/m = 0.01$, 0.005, 0.001, 0.0005, and 0.0001 starting from the bottom plot in each graph. It is clear that the Hill estimates become closer to the theoretical value $\rho = 3$ as q increases (even thouugh it reduces the effective sample size). In fact, the effect of inceased sample size seems to be ambiguous, especially for small q. In Figure 2 we keep a constant number of observations $n = 1{,}000{,}000$, no matter what the separation q between the observations is, and plot Hill estimates $1/H_{k,n}$ for several values of k/n (k is, once again, the number of tail order statistics used in the Hill estimator), and increasing values of q. We cannot explain at the moment why the most accurate estimate of the theoretical value $\rho = 3$ has been obtained for $k/n = 0.005$.

Another, often heuristic, approach to obtain information about the distribution of the solution $X(t)$ to an equation like (2.1), is by trying to derive an equation the characteristic function of of $X(t)$ satisfies. Denote the characteristic function by $\varphi(u;t) = E[\exp(i\,u\,X(t))]$. Consider, for

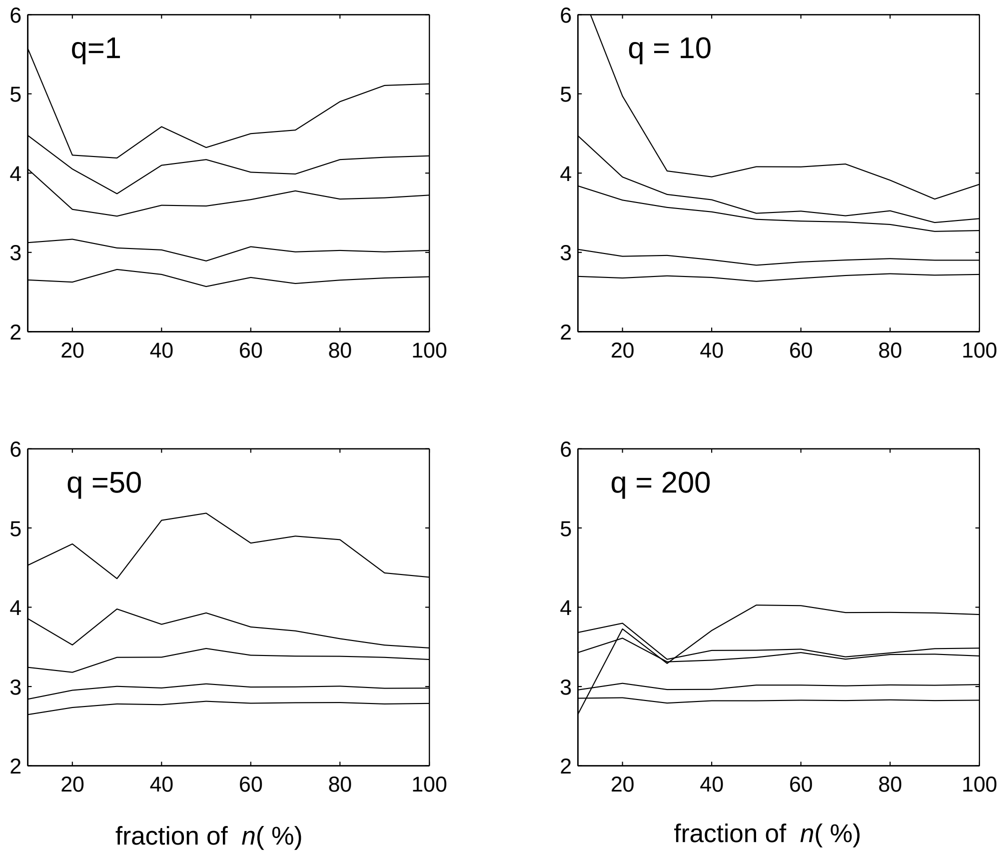

FIGURE 1. The dependence of the Hill estimates on the sample size and the ratio k/m.

the moment, the linear version of equation (2.1), with $f(x) = x$, $x \in \mathbb{R}$, and assume that L is a symmetric α-stable motion with $0 < \alpha < 2$, with Lévy measure given by (1.6) with $c_+ = c_- = 1$.

