The Annals of Applied Probability
2003, Vol. 13, No. 1, 1–36

RUIN PROBLEM AND HOW FAST STOCHASTIC PROCESSES MIX[1]

BY PAUL EMBRECHTS AND GENNADY SAMORODNITSKY[2]

ETHZ and Cornell University

The recent increasing interplay between actuarial and financial mathematics has led to a surge in risk theoretic modeling. Especially actuarial ruin models under fairly general conditions on the underlying risk process have become a focus of attention. Motivated by applications such as the modeling of operational risk losses in financial risk management, we investigate the stability of classical asymptotic ruin estimates when claims are heavy, and this under variability of the claim intensity process. Various examples are discussed.

1. Introduction. Over the recent years, we have seen an increasing interest in the finer analysis of actuarial risk models. One of the main reasons is the growing importance of integrated risk management (IRM) and the resulting stochastic modeling of financial solvency; see, for instance, Doherty (2000), Briys and de Varenne (2001) and Kaufmann, Gadmer and Klett (2001). One important class of such models concerns ruin theory as it is known in the actuarial literature; see Rolski, Schmidli, Teugels and Schmidt (1999) for a detailed overview of risk theory in general, and Asmussen (2000) for an up-to-date account of ruin theory. A further motivation for the results presented here stems from the area of operational risk as, for instance, discussed in Crouhy, Mark and Galai (2000) and Medova (2000). The loss process over the time period $[0,t]$ will be denoted by $(Y(t),\ t \geq 0)$. Typically, in insurance or for the modeling of operational risk losses, $Y(t)$ will have the form

$$Y(t) = \sum_{k=1}^{N(t)} Y_k, \qquad t \geq 0,$$

for some counting process $(N(t),\ t \geq 0)$ and claim process $(Y_k,\ k = 1, 2, \ldots)$. Depending on the application, one could then think of some premium rate system that compensates the expected losses; that is, we are looking at the process $(Y(t) - ct,\ t \geq 0)$. For such a process, a risk capital u_ε can be defined as that initial capital that associates a given, small probability ε to the event "over a given

Received October 2001; revised April 2002.

[1]Supported by Forschungsinstitut für Mathematik of ETH, Zurich.

[2]Supported in part by NSF Grant DMS-00-71073 and by NSF/SRC Grant DMI-97-13549.

AMS 2000 subject classifications. Primary 60E07, 60G10; secondary 60K30.

Key words and phrases. Ruin probability, heavy tails, supremum, negative drift, insurance risk, speed of mixing.

accounting period, $Y(t) - ct$ will be larger than u_ε." How does one estimate such u_ε? How can one use classical ruin theoretic estimates assuming that the loss intensity is random, only satisfying some very mild conditions? The results given in this paper mainly address the latter question. We now make the mathematical setup more precise.

Let $(Y(t),\ t \geq 0)$ be a general separable stochastic process, which, as, for instance, in the insurance example above, we view as a claim process. That is, for $t \geq 0$ we view $Y(t)$ as the total amount of claims received in the time interval $[0, t]$. Let $c > 0$ be the premium rate. We assume that

$$P\left(\lim_{t\to\infty}\big(Y(t) - ct\big) = -\infty\right) = 1. \tag{1.1}$$

The *ruin probability*

$$\psi_0(u) = P\left(\sup_{t\geq 0}\big(Y(t) - ct\big) > u\right) \tag{1.2}$$

and its modifications are the main objects of interest in risk theory; $\psi_0(u)$ describes the likelihood of eventual ruin when the initial capital is $u > 0$. The assumption (1.1) means that the ruin is not certain if the initial capital is large enough. Most often this assumption is implied by the assumption that the long-run claim intensity is smaller than the premium rate (*positive loading*). That is, for some $0 < \mu < c$,

$$\lim_{t\to\infty} \frac{Y(t)}{t} = \mu \qquad \text{a.s.} \tag{1.3}$$

In various applications one has to deal, for instance, with losses that occur in a nonhomogeneous way over time. To understand the effect of this phenomenon on the likelihood of ruin, we introduce a stochastic process $(\Delta(t),\ t \geq 0)$, which is a right-continuous nondecreasing stochastic process, satisfying $\Delta(0) = 0$ almost surely, defined on the same probability space as $(Y(t),\ t \geq 0)$. We view $(\Delta(t),\ t \geq 0)$ as a time change; if time runs faster, then losses occur faster as well.

We will assume that

$$\lim_{t\to\infty} \frac{\Delta(t)}{t} = 1 \qquad \text{a.s.} \tag{1.4}$$

This assumption says that, in the long run, the clock given by $\Delta(t)$ runs at the same speed as the natural clock. Clearly, any positive limit different from 1 in (1.4) can be made to fit into our discussion by a suitable modification of the claim process.

Consider the modified ruin probability

$$\psi(u) = P\left(\sup_{t\geq 0}\big(Y\big(\Delta(t)\big) - ct\big) > u\right). \tag{1.5}$$

Certain assumptions will have to be imposed to make sure that $\psi(u) \to 0$ as $u \to \infty$ and, hence, that the ruin is not certain in this modified situation.

An important question is: what is the effect of the time change Δ on the ruin probability? Is the modified ruin probability $\psi(u)$ asymptotically equivalent to the original ruin probability $\psi_0(u)$? Are the two of the same order as u goes to ∞? We will see that the answers to these questions [as well as the actual behavior of $\psi(u)$ for large u if the answers to the preceding questions are negative] depend heavily on certain mixing properties of the time change Δ. Specifically, how fast is the convergence of $\Delta(t)/t$ in (1.4) to 1? We will see that if $\Delta(t)/t$ converges to 1 fast enough, then the effect of the time change on the ruin probability is negligible for large values of initial capital u. On the other hand, if $\Delta(t)/t$ converges to 1 sufficiently slowly, then the modified ruin probability $\psi(u)$ may be of a different order of magnitude than $\psi_0(u)$.

In this article we deal with the so-called heavy-tailed case, which arises when the claim process $(Y(t),\ t \geq 0)$ is heavy tailed. The literature on the ruin probability in the heavy-tailed case is vast; see Embrechts, Klüppelberg and Mikosch (1997) for a discussion of the situation with iid heavy-tailed claims, as well as for numerous additional references, and Asmussen, Schmidli and Schmidt (1999) and Mikosch and Samorodnitsky (2000a, b) for more complicated heavy-tailed claim processes. For our purposes in this paper it is not particularly important, most of the time, what kind of a heavy-tailed claim process we are dealing with, and our main assumption of heavy-tailedness is in terms of the ruin probability itself. Specifically, we will assume that

$$\psi_0(u) \in \operatorname{Reg}(-\beta) \qquad \text{as } u \to \infty \tag{1.6}$$

for some $\beta \geq 0$, where $\operatorname{Reg}(-\beta)$ is the collection of all functions of the type $f(u) = u^{-\beta} L(u)$, with L slowly varying at ∞.

2. Fast and slow mixing of the time change.

One way to measure how fast the average clock rate $\Delta(t)/t$ converges to its limit is by studying the probability

$$g_\varepsilon(u) = P\left(\left|\frac{\Delta(t)}{t} - 1\right| > \varepsilon \text{ for some } t > u\right) \tag{2.1}$$

as $u \to \infty$ for fixed $\varepsilon > 0$. The main point of our first result, Theorem 2.1, is that if $g_\varepsilon(u)$ is of a smaller order than the original ruin probability $\psi_0(u)$, then the effect of the change is negligible. To state the result precisely, we need to introduce some new notation. For $\varepsilon \in \mathbb{R}$ let

$$\psi_{0,\varepsilon}(u) = P\left(\sup_{t\geq 0}\big(Y(t) - c\varepsilon t\big) > u\right); \tag{2.2}$$

clearly, $\psi_{0,1}(u) = \psi_0(u)$. We assume that

$$\lim_{\varepsilon \downarrow 1} \limsup_{u\to\infty} \frac{\psi_0(u)}{\psi_{0,\varepsilon}(u)} = \lim_{\varepsilon \uparrow 1} \liminf_{u\to\infty} \frac{\psi_0(u)}{\psi_{0,\varepsilon}(u)} = 1. \tag{2.3}$$

We also need to assume that ruin does not happen *too soon*. Specifically, assume that

$$\lim_{\delta\downarrow 0}\limsup_{u\to\infty}\frac{P(\sup_{0\le t\le\delta u}(Y(t)-ct)>u)}{\psi_0(u)}=0. \tag{2.4}$$

Conditions (2.3) and (2.4) turn out to hold in virtually all examples of interest when the ruin probability $\psi_0(u)$ is regularly varying. See the following examples.