We know from Example 4.2 above that

$$\varphi(u;t) = \exp\Big(i\,u\,x\,e^{-t} - C_\alpha\,(1 - e^{-\alpha\,t})\,|u|^\alpha\Big)\,, \tag{5.1}$$

where $C_\alpha = 2\int_0^\infty (1 - \cos y) y^{-(1+\alpha)}\,dy$.

The Itô formula applied to the function $e^{i\,u\,X(t)}$ gives us

$$e^{i\,u\,X(t)} - e^{i\,u\,x} = -i\,u\int_0^t X(s)\,e^{i\,u\,X(s)}\,ds + \sum_{0<s\le t}\Big[e^{i\,u\,X(s)} - e^{i\,u\,X(s-)}\Big]\,. \tag{5.2}$$

It is not difficult to check that the second term in the right hand side of (5.2) converges a.s. (as one performes the summation over the jumps of the magnitude greater than $\epsilon > 0$ and then let ϵ go to zero), and that its expectation is equal to

$$\int_{\mathbb{R}}\Big(e^{i\,u\,y} - 1\Big)\,\nu(dy)\int_0^t \varphi(u;s)\,ds = -\alpha C_\alpha |u|^\alpha \int_0^t \varphi(u;s)\,ds\,.$$

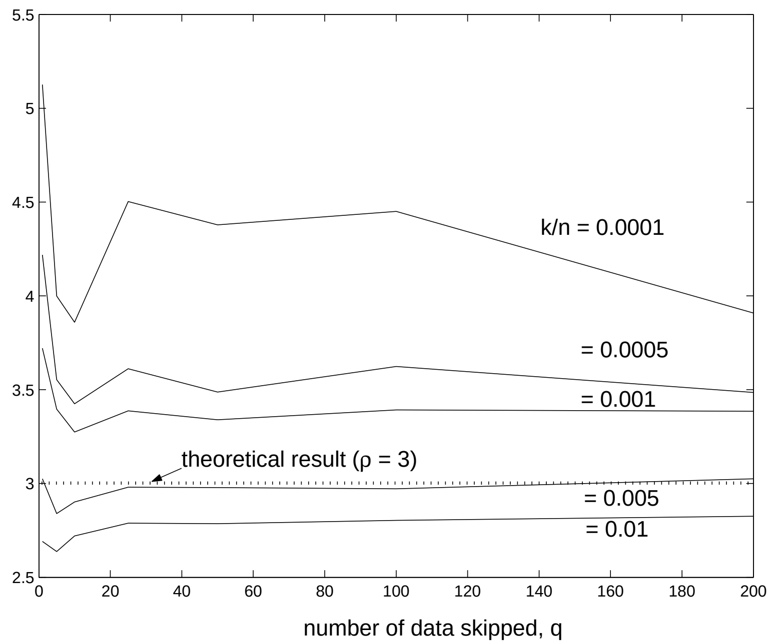

FIGURE 2. Hill estimates for $n = 1,000,000$ and increasing values of q.

Of course, $X(s)\, e^{i\,u\,X(s)}$ is not integrable for $s > 0$. If it were, however, then its expectation would be just $-i\partial\varphi(u;s)/\partial u$, and so taking expectations in (5.2) one would get formally

$$\varphi(u;t) = e^{i\,u\,x} - u\int_0^t \frac{\partial\varphi(u;s)}{\partial u}\, ds - \alpha C_\alpha |u|^\alpha \int_0^t \varphi(u;s)\, ds\,,$$

and differentiating with respect to time will formally require φ to satisfy the partial differential equation

$$\frac{\partial\varphi(u;t)}{\partial t} = -u\,\frac{\partial\varphi(u;t)}{\partial u} - \alpha C_\alpha |u|^\alpha\, \varphi(u;t) \tag{5.3}$$

with the initial condition $\varphi(u;0) = e^{i\,u\,x}$, *provided the infinities cancel the right way in the above "improper" computation of the expectation.*

It is easy to check that the characteristic function in (5.1) is a solution of this equation over $u \in [0,\infty)$ and $u \in (-\infty, 0]$. Hence, in this case the infinities seem to have canceled each other in the right way.

Such heuristics may, however, fail in other cases. Continuing the above example, let L be the Cauchy motion ($\alpha = 1$) with $c_+ = c_- = 1$ in (1.6), and let this time $f(x) = x^3 + ax$, $x \in \mathbb{R}$ for some $a > 0$. Heuristic considerations similar to the above lead to the following partial differential equation

$$\frac{\partial\varphi(u;t)}{\partial t} = u\,\frac{\partial^3\varphi(u;t)}{\partial u^3} - a\,u\,\frac{\partial\varphi(u;t)}{\partial u} - C_1 |u|\, \varphi(u;t). \tag{5.4}$$

The next heuristic step one takes is that to consider the "stationary case" by assuming existence of a stationary distribution, and writing $\varphi_s(u)$ for the characteristic function of the stationary solution. Then, hoping that interchanging various derivatives and integrals works fine, one arrives at a stationary version of the above partial differential equation, which is now an ordinary differential equation,

$$u\,\varphi_s'''(u) - au\,\varphi_s'(u) - C_1|u|\,\varphi_s(u) = 0. \tag{5.5}$$

However, all solutions of the equation (5.5) whose absolute values do not exceed one are the characteristic functions of (perhaps, shifted) Cauchy laws, and this is inconsistent with the result of Theorem 4.1 showing that, in this case, the tail probability $P(X(t) > \lambda)$ of the stationary state $X(t)$ decays as λ^{-3} as $\lambda \to \infty$.

6. Lemmas and auxiliary results

We start with some simple facts about certain ordinary differential equations and related dynamical systems. Let f be a function satisfying (2.2), (2.3) and (2.4). Denote

$$g(u) = \int_u^\infty \frac{1}{f(y)}\,dy, \quad u \geq 0\,. \tag{6.1}$$

It follows from the assumptions (2.2) – (2.4) the function g is finite, continuous and strictly increasing for large u. In particular, it has an inverse, g^{-1} defined in a right neighborhood of the origin. Furthermore, $g(u) \uparrow \infty$ as $u \downarrow 0$, and, moreover,

$$g \text{ is regularly varying at infinity with exponent } -\beta + 1\,. \tag{6.2}$$

Let $A > 0$ and $0 = S_0 < S_1 < \ldots < S_n < A < S_{n+1}$. Consider the system $y(0) = y_0$,

$$y'(t) = -f\Big(y(t)\Big)\,,\ t \in (S_{i-1}, S_i),\ i = 1, \ldots, n, n+1\,, \tag{6.3}$$

and

$$y(S_i) = y(S_i-) + j_i,\ i = 1, \ldots, n \tag{6.4}$$

for some sequence of real numbers $j_1, \ldots, j_n$. This system of equations has a unique solution. This solution satisfies, in particular,

$$g\Big(y(A)\Big) = g\Big(y(S_n)\Big) + A - S_n\,. \tag{6.5}$$

In particular, for any $u > 0$

$$\text{if } y(A) > u \text{ then } A - S_n \leq g(u)\,. \tag{6.6}$$

Lemma 6.1. *Let f be a function satisfying (2.2) and (2.3).*

(i) *Let*

$$x'(t) = -f\Big(x(t)\Big)\,,\ t>0\,,\ x(0)=x_0\,, \tag{6.7}$$

$$z'(t) = -f\Big(z(t)\Big)\,,\ t>0\,,\ z(0)=z_0\,. \tag{6.8}$$

If $x_0 \geq z_0$ then for all $t\geq 0$ we have $0 \leq x(t)-z(t) \leq x_0 - z_0$.