THEOREM 2.1. *Assume that* (2.3) *and* (2.4) *hold. Under the assumption of heavy tails* (1.6), *assume that, for every* $\varepsilon>0$ *and* $\delta>0$,

$$\lim_{u\to\infty}\frac{g_\varepsilon(\delta u)}{\psi_0(u)}=0. \tag{2.5}$$

Assume, further, either that Δ *is continuous on a set of probability* 1, *or that, for some* $a\ge 0$, *on a set of probability* 1,

$$\text{the process } \{Y(t)+at,\ t\ge 0\} \text{ is eventually nondecreasing.} \tag{2.6}$$

Then

$$\lim_{u\to\infty}\frac{\psi(u)}{\psi_0(u)}=1. \tag{2.7}$$

PROOF. For $t\ge 0$ let

$$\Delta^{-1}(t):=\inf\{s\ge 0:\Delta(s)\ge t\}.$$

It is elementary to check that (2.5) implies that, for every $\varepsilon>0$ and $\delta>0$,

$$\lim_{u\to\infty}\frac{P(|\Delta^{-1}(t)/t-1|>\varepsilon \text{ for some } t>\delta u)}{\psi_0(u)}=0. \tag{2.8}$$

Suppose first that Δ is continuous on a set of probability 1. Then, for every $\delta>0$ and $\varepsilon>0$,

$$\begin{aligned}\psi(u)&\ge P\left(\sup_{t\ge\delta u}\big(Y(t)-c\Delta^{-1}(t)\big)>u\right)\\&\ge P\left(\sup_{t\ge\delta u}\big(Y(t)-c(1+\varepsilon)t\big)>u\right)-P\big(\Delta^{-1}(t)>(1+\varepsilon)t \text{ for some } t>\delta u\big),\end{aligned}$$

and, hence, using (2.8), we conclude that

$$\liminf_{u\to\infty}\frac{\psi(u)}{\psi_0(u)}\ge\liminf_{u\to\infty}\frac{P(\sup_{t\ge\delta u}(Y(t)-c(1+\varepsilon)t)>u)}{\psi_0(u)}.$$

On the other hand,

$$P\left(\sup_{t\ge\delta u}\big(Y(t)-c(1+\varepsilon)t\big)>u\right)\ge\psi_{0,1+\varepsilon}(u)-P\left(\sup_{0\le t\le\delta u}\big(Y(t)-ct\big)>u\right),$$

and so, by (2.4),

$$\liminf_{u\to\infty} \frac{\psi(u)}{\psi_0(u)} \geq \liminf_{u\to\infty} \frac{\psi_{0,1+\varepsilon}(u)}{\psi_0(u)}.$$

Letting $\varepsilon \downarrow 0$ and using (2.3), we see that

$$\liminf_{u\to\infty} \frac{\psi(u)}{\psi_0(u)} \geq 1. \tag{2.9}$$

On the other hand, suppose that (2.6) holds. Write, for $\delta > 0$ and $0 < \varepsilon < 1$,

$$\begin{aligned}
\psi(u) &\geq P\left(\sup_{t\geq\delta u}\Big(\big(Y\big(\Delta(t)\big) + a\Delta(t)\big) - a\Delta(t) - ct\Big) > u\right) \\
&\geq P\left(\sup_{t\geq\delta u}\Big(Y\big((1-\varepsilon)t\big) + a(1-\varepsilon)t - a(1+\varepsilon)t - ct\Big) > u\right) \\
&\quad - P\left(\left|\frac{\Delta(t)}{t} - 1\right| > \varepsilon \text{ for some } t > u\right) \\
&= P\left(\sup_{t\geq(1-\varepsilon)\delta u}\left(Y(t) - \frac{c+2\varepsilon a}{1-\varepsilon}t\right) > u\right) - g_\varepsilon(\delta u),
\end{aligned}$$

and using once again (2.8) and (2.4), we conclude that, for every $\varepsilon > 0$,

$$\liminf_{u\to\infty} \frac{\psi(u)}{\psi_0(u)} \geq \liminf_{u\to\infty} \frac{\psi_{0,(1+2\varepsilon ac^{-1})/(1-\varepsilon)}(u)}{\psi_0(u)},$$

and, letting once again $\varepsilon \downarrow 0$ and using (2.3), we see that (2.9) still holds.

In the other direction, for every $\delta > 0$,

$$\begin{aligned}
\psi(u) &\leq P\left(\sup_{t\geq 0}\big(Y(t) - c\Delta^{-1}(t)\big) > u\right) \\
&\leq P\left(\sup_{0\leq t\leq\delta u}\big(Y(t) - c\Delta^{-1}(t)\big) > u\right) + P\left(\sup_{t\geq\delta u}\big(Y(t) - c\Delta^{-1}(t)\big) > u\right).
\end{aligned} \tag{2.10}$$

Observe that

$$\begin{aligned}
P\left(\sup_{0\leq t\leq\delta u}\big(Y(t) - c\Delta^{-1}(t)\big) > u\right) &\leq P\left(\sup_{0\leq t\leq\delta u} Y(t) > u\right) \\
&\leq P\left(\sup_{0\leq t\leq\delta u}\big(Y(t) - ct\big) > u(1-\delta c)\right),
\end{aligned}$$

and so, using (2.4) and regular variation of ψ_0, we conclude that

$$\lim_{u\to\infty} \frac{P(\sup_{0\leq t\leq\delta u}(Y(t) - c\Delta^{-1}(t)) > u)}{\psi_0(u)} = 0. \tag{2.11}$$

On the other hand, for every $\varepsilon \in (0,1)$,

$$P\left(\sup_{t\geq\delta u}\big(Y(t)-c\Delta^{-1}(t)\big)>u\right)$$
$$\leq P\left(\sup_{t\geq\delta u}\big(Y(t)-c(1-\varepsilon)t\big)>u\right)+P\big(\Delta^{-1}(t)<(1-\varepsilon)t \text{ for some } t>\delta u\big)$$

and, hence, using (2.8), we conclude that

$$\limsup_{u\to\infty}\frac{\psi(u)}{\psi_0(u)}\leq\limsup_{u\to\infty}\frac{\psi_{0,1-\varepsilon}(u)}{\psi_0(u)}.$$

Letting $\varepsilon\downarrow 0$ and using (2.3), we see that

$$\limsup_{u\to\infty}\frac{\psi(u)}{\psi_0(u)}\leq 1. \tag{2.12}$$

Finally, comparing the bounds (2.9) and (2.12), we obtain the statement of the theorem. □

One may suspect that the assumption of continuity of the time change Δ in Theorem 2.1 is superfluous and is only an artifact of our proof. In fact, this assumption *can be removed* in a variety of situations; the alternative assumption (2.6) provides a naturally occurring situation where this is possible. However, Example 2.2 shows that Theorem 2.1 is, in general, *false* without the assumption of continuity of the time change Δ. We also note that the asymptotic upper bound on the ruin probability $\psi(u)$ in Theorem 2.1 does not require the continuity assumption.

EXAMPLE 2.2. Let $X>0$ be a random variable such that $P(X>u)\in \mathrm{Reg}(-\beta)$ as $u\to\infty$ for some $\beta>0$. Define

$$Y(t)=\begin{cases}0, & \text{for } 0\leq t<X,\\ 2X, & \text{for } X\leq t<X+1,\\ 0, & \text{for } t\geq X+1.\end{cases}$$

Let $c=1$. Then

$$\psi_0(u)=P(2X-X>u)=P(X>u)\in\mathrm{Reg}(-\beta)$$

as $u\to\infty$, while, for $\varepsilon\in(0,2)$,

$$\psi_{0,\varepsilon}(u)=P(2X-\varepsilon X>u)=P\big(X>(2-\varepsilon)^{-1}u\big),$$

and so (2.3) holds. Moreover, for every $0<\delta<1$,

$$P\left(\sup_{0\leq t\leq\delta u}\big(Y(t)-ct\big)>u\right)\leq P(X\leq\delta u,\ X>u)=0,$$

and so (2.4) holds as well.

Define

$$\Delta(t) = \begin{cases} t, & \text{for } 0 \le t < X, \\ t+1, & \text{for } t \ge X. \end{cases}$$

Obviously, (1.4) holds. Moreover, for every $\varepsilon > 0$,

$$g_\varepsilon(u) \le P\left(\frac{1}{t} > \varepsilon \text{ for some } t \ge \max(X, u)\right) \le P\left(\frac{1}{u} > \varepsilon\right) = 0$$

for all u large enough. Therefore, (2.5) holds as well. However, $Y(\Delta(t)) = 0$ for all t, so that $\psi(u) = 0$ for all $u > 0$ and (2.7) fails.