(ii) *Let $x(t), t\geq 0$ satisfy (6.7) and let $y(t), t\geq 0$ satisfy (6.3) and (6.3) with $y_0 = x_0$, Then*

$$-\max_{k=1,\dots,n}\left(\sum_{i=k}^{n} j_i\right)_- \leq y(A)-x(A) \leq \max_{k=1,\dots,n}\left(\sum_{i=k}^{n} j_i\right)_+ . \tag{6.9}$$

(iii) *In the system (6.3) and (6.4) let $N(a) = \sup\{i \leq n:\ j_i + \ldots + j_n > a\}$ ($=0$ if the set is empty) for $a>0$. For $a,b>0$ and $S_{N(a)} \leq s \leq A$, if $y(s)\leq b$ then $y(A) \leq a+b$.*

Proof. Part (i) is obvious. For part (ii) we proceed by induction. For $n=1$ the statement of part (ii) is an immediate consequence of part (i). Assuming that the statement of part (ii) holds for some $n=n_0\geq 1$ we have, in the case $n=n_0+1$, by the assumption of the induction

$$-\max_{k=1,\dots,n_0}\left(\sum_{i=k}^{n_0} j_i\right)_- \leq y(S_{n_0+1}-) - x(S_{n_0+1}-) \leq \max_{k=1,\dots,n_0}\left(\sum_{i=k}^{n_0} j_i\right)_+$$

and, hence,

$$y(S_{n_0+1}) - x(S_{n_0+1}) \leq \max_{k=1,\dots,n_0}\left(\sum_{i=k}^{n_0} j_i\right)_+ + j_{n_0+1} \leq \max_{k=1,\dots,n_0+1}\left(\sum_{i=k}^{n_0+1} j_i\right)_+$$

and

$$y(S_{n_0+1}) - x(S_{n_0+1}) \geq -\max_{k=1,\dots,n_0}\left(\sum_{i=k}^{n_0} j_i\right)_- + j_{n_0+1} \geq -\max_{k=1,\dots,n_0+1}\left(\sum_{i=k}^{n_0+1} j_i\right)_- .$$

Now the statement of part (ii) for $n=n_0+1$ follows by part (i).

Finally, part (iii) of the lemma follows from its part (ii) (specifically, from the right hand side inequality in (6.9)). □

The next lemma is very useful when dealing with sums of truncated random variables. It is due to Prokhorov (1959). See also Petrov (1995).

Lemma 6.2. *Let $Y_1,\dots,Y_n$ be independent zero mean random variables, such that $|Y_j|\leq c$ for some $c>0$. Then*

$$P\left(Y_1+\ldots Y_n > \lambda\right) \leq \exp\left\{-\frac{\lambda}{2c}\,\text{arsinh}\,\frac{c\lambda}{2\,\text{var}(Y_1+\ldots+Y_n)}\right\}\,,\quad \lambda>0\,.$$

A trivial but useful fact in connection with this lemma is that

$$\operatorname{arsinh}(x) \sim \log x \quad \text{as } x \to \infty. \tag{6.10}$$

Let, for $\lambda > 0$, X_λ be a Poisson random variable with mean λ. A direct inspection of the limiting cases $\lambda \to 0$ and $\lambda \to \infty$ shows that for any $p \ge 1$ there is a constant $C_p > 0$ such that

$$EX_\lambda^p \le C_p(\lambda + \lambda^p), \quad \text{any } \lambda > 0. \tag{6.11}$$

The next lemma provides a uniform lower bound for the probabilities of balls centered at the origin for the solution of the equation (2.1). Note that the argument does not require any assumptions on the Lévy process L (apart from its symmetry) and it does not use regular variation of the function f (only finiteness of the function g in (6.1) is required). More precisely, let

$$g_1(u) = \int_u^\infty \frac{1}{\min\bigl(f(y), -f(-y)\bigr)}\, dy, \quad u \ge 0\,. \tag{6.12}$$

Lemma 6.3. *Let $(X(t), t \ge 0)$ be the solution to stochastic differential equation (2.1). The for every $\epsilon > 0$ there is $p_\epsilon > 0$ such that*

$$P_x\Bigl(X(t) \in [-\epsilon, \epsilon]\Bigr) \ge p_\epsilon \;\; \textit{for all } t \ge g_1\left(\frac{\epsilon}{2}\right) \textit{ and } x \in \mathbb{R}.$$

Proof. By Theorem 5.4 in Kurtz and Protter (1991) our claim will follow once we show that for every $\epsilon > 0$ there is $\sigma_0(\epsilon) > 0$ and $p_\epsilon > 0$ such that

$$P_x\Bigl(X_\sigma(t) \in [-\epsilon, \epsilon]\Bigr) \ge p_\epsilon \;\; \text{for all } 0 < \sigma \le \sigma_0(\epsilon) > 0,\ t \ge g_1\left(\frac{\epsilon}{2}\right) \text{ and } x \in \mathbb{R}. \tag{6.13}$$

To this end note that it follows from (6.6) that, if $X_\sigma\bigl(t - g_1(\epsilon/2)\bigr) \ge 0$, then without any jumps of L_σ in the interval $\bigl[t - g_1(\epsilon/2), t\bigr]$ one would have $0 \le X(t) \le \epsilon/2$. Similarly, if $X_\sigma\bigl(t - g_1(\epsilon/2)\bigr) \le 0$, then without any jumps of L_σ in the interval $\bigl[t - g_1(\epsilon/2), t\bigr]$ one would have $-\epsilon/2 \le X(t) \le 0$. We conclude by part (ii) of Lemma 6.1 that

$$P_x\Bigl(X_\sigma(t) \in [-\epsilon, \epsilon]\Bigr) \ge P\left(\max_{j \le N_{\epsilon,\sigma}} \Bigl|W_1 + \ldots + W_j\Bigr| \le \frac{\epsilon}{2}\right), \tag{6.14}$$

where $W_1, W_2, \ldots$ are iid random variables with the common law F_σ given by (3.6) independent of a Poisson random variable $N_{\epsilon,\sigma}$ with the mean $g_1(\epsilon/2)\lambda_\sigma$.

Let $K > 0$. We have for all $0 < \sigma \le \epsilon/K$

(6.15)

$$P\left(\max_{j \le N_{\epsilon,\sigma}} \Bigl|W_1 + \ldots + W_j\Bigr| \le \frac{\epsilon}{2}\right) \ge P\left(\max_{j \le N_{\epsilon,\sigma}} \bigl|W_j\bigr| \le \frac{\epsilon}{K}\right) P\left(\max_{j \le N^{(1)}_{\epsilon,\sigma}} \Bigl|W_1^{1)} + \ldots + W_j^{(1)}\Bigr| \le \frac{\epsilon}{2}\right),$$

where $W_1^{1)}, W_2^{1)}, \ldots$ are iid random variables with the common law

$$F_{\sigma,K}(A) = \frac{F_\sigma\left(A \cap \left[-\frac{\epsilon}{K}, \frac{\epsilon}{K}\right]\right)}{F_\sigma\left(\left[-\frac{\epsilon}{K}, \frac{\epsilon}{K}\right]\right)}, \quad A \text{ a Borel set},$$

independent of a Poisson random variable $N_{\epsilon,\sigma}^{(1)}$ with the mean $g_1(\epsilon/2)\lambda_\sigma F_\sigma\big([-\epsilon/K,\epsilon/K]\big)$. By the choice of σ we have

$$P\left(\max_{j\le N_{\epsilon,\sigma}}\left|W_j\right|\le\frac{\epsilon}{K}\right)=e^{-g_1(\epsilon/2)\nu((\epsilon/K,\infty))}\,. \tag{6.16}$$

Furthermore, by the symmetry

$$P\left(\max_{j\le N_{\epsilon,\sigma}^{(1)}}\left|W_1^{1)}+\ldots+W_j^{(1)}\right|>\frac{\epsilon}{2}\right)\le 4P\left(\left|W_1^{1)}+\ldots+W_{N_{\epsilon,\sigma}^{(1)}}\right|>\frac{\epsilon}{2}\right)\,. \tag{6.17}$$

Observe that $W_1^{1)}+\ldots+W_{N_{\epsilon,\sigma}^{(1)}}$ is an infinitely divisible random variable with the Lévy measure

$$g_1\left(\frac{\epsilon}{2}\right)\lambda_\sigma F_\sigma\left(\cdot\cap\left[-\frac{\epsilon}{K},\frac{\epsilon}{K}\right]\right)=g_1\left(\frac{\epsilon}{2}\right)\,\nu\left(\cdot\cap\left(\left[-\frac{\epsilon}{K},-\sigma\right]\cup\left[\sigma,\frac{\epsilon}{K}\right]\right)\right)\,.$$