Roughly speaking, Theorem 2.1 shows that if the rate of "mixing" of the time change Δ is fast enough, meaning that $\Delta(t)/t$ is not "very far" from 1, measured with respect to the original ruin probability ψ_0, then the ruin probability is not much affected by the time change. The next result can be viewed as a counterpart of this statement: if the time change "mixes" slowly enough, once again, in the context of the original ruin probability, then the ruin probability is affected significantly by the time change. Of course, technical conditions are required in both cases. The speed of mixing in the latter result is measured differently than in the former result, and the way we measure it turns out to be naturally related to the ruin probability after the time change.

Assume the positive loading condition (1.3) and define

$$\psi_1(u) = P\left(\sup_{t\ge 0}\big(\mu\Delta(t) - ct\big) > u\right). \tag{2.13}$$

One can view ψ_1 as a version of the ruin probability ψ in which the stream of the claims $(Y(\Delta(t)),\ t \ge 0)$ is replaced by its long-run average stream $(\mu t,\ t \ge 0)$. Let us introduce the ε-modification of ψ_1. Let

$$\psi_{1,\varepsilon}(u) = P\left(\sup_{t\ge 0}\big(\mu\varepsilon\Delta(t) - ct\big) > u\right). \tag{2.14}$$

Assume that

$$\lim_{\varepsilon\downarrow 1}\liminf_{u\to\infty}\frac{\psi_1(u)}{\psi_{1,\varepsilon}(u)} = \lim_{\varepsilon\uparrow 1}\limsup_{u\to\infty}\frac{\psi_1(u)}{\psi_{1,\varepsilon}(u)} = 1. \tag{2.15}$$

We also assume that, for every $\varepsilon > \mu/c$,

$$\limsup_{u\to\infty}\frac{\psi_{0,\varepsilon}(u)}{\psi_0(u)} < \infty. \tag{2.16}$$

As before, conditions (2.15) and (2.16) turn out to hold in virtually all cases when the ruin probabilities we are dealing with are regularly varying. See the following examples.

We define

$$h_\varepsilon(u) = P\left(\left(\frac{Y(t)}{t} - \mu\right) < -\varepsilon \text{ for some } t > u\right). \tag{2.17}$$

THEOREM 2.3. *Assume that* (2.15) *and* (2.16) *hold. Assume, further, either that the processes* $(Y(t),\ t\ge 0)$ *and* $(\Delta(t),\ t\ge 0)$ *are independent or that, for any* $\varepsilon>0$ *and* $\delta>0$,

$$\lim_{u\to\infty}\frac{h_\varepsilon(\delta u)}{\psi_1(u)}=0. \tag{2.18}$$

If

$$\lim_{u\to\infty}\frac{\psi_0(u)}{\psi_1(u)}=0, \tag{2.19}$$

then

$$\lim_{u\to\infty}\frac{\psi(u)}{\psi_1(u)}=1. \tag{2.20}$$

PROOF. Define, for $0<\varepsilon<1$,

$$T_{\varepsilon,u}=\inf\{t\ge 0:\mu\varepsilon\Delta(t)-ct>u\}.$$

It is clear that $T_{\varepsilon,u}\to\infty$ with probability 1 as $u\to\infty$. We have

$$\psi(u)\ge P\left(T_{\varepsilon,u}<\infty,\ Y(s)-\varepsilon\mu s\ge 0 \text{ for all } s\ge\frac{u}{\varepsilon\mu}\right). \tag{2.21}$$

If $(Y(t),\ t\ge 0)$ and $(\Delta(t),\ t\ge 0)$ are independent, then it follows from (2.21) that

$$\psi(u)\ge\psi_{1,\varepsilon}(u)\bigl(1-h_{(1-\varepsilon)\mu}(u/(\varepsilon\mu))\bigr),$$

and we conclude by (1.3) that

$$\liminf_{u\to\infty}\frac{\psi(u)}{\psi_{1,\varepsilon}(u)}\ge 1.$$

Letting $\varepsilon\to 1$ and using (2.15), we conclude that

$$\liminf_{u\to\infty}\frac{\psi(u)}{\psi_1(u)}\ge 1. \tag{2.22}$$

On the other hand, it also follows from (2.21) that

$$\psi(u)\ge\psi_{1,\varepsilon}(u)-h_{(1-\varepsilon)\mu}(u/(\varepsilon\mu)),$$

and now (2.22) follows as before, using (2.18) (instead of independence) and (2.15).

In the other direction, for $\varepsilon > 1$ we obtain, from (2.21),

$$\begin{aligned}\psi(u) &\le P\Big(\sup_{t\ge 0}\big(\mu\varepsilon\Delta(t) - ct\big) > u\Big)\\ &\quad + P\Big(\sup_{t\ge 0}\big(Y\big(\Delta(t)\big) - ct\big) > u,\ \mu\varepsilon\Delta(t) - ct \le u \text{ for all } t \ge 0\Big)\\ &\le \psi_{1,\varepsilon}(u) + P\Big(\sup_{t\ge 0}\big(Y\big(\Delta(t)\big) - \varepsilon^{1/2}\mu\Delta(t)\big) > \Big(1 - \frac{1}{\varepsilon^{1/3}}\Big)u\Big)\\ &\le \psi_{1,\varepsilon}(u) + \psi_{0,\varepsilon^{1/2}\mu c^{-1}}\Big(\Big(1 - \frac{1}{\varepsilon^{1/3}}\Big)u\Big).\end{aligned} \tag{2.23}$$

To check the second inequality above, note the simple fact that, for $\varepsilon > 1$, $A, B, u > 0$, the inequality $\varepsilon A \le B + u$ implies the inequality $B \ge \varepsilon^{1/2}A - \varepsilon^{-1/3}u$. Hence, if for some $t > 0$ we have $Y(\Delta(t)) - ct > u$ and $\mu\varepsilon\Delta(t) - ct \le u$, then, denoting $A = \mu\Delta(t)$ and $B = ct$, we see that

$$u + ct = u + B \ge u + \varepsilon^{1/2}A - \varepsilon^{-1/3}u = \varepsilon^{1/2}\mu\Delta(t) + u\Big(1 - \frac{1}{\varepsilon^{1/3}}\Big).$$

Using (2.16) and (2.19), we see that, for all $\varepsilon > 1$,

$$\limsup_{u\to\infty} \frac{\psi(u)}{\psi_{1,\varepsilon}(u)} \le 1.$$

Finally, letting $\varepsilon \downarrow 1$, we conclude that

$$\limsup_{u\to\infty} \frac{\psi(u)}{\psi_1(u)} \le 1,$$

which, together with (2.22), completes the proof of the theorem. □

EXAMPLE 2.4. For $0 < H < 1$ let $\{W_H(t),\ t \ge 0\}$ be fractional Brownian motion with parameter H—the H-self-similar Gaussian process with stationary increments; see, for example, Samorodnitsky and Taqqu (1994). Define, for a $\sigma > 0$,

$$\Delta(t) = t + \sigma \sup_{0\le x\le t} |W_H(x)|, \qquad t \ge 0. \tag{2.24}$$

This time change is continuous on a set of probability 1, and (1.4) holds. Moreover, this time change mixes very fast. Indeed, we have, for $\varepsilon > 0$ by the H-self-similarity of W_H,

$$\begin{aligned}g_\varepsilon(u) &= P\Big(\Big|\frac{\Delta(t)}{t} - 1\Big| > \varepsilon \text{ for some } t > u\Big)\\ &= P\Big(\sup_{0\le x\le t} |W_H(x)| > \frac{\varepsilon}{\sigma} t u^{1-H} \text{ for some } t > 1\Big)\\ &\le P\Big(\sup_{0\le x\le 1} |W_H(x)| > \frac{\varepsilon}{\sigma} u^{1-H}\Big) + P\Big(\sup_{x\ge 1} \frac{|W_H(x)|}{x} > \frac{\varepsilon}{\sigma} u^{1-H}\Big).\end{aligned}$$

However, both $\{W_H(t),\ 0 \le t \le 1\}$ and $\{W_H(t)/t,\ t \ge 1\}$ are bounded Gaussian processes. Hence, for some $C_1, C_2 > 0$,

$$g_\varepsilon(u) \le C_1 \exp\left\{-\frac{\varepsilon^2}{2C_2\sigma^2}u^{2(1-H)}\right\}$$

[see Adler (1990)] and so (2.5) holds for every $\delta > 0$ under the assumption (1.6) of heavy tails. This fast mixing of W_H is somewhat surprising for $H > 1/2$ because the increments of fractional Brownian motion are long range dependent in that range of H. See, for example, Beran (1994).

EXAMPLE 2.5. Let

$$Y(t) = Y_0(t) + \mu t, \qquad t \ge 0, \tag{2.25}$$

where $(Y_0(t),\ t \ge 0)$ is an H-self-similar symmetric α-stable (SαS) process with stationary increments with $0 < H < 1$ and $1 < \alpha < 2$. We refer the reader to Samorodnitsky and Taqqu (1994) for more details on both α-stable processes, their representations discussed below and self-similarity. We assume that $(Y_0(t),\ t \ge 0)$ can be represented in the form

$$Y_0(t) = \int_S f_t(x)M(dx), \qquad t \ge 0, \tag{2.26}$$

where $(S, \mathcal{S})$ is a measurable space, M an SαS random measure on this space with control measure m and $f_t \in L^\alpha(m)$ for all $t \ge 0$, such that

$$\int_S \sup_{t\ge0}\left|\frac{f_t(x)}{1+t}\right|^\alpha m(dx) < \infty. \tag{2.27}$$

It is known that (2.26) with (2.27) holds for all H-self-similar SαS process with stationary increments with $1 < \alpha < 2$ if $1/\alpha < H < 1$, whereas, in the case $0 < H \le 1/\alpha$, (2.26) with (2.27) is an assumption that holds in some cases and does not hold in other cases; see Section 12.4 in Samorodnitsky and Taqqu (1994).

Observe that, by the self-similarity of Y_0,

$$\psi_0(u) = P\left(\sup_{t\ge0}\frac{Y_0(t)}{1+(c-\mu)t} > u^{1-H}\right) \sim C_\alpha K u^{-\alpha(1-H)}$$

as $u \to \infty$, where C_α is a finite positive constant that depends only on α and

$$K = \int_S \sup_{t\ge0}\left|\frac{f_t(x)}{1+(c-\mu)t}\right|^\alpha m(dx),$$

and so (1.6) holds. Similarly, for all $\varepsilon > \mu/c$,

$$\psi_{0,\varepsilon}(u) \sim C_\alpha K_\varepsilon u^{-\alpha(1-H)}$$

as $u \to \infty$, with

$$K_\varepsilon = \int_S \sup_{t \geq 0} \left| \frac{f_t(x)}{1 + (\varepsilon c - \mu)t} \right|^\alpha m(dx),$$

demonstrating that (2.3) holds as well. Furthermore, for any $\delta > 0$,

$$P\left(\sup_{0 \leq t \leq \delta u} \big(Y(t) - ct\big) > u\right) = P\left(\sup_{0 \leq t \leq \delta} \frac{Y_0(t)}{1 + (c - \mu)t} > u^{1-H}\right)$$
$$\sim C_\alpha k(\delta) u^{-\alpha(1-H)},$$

where

$$k(\delta) = \int_S \sup_{0 \leq t \leq \delta} \left| \frac{f_t(x)}{1 + (c - \mu)t} \right|^\alpha m(dx).$$

Since $k(\delta) \to 0$ as $\delta \to 0$, the assumption (2.4) holds.

No nondegenerate processes Y of this kind will satisfy (2.6) so in order to apply Theorem 2.1 one will have to assume continuity of the time change Δ as, say, in Example 2.4.

EXAMPLE 2.6. Let Y be as in (2.25), but this time $(Y_0(t),\ t \geq 0)$ is a zero mean Lévy process with Lévy measure ρ. We assume that

$$\rho\big((u, \infty)\big) \in \mathrm{Reg}(-\alpha) \qquad \text{as } u \to \infty \tag{2.28}$$

for some $\alpha > 1$, and that, for some $C > 0$,

$$\rho\big((-\infty, u]\big) \leq C\rho\big((u, \infty)\big) \tag{2.29}$$

for all $u \geq 1$. It follows from Theorem 5.3 in Braverman, Mikosch and Samorodnitsky (2000) that, for all $\varepsilon > \mu/c$,

$$\psi_{0,\varepsilon}(u) \sim K_\varepsilon u \rho\big((u, \infty)\big)$$

as $u \to \infty$, where

$$K_\varepsilon = \frac{1}{(\alpha - 1)(\varepsilon c - \mu)}.$$

Therefore, (1.6) holds with $\beta = \alpha - 1$, and (2.3) holds as well. On the other hand, for all $\delta > 0$,

$$P\left(\sup_{0 \leq t \leq \delta u} \big(Y(t) - ct\big) > u\right) \leq P\left(\sup_{0 \leq t \leq \delta u} Y_0(t) > u\right)$$
$$\leq P\left(\sum_{j=1}^{\lceil u \rceil} \sup_{(j-1)\delta \leq t \leq j\delta} \big(Y_0(t) - Y_0\big((j-1)\delta\big) > u\big)\right).$$

Here $\lceil a \rceil$ is the smallest integer greater than or equal to a. Since

$$P\Big(\sup_{0\le t\le\delta} Y_0(t) > u\Big) \sim \delta\rho\big((u,\infty)\big)$$

as $u \to \infty$ [see, e.g., Embrechts, Goldie and Veraverbeke (1979)], we can then apply the usual large-deviation results [see, e.g., Nagaev (1979)] to see that

$$P\Big(\sum_{j=1}^{\lceil u\rceil} \sup_{(j-1)\delta\le t\le j\delta} \big(Y_0(t) - Y_0\big((j-1)\delta\big) > u\big)\Big)$$

$$\sim uP\Big(\sup_{0\le t\le\delta} Y_0(t) > u\Big) \sim \delta u\rho\big((u,\infty)\big)$$

as $u \to \infty$. Therefore, the assumption (2.4) holds.

3. Mixing of Markov chain switching models. A very important and widely used class of stochastic models in almost every area of application is that of Markov switching (Markov modulated, Markov renewal) models. We refer the reader to Çinlar (1975) for a general theory of such models. In the context of insurance, it is natural to consider a class of time change processes Δ in which time runs at different rates in different time intervals, depending on the state of a certain Markov chain, and the Markov chain stays in each state a random amount of time, with a distribution that depends on the state. It turns out that this class of models is very flexible and mixing in this class of models can be either fast or slow. The speed of mixing in this class of models is our subject in this section.

Here is the formal setup. Let $(Z_n,\ n \ge 0)$ be an irreducible Markov chain with a finite state space $\{1,\dots,K\}$, transition matrix P and stationary probabilities $\pi_1,\dots,\pi_K$. Let $F_j,\ j = 1,\dots,K$, be probability distributions on $(0,\infty)$. Note that F_j is the law of the holding time the system spends in state j, whose mean μ_j is assumed to be finite for $j = 1,\dots,K$. We denote $\overline{\mu} = \sum_{j=1}^K \mu_j\pi_j$.

Let $(H_i^{(j)},\ i \ge 1),\ j = 1,\dots,K$, be K independent sequences of iid random variables such that $H_i^{(j)}$ has the distribution F_j and describes the length of the ith sojourn in state j. Transitions from state to state are governed by the transition matrix P, and, given the present state of the Markov chain, its next state is independent of the sojourn times sequences.

Let $r_1, r_2,\dots,r_K$ be nonnegative numbers such that

$$\frac{\sum_{j=1}^K r_j\mu_j\pi_j}{\sum_{j=1}^K \mu_j\pi_j} = 1. \tag{3.1}$$

We define the time change Δ to be an absolutely continuous process with

$$\Delta(0) = 0, \qquad \frac{d\Delta}{dt}(t) = r_j \quad \text{if at time } t \text{ the Markov chain is in state } j. \tag{3.2}$$

Clearly, a complete definition of the time change Δ requires a specification of the initial distribution $p_1, \dots, p_K$ of the Markov chain. This initial distribution, however, has no effect on the speed of mixing of the time change.

A simple renewal argument shows that (3.1) guarantees (1.4). It is our goal to show how the holding time distributions F_j, $j = 1, \dots, K$, and the parameters of the Markov chain affect the rate at which the average clock $\Delta(t)/t$ converges to 1.