As $\sigma\to 0$ this Lévy measure converges vaguely to

$$g_1\left(\frac{\epsilon}{2}\right)\,\nu\left(\cdot\cap\left[-\frac{\epsilon}{K},\frac{\epsilon}{K}\right]\right)\,,$$

which is the Lévy measure of some infinitely divisible random variable, say, Z_K. Notice that $Z_K\Rightarrow 0$ as $K\to\infty$. Hence there is K large enough so that $P(|Z_K|\ge\epsilon/2)\le 1/16$. Fixing that K and noticing that $W_1^{1)}+\ldots+W_{N_{\epsilon,\sigma}^{(1)}}\Rightarrow Z_K$ as $\sigma\to 0$ we see that there is $0<\sigma_0(\epsilon)=\sigma_0(\epsilon,K)\le\epsilon/K$ such that

$$P\left(\left|W_1^{1)}+\ldots+W_{N_{\epsilon,\sigma}^{(1)}}\right|>\frac{\epsilon}{2}\right)\le\frac{1}{8}\ \text{ for all } 0<\sigma\le\sigma_0(\epsilon). \tag{6.18}$$

It follows from (6.15) – (6.18) that

$$P\left(\max_{j\le N_{\epsilon,\sigma}}\left|W_1+\ldots+W_j\right|\le\frac{\epsilon}{2}\right)\ge\frac{1}{2}e^{-g_1(\epsilon/2)\nu((\epsilon/K,\infty))}\,. \tag{6.19}$$

Now the claim (6.13) with $p_\epsilon=\exp\{-g_1(\epsilon/2)\nu((\epsilon/K,\infty))\}/2$ follows from (6.14) and (6.19). □

References

S. Asmussen (1998): Subexponential asymptotics for stochastic processes: Extremal behavior, stationary distributions and first passage probabilities. *The Annals of Applied Probability* 8:354–374.

J. Bertoin (1996): *Lévy Processes*. Cambridge University Press, Cambridge.

P. Billingsley (1999): *Convergence of Probability Measures*. Wiley, New York, 2nd edition.

P. Brockwell, S. Resnick and R. Tweedie (1982): Storage processes with general release rule and additive inputs. *Advances in Applied Probability* 14:392–433.

H. Drees, L. de Haan and S. Resnick (2000): How to make a Hill plot. *Annals of Statistics* 28:254–274.

P. EMBRECHTS, C. GOLDIE and N. VERAVERBEKE (1979): Subexponentiality and infinite divisibility. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 49:335–347.

B. HILL (1975): A simple general approach to inference about the tail of a distribution. *Ann. Statist.* 3:1163–1174.

T. KURTZ and P. PROTTER (1991): Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability* 19:1035–1070.

T. LINDVALL (1992): *Lectures on the Coupling Method*. Wiley, New York.

V. PETROV (1995): *Limit Theorems of Probability Theory*. Oxford University Press, Oxford.

Y. PROKHOROV (1959): An extremal problem in probability theory. *Theory of Probability and Its Applications* 4:201–204.

P. PROTTER (1991): *Stochastic integration and differential equations. A new approach*. Springer-Verlag, Berlin.

S. RESNICK (1987): *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York.

S. RESNICK (1992): *Adventures in Stochastic Processes*. Birkhäuser, Boston.

S. RESNICK and C. STĂRICĂ (1995): Consistency of Hill estimator for dependent data. *Journal of Applied Probability* 32:139–167.

G. SAMORODNITSKY and M. TAQQU (1994): *Stable Non-Gaussian Random Processes*. Chapman and Hall, New York.

K. SATO (1999): *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

H. THORISSON (2000): *Coupling, Stationarity and Regeneration*. Springer, New York.

SCHOOL OF OPERATIONS RESEARCH AND INDUSTRIAL ENGINEERING, AND DEPARTMENT OF STATISTICAL SCIENCE, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: gennady@orie.cornell.edu

SCHOOL OF CIVIL AND ENVIRONMENTAL ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: mdg12@cornell.edu