Our main assumption is that of *heavy-tailed* holding times. Specifically, we assume that there is a distribution F on $(0, \infty)$ such that

$$\overline{F}(x) \in \mathrm{Reg}(-\gamma) \quad \text{as } x \to \infty \quad \text{and} \quad \lim_{x\to\infty} \frac{\overline{F}_j(x)}{\overline{F}(x)} = \theta_j \quad \text{for } j = 1, \dots, K \tag{3.3}$$

for some $\gamma > 1$ and some $\theta_1, \dots, \theta_K \in [0, \infty)$, not all of which are equal to 0.

Let $\varepsilon > 0$. The following theorem addresses the rate of decay of the probability $g_\varepsilon(u)$ in (2.1) as $u \to \infty$. We mention that ideas similar to those we use in the proof of this theorem can also be used to obtain the asymptotic behavior of various other probabilities related to the time change process Δ, for example, the probabilities $\psi_1(u)$ and $\psi_{1,\varepsilon}(u)$ in (2.13) and (2.14). We put some of the auxiliary statements as lemmas at the end of this section.

THEOREM 3.1. *Let $\varepsilon > 0$ be such that $\{j = 1, \dots, K : |r_j - 1| = \varepsilon\} = \varnothing$ and let*

$$\begin{aligned} J_+(\varepsilon) &= \{j = 1, \dots, K : r_j > 1 + \varepsilon\}, \\ J_-(\varepsilon) &= \{j = 1, \dots, K : r_j < 1 - \varepsilon\}. \end{aligned} \tag{3.4}$$

Then

$$\begin{aligned} \lim_{u\to\infty} \frac{g_\varepsilon(u)}{u\overline{F}(u)} = \frac{1}{\varepsilon^\gamma \overline{\mu}} \Bigg[&\sum_{j \in J_+(\varepsilon)} \theta_j \pi_j (r_j - 1 - \varepsilon)(r_j - 1)^{\gamma - 1} \\ &+ \sum_{j \in J_-(\varepsilon)} \theta_j \pi_j (1 - r_j - \varepsilon)(1 - r_j)^{\gamma - 1} \Bigg]. \end{aligned} \tag{3.5}$$

REMARK 3.2. If $\theta_j > 0$ for at least one j such that $r_j \neq 1$, then it follows immediately from Theorem 3.1 that $g_\varepsilon(u)$ is regularly varying with exponent $\gamma - 1$ as $u \to \infty$ for all $\varepsilon > 0$ small enough.

REMARK 3.3. The conclusion of the theorem is independent of the initial state or, indeed, of the initial distribution of the underlying Markov chain. Where it is convenient, we will denote in the following proof by j_0 the initial state of the Markov chain and assume it to be nonrandom. In most cases we will not use the explicit notation P_{j_0}, E_{j_0}; the initial state will be kept implicit in most cases.

REMARK 3.4. The proof of Theorem 3.1 is fairly technical. Its idea is, however, very simple. Under the assumption (3.3) of heavy tails, the event $\{|\Delta(t)/t - 1| > \varepsilon \text{ for some } t > u\}$, if it occurs for a large u, is caused, most likely, by a single long holding time, either with a state $j \in J_+(\varepsilon)$ or with $j \in J_-(\varepsilon)$. The reader can easily realize what is happening by checking two possibilities: the long holding time can end either before time u or after time u. In both cases one figures out just how long this long holding time has to be by pretending that before the start of the long holding time and its end all the random quantities are about equal to their averages. The technical details in the proof are required to justify the above statements. We provide most of the details, but try to avoid duplication of the argument.

PROOF OF THEOREM 3.1. Denote

$$E_\varepsilon^+(u) = \left\{\frac{\Delta(t)}{t} - 1 > \varepsilon \text{ for some } t > u\right\}$$

and

$$E_\varepsilon^-(u) = \left\{\frac{\Delta(t)}{t} - 1 < -\varepsilon \text{ for some } t > u\right\}$$

and let, for $\tau > 0$ small enough (we will specify just how small τ has to be later),

$$B_\tau(u) = \Big\{\text{for at most one pair } (i, j) \in \{1, 2, \dots\} \times \{1, \dots, K\},\ H_i^{(j)} > (u + i)\tau\Big\}$$
$$\supset \big(B_\tau(u) \cap B_\tau^{(\varepsilon,+)}(u)\big) \cup \big(B_\tau(u) \cap B_\tau^{(\varepsilon,-)}(u)\big),$$

where

$$B_\tau^{(\varepsilon,+)}(u) = \Big\{\text{for exactly one pair } (i, j) \in \{1, 2, \dots\} \times J_+(\varepsilon),\ H_i^{(j)} > (u+)\tau\Big\}$$

and

$$B_\tau^{(\varepsilon,-)}(u) = \Big\{\text{for exactly one pair } (i, j) \in \{1, 2, \dots\} \times J_-(\varepsilon),\ H_i^{(j)} > (u+)\tau\Big\}.$$

We will show first that

$$\liminf_{u\to\infty} \frac{P(E_\varepsilon^+(u) \cap B_\tau(u) \cap B_\tau^{(\varepsilon,+)}(u))}{u\overline{F}(u)} \geq \frac{1}{\varepsilon^\gamma \overline{\mu}} \sum_{j\in J_+(\varepsilon)} \theta_j \pi_j (r_j - 1 - \varepsilon)(r_j - 1)^{\gamma-1} \tag{3.6}$$

and that

$$\liminf_{u\to\infty} \frac{P(E_\varepsilon^-(u) \cap B_\tau(u) \cap B_\tau^{(\varepsilon,-)}(u))}{u\overline{F}(u)} \geq \frac{1}{\varepsilon^\gamma \overline{\mu}} \sum_{j\in J_-(\varepsilon)} \theta_j \pi_j (1 - r_j - \varepsilon)(1 - r_j)^{\gamma-1}. \tag{3.7}$$

Since the sets $B_\tau(u) \cap B_\tau^{(\varepsilon,+)}(u)$ and $B_\tau(u) \cap B_\tau^{(\varepsilon,-)}(u)$ are disjoint, this will imply that

$$\liminf_{u\to\infty} \frac{g_\varepsilon(u)}{u\overline{F}(u)} \geq \frac{1}{\varepsilon^\gamma \overline{\mu}} \Bigg[\sum_{j\in J_+(\varepsilon)} \theta_j \pi_j (r_j - 1 - \varepsilon)(r_j - 1)^{\gamma-1} + \sum_{j\in J_-(\varepsilon)} \theta_j \pi_j (1 - r_j - \varepsilon)(1 - r_j)^{\gamma-1} \Bigg].$$

Note also that the statements (3.6) and (3.7) are of the same nature, and so we really need to prove only one of the two. We choose to prove (3.6), and this is what we proceed to do now.

For $j = 1, \ldots, K$ and $i = 1, 2, \ldots$, let $T_i^{(j)}$ be the starting time of the ith sojourn in state j of the underlying process. For $j \in J_+(\varepsilon)$ and $i = 1, 2, \ldots$, consider the events

$$E^{(1)}_{\varepsilon,\tau,i,j}(u) = \Big\{ T_i^{(j)} + H_i^{(j)} \geq u,\ r_j H_i^{(j)} + \Delta\big(T_i^{(j)}\big) > \big(T_i^{(j)} + H_i^{(j)}\big)(1+\varepsilon),\ H_i^{(j)} > (u+i)\tau \Big\} \tag{3.8}$$

and

$$E^{(2)}_{\varepsilon,\tau,i,j}(u) = \Big\{ T_i^{(j)} + H_i^{(j)} < u,\ r_j H_i^{(j)} + \Delta\big(T_i^{(j)}\big) + \Big(\Delta(u) - \Delta\big(T_i^{(j)} + H_i^{(j)}\big)\Big) > (1+\varepsilon)u,\ H_i^{(j)} > (u+i)\tau \Big\}. \tag{3.9}$$

Notice that

$$\begin{aligned} E_\varepsilon^+(u) \cap B_\tau(u) \cap B_\tau^{(\varepsilon,+)}(u) &\supset \bigg(\bigcup_{j\in J_+(\varepsilon)} \bigcup_{i=1}^{\infty} \Big(E^{(1)}_{\varepsilon,\tau,i,j}(u) \cap B_\tau(u) \Big) \bigg) \\ &\quad \cup \bigg(\bigcup_{j\in J_+(\varepsilon)} \bigcup_{i=1}^{\infty} \Big(E^{(2)}_{\varepsilon,\tau,i,j}(u) \cap B_\tau(u) \Big) \bigg) \\ &:= E^{1,+}_{\varepsilon,\tau}(u) \cup E^{2,+}_{\varepsilon,\tau}(u). \end{aligned} \tag{3.10}$$

Note that all the events in the above unions are disjoint and, in particular, the events $E^{1,+}_{\varepsilon,\tau}(u)$ and $E^{2,+}_{\varepsilon,\tau}(u)$ are disjoint.

We start by estimating the probability of the event $E^{(1)}_{\varepsilon,\tau,i,j}(u) \cap B_\tau(u)$. Assume that $\theta_j > 0$. For $\delta > 0$ we have

$$P\Big(E^{(1)}_{\varepsilon,\tau,i,j}(u) \cap B_\tau(u)\Big) \geq P\Big(E^{(1)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Big), \tag{3.11}$$

where

$$E^{(1)}_{\varepsilon,\tau,\delta,i,j}(u) = \Big\{ H^{(j)}_i > \max\Big(u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big),\ i\frac{\varepsilon(\overline{\mu}/\pi_j) + \varepsilon\delta + 2\delta}{r_j - 1 - \varepsilon},\ (u+i)\tau\Big),$$

$$i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big) \le T^{(j)}_i \le i\Big(\frac{\overline{\mu}}{\pi_j} + \delta\Big),\ \Delta(T^{(j)}_i) > i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big\},$$

$j \in J_+(\varepsilon)$, $i = 1, 2, \ldots$. We now estimate the probability in the right-hand side of (3.11) in different ranges of i. Denote, for $j \in J_+(\varepsilon)$,

$$s^+_j(\varepsilon) = \frac{\varepsilon}{r_j - 1 - \varepsilon} > 0 \tag{3.12}$$

and

$$D^{+,\varepsilon}_j(\delta) = \delta\Big(\frac{2\varepsilon + 3 - r_j}{r_j - 1 - \varepsilon}\Big). \tag{3.13}$$

We consider only $\delta > 0$ so small that

$$(1 + s^+_j(\varepsilon))\frac{\overline{\mu}}{\pi_j} + D^{+,\varepsilon}_j(\delta) > \Big(1 + \frac{1}{2}s^+_j(\varepsilon)\Big)\frac{\overline{\mu}}{\pi_j} > 0.$$

Let $\lambda > 0$ be a small positive number. The first range of i we consider is

$$\lambda u \le i \le \Big(\frac{1}{(\overline{\mu}/\pi_j)(1 + s^+_j(\varepsilon)) + D^{+,\varepsilon}_j(\delta)}\Big)u. \tag{3.14}$$

Notice that, for some $\tau_1(\varepsilon) > 0$, if $0 < \tau < \tau_1(\varepsilon)$, then, in our range of i,

$$u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big) \ge \max\Big(i\frac{\varepsilon(\overline{\mu}/\pi_j) + \varepsilon\delta + 2\delta}{r_j - 1 - \varepsilon},\ (u+i)\tau\Big).$$

Therefore,

$$\begin{aligned} P\Big(E^{(1)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Big) &\\ \ge P\Big(H^{(j)}_i > u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big) &\\ - P\Big(T^{(j)}_i \le i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big) - P\Big(\Delta(T^{(j)}_i) \le i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big) &\\ - P\Big(T^{(j)}_i > i\Big(\frac{\overline{\mu}}{\pi_j} + \delta\Big),\ H^{(j)}_i > (u+i)\tau\Big) &\\ - P\Big(\{H^{(j)}_i > (u+i)\tau\} \cap \big(B_\tau(u)\big)^c\Big) &\\ := \overline{F}_j\Big(u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big) - \sum_{l=1}^{4} e_{l,i,j}(u). & \end{aligned} \tag{3.15}$$

By Lemma 3.5 and (3.14), we have

$$e_{l,i,j}(u) \le C_1^{(j)} e^{-C_2^{(j)} i} \le C_1^{(j)} e^{-C_2^{(j)} \lambda u} \tag{3.16}$$

for $l = 1, 2$. Now, by the ergodic theorem, for every $j = 1, \dots, K$,

$$\frac{T_i^{(j)}}{i} \to \frac{\overline{\mu}}{\pi_j} \qquad \text{a.s. as } i \to \infty.$$

Therefore, for all u large enough and, hence, i large enough,

$$\begin{aligned} e_{3,i,j}(u) &= P\Big(T_i^{(j)} > i\Big(\frac{\overline{\mu}}{\pi_j} + \delta\Big)\Big) P\Big(H_i^{(j)} > (u+i)\tau\Big) \\ &\le \delta \overline{F}\Big(u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big), \end{aligned} \tag{3.17}$$

where we have used regular variation of $\overline{F}$, which also shows that, for all u large enough,

$$\begin{aligned} e_{4,i,j}(u) &\le P\Big(H_i^{(j)} > (u+i)\tau\Big) \sum_{m=1}^{\infty} \sum_{k=1}^{K} P\Big(H_m^{(k)} > (u+m)\tau\Big) \\ &\le \delta \overline{F}\Big(u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big). \end{aligned} \tag{3.18}$$

We conclude by (3.16)–(3.18) that, for all u large enough and all i in the range (3.14),

$$P\Big(E^{(1)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Big) \ge (\theta_j - 5\delta)\overline{F}\Big(u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big)\Big). \tag{3.19}$$

We next consider i in the range

$$i > \Big(\frac{1}{(\overline{\mu}/\pi_j)(1 + s_j^{+}(\varepsilon)) + D_j^{+,\varepsilon}(\delta)}\Big) u. \tag{3.20}$$

Notice that, for some $\tau_2(\varepsilon) > 0$, if $0 < \tau < \tau_2(\varepsilon)$, then, in our range of i,

$$i \frac{\varepsilon(\overline{\mu}/\pi_j) + \varepsilon\delta + 2\delta}{r_j - 1 - \varepsilon} \ge \max\Big(u - i\Big(\frac{\overline{\mu}}{\pi_j} - \delta\Big),\ (u+i)\tau\Big).$$

Therefore, we can write, as in (3.15),

$$P\Big(E^{(1)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Big) \ge \overline{F}_j\Big(i \frac{\varepsilon(\overline{\mu}/\pi_j) + \varepsilon\delta + 2\delta}{r_j - 1 - \varepsilon}\Big) - \sum_{l=1}^{4} e_{l,i,j}(u),$$

and the same argument as above shows that, for all u large enough and all i in the range (3.20),

$$P\Big(E^{(1)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Big) \ge (\theta_j - 5\delta)\overline{F}\Big(i \frac{\varepsilon(\overline{\mu}/\pi_j) + \varepsilon\delta + 2\delta}{r_j - 1 - \varepsilon}\Big). \tag{3.21}$$

We conclude that, for all u large enough,

$$
\begin{aligned}
&P\big(E^{1,+}_{\varepsilon,\tau}(u)\big)\\
&\quad\geq \sum_{j\in J_+(\varepsilon)}\sum_{i\geq\lambda u} P\big(E^{(1)}_{\varepsilon,\tau,i,j}(u)\cap B_\tau(u)\big)\\
&\quad\geq \sum_{j\in J_+(\varepsilon)}(\theta_j-5\delta)\left(\sum_{\lambda u\leq i\leq u((\overline{\mu}/\pi_j)(1+s_j^+(\varepsilon))+D_j^{+,\varepsilon}(\delta))^{-1}} \overline{F}\left(u-i\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)\right)\right.\\
&\qquad\qquad\left.+\sum_{i>u((\overline{\mu}/\pi_j)(1+s_j^+(\varepsilon))+D_j^{+,\varepsilon}(\delta))^{-1}} \overline{F}\left(i\frac{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta}{r_j-1-\varepsilon}\right)\right)\\
&\quad:= \sum_{j\in J_+(\varepsilon)}(\theta_j-5\delta)\big(S_{1,j}(u)+S_{2,j}(u)\big).
\end{aligned}
\tag{3.22}
$$

Notice that, for u large enough,

$$
\begin{aligned}
S_{1,j}(u)&\geq\int_{\lambda u+2}^{u(\overline{\mu}(1+s_j^+(\varepsilon))\pi_j^{-1}+D_j^{+,\varepsilon}(\delta))^{-1}-1}\overline{F}\left(u-x\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)\right)dx\\
&\geq\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)^{-1}\int_{u(1-(\overline{\mu}/\pi_j-2\delta)(\overline{\mu}(1+s_j^+(\varepsilon))\pi_j^{-1}+D_j^{+,\varepsilon}(\delta))^{-1})}^{u(1-\lambda(\overline{\mu}/\pi_j-2\delta))}\overline{F}(x)\,dx\\
&\sim\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)^{-1}\frac{1}{\gamma-1}\left[u\left(1-\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\left(\frac{\overline{\mu}}{\pi_j}\big(1+s_j^+(\varepsilon)\big)+D_j^{+,\varepsilon}(\delta)\right)^{-1}\right)\right.\\
&\qquad\qquad\times\overline{F}\left(u\left(1-\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\right.\right.\\
&\qquad\qquad\qquad\left.\left.\times\left(\frac{\overline{\mu}}{\pi_j}\big(1+s_j^+(\varepsilon)\big)+D_j^{+,\varepsilon}(\delta)\right)^{-1}\right)\right)\\
&\qquad\qquad\left.-u\left(1-\lambda\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\right)\overline{F}\left(u\left(1-\lambda\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\right)\right)\right]\\
&\sim\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)^{-1}\frac{1}{\gamma-1}u\overline{F}(u)\\
&\quad\times\left[\left(1-\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\left(\frac{\overline{\mu}}{\pi_j}\big(1+s_j^+(\varepsilon)\big)+D_j^{+,\varepsilon}(\delta)\right)^{-1}\right)^{-(\gamma-1)}\right.\\
&\qquad\qquad\left.-\left(1-\lambda\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\right)^{-(\gamma-1)}\right]
\end{aligned}
\tag{3.23}
$$

as $u \to \infty$, by the regular variation of $\overline{F}$, where we used Karamata's theorem [see, e.g., Theorem 0.6 in Resnick (1987)]. Similarly, for u large enough,

$$
\begin{aligned}
S_{2,j}(u) &\geq \int_{u(\overline{\mu}(1+s_j^{+}(\varepsilon))\pi_j^{-1}+D_j^{+,\varepsilon}(\delta))^{-1}+2}^{\infty} \overline{F}\left(x \frac{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta}{r_j-1-\varepsilon}\right) dx \\
&\geq \frac{r_j-1-\varepsilon}{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta} \\
&\quad \times \int_{u(\overline{\mu}(1+s_j^{+}(\varepsilon))\pi_j^{-1}+D_j^{+,\varepsilon}(\delta))^{-1}(\varepsilon\overline{\mu}\pi_j^{-1}+\varepsilon\delta+3\delta)(r_j-1-\varepsilon)^{-1}}^{\infty} \overline{F}(x)\, dx \\
&\sim \frac{r_j-1-\varepsilon}{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta} \frac{1}{\gamma-1} u \left(\frac{\overline{\mu}}{\pi_j}\left(1+s_j^{+}(\varepsilon)\right)+D_j^{+,\varepsilon}(\delta)\right)^{-1} \\
&\quad \times \frac{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta}{r_j-1-\varepsilon} \\
&\quad \times \overline{F}\left(u\left(\frac{\overline{\mu}}{\pi_j}\left(1+s_j^{+}(\varepsilon)\right)+D_j^{+,\varepsilon}(\delta)\right)^{-1} \frac{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta}{r_j-1-\varepsilon}\right) \\
&\sim \frac{r_j-1-\varepsilon}{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta} \frac{1}{\gamma-1} u \overline{F}(u) \\
&\quad \times \left(\left(\frac{\overline{\mu}}{\pi_j}\left(1+s_j^{+}(\varepsilon)\right)+D_j^{+,\varepsilon}(\delta)\right)^{-1} \frac{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta}{r_j-1-\varepsilon}\right)^{-(\gamma-1)}
\end{aligned}
\tag{3.24}
$$

as $u \to \infty$, once again by the regular variation of $\overline{F}$ and Karamata's theorem.

We conclude by (3.22)–(3.24) that, for all $\delta > 0$ and $\lambda > 0$ small enough,

$$
\begin{aligned}
&\liminf_{u\to\infty} \frac{P(E_{\varepsilon,\tau}^{1,+}(u))}{u\overline{F}(u)} \\
&\quad \geq \frac{1}{\gamma-1} \sum_{j\in J_{+}(\varepsilon)} (\theta_j - 5\delta) \\
&\qquad \times \left\{ \left(\frac{\overline{\mu}}{\pi_j}-\delta\right)^{-1} \left[\left(1-\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\left(\frac{\overline{\mu}}{\pi_j}\left(1+s_j^{+}(\varepsilon)\right)+D_j^{+,\varepsilon}(\delta)\right)^{-1}\right)^{-(\gamma-1)} \right.\right. \\
&\qquad\qquad \left. - \left(1-\lambda\left(\frac{\overline{\mu}}{\pi_j}-2\delta\right)\right)^{-(\gamma-1)} \right] + \frac{r_j-1-\varepsilon}{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta} \\
&\qquad\qquad \left. \times \left(\left(\frac{\overline{\mu}}{\pi_j}\left(1+s_j^{+}(\varepsilon)\right)+D_j^{+,\varepsilon}(\delta)\right)^{-1} \frac{\varepsilon(\overline{\mu}/\pi_j)+\varepsilon\delta+2\delta}{r_j-1-\varepsilon}\right)^{-(\gamma-1)} \right\}.
\end{aligned}
$$

Letting $\delta \to 0$ and $\lambda \to 0$, we conclude that

$$\liminf_{u\to\infty} \frac{P(E^{1,+}_{\varepsilon,\tau}(u))}{u\overline{F}(u)} \geq \frac{1}{\overline{\mu}(\gamma-1)} \sum_{j\in J_+(\varepsilon)} \theta_j \pi_j \biggl(\biggl(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\biggr)^{-\gamma} - 1\biggr). \tag{3.25}$$

Notice that (3.25) has been proved under the assumption that $\theta_j > 0$ for every $j \in J_+(\varepsilon)$, but it is entirely obvious that if $\theta_j = 0$ for some $j \in J_+(\varepsilon)$ then (3.25) still follows, with the sum on its right-hand side having appropriately fewer nonzero terms.

We proceed with estimating, in a similar manner, the probability of $E^{(2)}_{\varepsilon,\tau,i,j}(u)$ in (3.10). Concentrating on the event $E^{(2)}_{\varepsilon,\tau,i,j}(u) \cap B_\tau(u)$ for $j \in J_+(\varepsilon)$, we still may and will assume that $\theta_j > 0$. For $\delta > 0$ we have, as before,

$$P\Bigl(E^{(2)}_{\varepsilon,\tau,i,j}(u) \cap B_\tau(u)\Bigr) \geq P\Bigl(E^{(2)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Bigr), \tag{3.26}$$

where

$$\begin{aligned} E^{(2)}_{\varepsilon,\tau,\delta,i,j}(u) = \biggl\{ & H_i^{(j)} \leq u - i\biggl(\frac{\overline{\mu}}{\pi_j} + 2\delta\biggr), \\ & H_i^{(j)} > \max\biggl(u\frac{\varepsilon+\delta}{r_j - 1 + \delta} + iD_j^+(\delta), (u+i)\tau\biggr), \\ & T_i^{(j)} \leq i\biggl(\frac{\overline{\mu}}{\pi_j} + \delta\biggr),\ \Delta(T_i^{(j)}) > i\biggl(\frac{\overline{\mu}}{\pi_j} - \delta\biggr), \\ & \Delta(u) - \Delta\bigl(T_i^{(j)} + H_i^{(j)}\bigr) > (1-\delta)\bigl(u - T_i^{(j)} - H_i^{(j)}\bigr)\biggr\}, \end{aligned}$$

where

$$D_j^+(\delta) = \delta\frac{2 - \delta - \overline{\mu}/\pi_j}{r_j - 1 + \delta}. \tag{3.27}$$

Let once again $\lambda > 0$ be a small positive number. Consider i in the range

$$\lambda u \leq i \leq \biggl(\frac{r_j - 1 - \varepsilon}{(r_j - 1 + \delta)(\overline{\mu}/\pi_j + 3\delta + D_j^+(\delta))}\biggr)u. \tag{3.28}$$

As before, for some $\tau_3(\varepsilon) > 0$, if $0 < \tau < \tau_3(\varepsilon)$, then, in our range of i,

$$u\frac{\varepsilon+\delta}{r_j - 1 + \delta} + iD_j^+(\delta) > (u+i)\tau$$

as long as $\delta > 0$ is small enough. Therefore,

$$
\begin{aligned}
&P\Big(E^{(2)}_{\varepsilon,\tau,\delta,i,j}(u) \cap B_\tau(u)\Big) \\
&\quad \ge P\bigg(u\frac{\varepsilon+\delta}{r_j-1+\delta} + iD_j^+(\delta) < H_i^{(j)} \le u - i\bigg(\frac{\overline{\mu}}{\pi_j}+2\delta\bigg)\bigg) \\
&\qquad - P\bigg(T_i^{(j)} > i\bigg(\frac{\overline{\mu}}{\pi_j}+\delta\bigg),\ H_i^{(j)} > (u+i)\tau\bigg) \\
&\qquad - P\bigg(\Delta\big(T_i^{(j)}\big) \le i\bigg(\frac{\overline{\mu}}{\pi_j}-\delta\bigg)\bigg) \qquad (3.29)\\
&\qquad - P\Big(\Delta(u) - \Delta\big(T_i^{(j)} + H_i^{(j)}\big) \le (1-\delta)\big(u - T_i^{(j)} - H_i^{(j)}\big), \\
&\qquad\qquad\qquad T_i^{(j)} + H_i^{(j)} \le u - i\delta,\ H_i^{(j)} > (u+i)\tau\Big) \\
&\quad := \overline{F}_j\bigg(u\frac{\varepsilon+\delta}{r_j-1+\delta} + iD_j^+(\delta)\bigg) - \overline{F}_j\bigg(u - i\bigg(\frac{\overline{\mu}}{\pi_j}+2\delta\bigg)\bigg) - \sum_{l=1}^{3} e_{l,i,j}(u).
\end{aligned}
$$

Notice that $e_{l,i,j}(u)$ with $l = 1, 2$ was handled in (3.16) and (3.17). Similarly, by the strong Markov property,

$$
e_{3,i,j}(u) \le P\big(\Delta(t) \le (1-\delta)t \text{ for some } t > i\delta\big) P\big(H_i^{(j)} > (u+i)\tau\big). \tag{3.30}
$$

By the ergodic theorem,

$$
\frac{\Delta(t)}{t} \to 1 \qquad \text{as } t \to \infty.
$$

Therefore, the first term on the right-hand side in (3.30) goes to 0 as $u \to \infty$ uniformly in i in the range (3.28). Using, once again, the regular variation of $\overline{F}$, we conclude that, for all u large enough,

$$
e_{l,i,j}(u) \le \delta u \overline{F}(u) \qquad \text{for } l = 1, 2, 3 \tag{3.31}
$$

for all i in the range (3.28) and $j \in J_+(\varepsilon)$. Letting

$$
d = \frac{r_j - 1 - \varepsilon}{(r_j - 1 + \delta)(\overline{\mu}/\pi_j + 3\delta + D_j^+(\delta))},
$$

we have, therefore, for all u large enough,

$$
\begin{aligned}
&P\big(E^{2,+}_{\varepsilon,\tau}(u)\big) \\
&\quad \ge \sum_{j\in J_+(\varepsilon)} \bigg((\theta_j - \delta) \sum_{\lambda u \le i \le du} \bigg(\overline{F}\bigg(u\frac{\varepsilon+\delta}{r_j-1+\delta} + iD_j^+(\delta)\bigg) \\
&\qquad\qquad - \overline{F}\bigg(u - i\bigg(\frac{\overline{\mu}}{\pi_j}+2\delta\bigg)\bigg)\bigg) - 3\delta u\overline{F}(u)\bigg) \qquad (3.32)
\end{aligned}
$$

$$
\begin{aligned}
&\geq \sum_{j\in J_+(\varepsilon)} (\theta_j-\delta)\Biggl(\sum_{\lambda u\leq i\leq du} \overline{F}\Bigl(u\frac{\varepsilon+\delta}{r_j-1+\delta}+iD_j^+(\delta)\Bigr) \\
&\qquad\qquad - \sum_{\lambda u\leq i\leq du} \overline{F}\Bigl(u-i\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)\Biggr) - 3\sum_{j\in J_+(\varepsilon)} \theta_j\delta u\overline{F}(u) \\
&:= \sum_{j\in J_+(\varepsilon)} (\theta_j-\delta)\bigl(S_{3,j}(u)-S_{4,j}(u)\bigr) - 3\sum_{j\in J_+(\varepsilon)} \theta_j\delta u\overline{F}(u).
\end{aligned}
$$

Clearly, if $\delta > 0$ is small enough, then

$$
S_{3,j}(u) \geq \bigl((d-\lambda)u-2\bigr)\overline{F}\Bigl(u\frac{\varepsilon+\delta^{1/2}}{r_j-1+\delta}\Bigr) \tag{3.33}
$$

for all u large enough. Observe, further, that as in (3.23) and (3.24), as $u\to\infty$,

$$
\begin{aligned}
S_{4,j}(u) &\leq \int_{\lambda u}^{du+1} \overline{F}\Bigl(u-x\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)\,dx \\
&\sim \Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)^{-1}\frac{1}{\gamma-1}\Bigl[u\Bigl(1-d\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)\overline{F}\Bigl(u\Bigl(1-d\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)\Bigr) \\
&\qquad - u\Bigl(1-\lambda\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)\overline{F}\Bigl(u\Bigl(1-\lambda\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)\Bigr)\Bigr].
\end{aligned} \tag{3.34}
$$

We conclude by (3.32)–(3.34) that, for all $\delta > 0$ and $\lambda > 0$ small enough,

$$
\begin{aligned}
&\liminf_{u\to\infty} \frac{P(E^{2,+}_{\varepsilon,\tau}(u))}{u\overline{F}(u)} \\
&\quad \geq \sum_{j\in J_+(\varepsilon)} (\theta_j-\delta)\Bigl[(d-\lambda)\Bigl(\frac{\varepsilon+\delta^{1/2}}{r_j-1+\delta}\Bigr)^{-\gamma} \\
&\qquad - \Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)^{-1}\frac{1}{\gamma-1}\Bigl(\Bigl(1-d\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)^{-(\gamma-1)} \\
&\qquad - \Bigl(1-\lambda\Bigl(\frac{\overline{\mu}}{\pi_j}+2\delta\Bigr)\Bigr)^{-(\gamma-1)}\Bigr)\Bigr] - 3\sum_{j\in J_+(\varepsilon)} \theta_j\delta.
\end{aligned}
$$

Letting $\delta\to 0$ and $\lambda\to 0$, we conclude that

$$
\begin{aligned}
\liminf_{u\to\infty} \frac{P(E^{2,+}_{\varepsilon,\tau}(u))}{u\overline{F}(u)} &\geq \frac{1}{\overline{\mu}}\sum_{j\in J_+(\varepsilon)} \theta_j\pi_j\Bigl[\frac{1}{1+s_j^+(\varepsilon)}\Bigl(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\Bigr)^{-\gamma} \\
&\qquad - \frac{1}{\gamma-1}\Bigl(\Bigl(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\Bigr)^{-(\gamma-1)}-1\Bigr)\Bigr].
\end{aligned} \tag{3.35}
$$

Now (3.6) follows from (3.10), (3.25) and (3.35). As mentioned above, this is enough to establish the asymptotic lower bound in our statement.

We will prove now the corresponding asymptotic upper bound. We will actually prove that

$$\limsup_{u\to\infty} \frac{P(E_\varepsilon^+(u))}{u\overline{F}(u)} \le \frac{1}{\varepsilon^\gamma \overline{\mu}} \sum_{j\in J_+(\varepsilon)} \theta_j \pi_j (r_j - 1 - \varepsilon)(r_j - 1)^{\gamma-1}. \tag{3.36}$$

Since the corresponding result for the event $E_\varepsilon^-(u)$ can be established in the same way, this will be enough to finish the proof of the theorem.

For $\tau > 0$ we define the event

$$A_\tau(u) = \Big\{ H_i^{(j)} \le \tau(u+i) \text{ for all pairs } (i,j) \in \{1,2,\dots\} \times \{1,\dots,K\} \Big\}. \tag{3.37}$$

Our first goal is to check that, for all τ small enough,

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^+ \cap A_\tau(u))}{u\overline{F}(u)} = 0. \tag{3.38}$$

Observe that

$$E_\varepsilon^+(u) = E_\varepsilon^{(1)}(u) \cup E_\varepsilon^{(2)}(u), \tag{3.39}$$

where

$$E_\varepsilon^{(1)}(u) = \Big\{ T_i^{(j)} + H_i^{(j)} \ge u,\ r_j H_i^{(j)} + \Delta(T_i^{(j)}) > (T_i^{(j)} + H_i^{(j)})(1+\varepsilon) \\ \text{for some } (i,j) \in \{1,2,\dots\} \times J_+(\varepsilon) \Big\}$$

and

$$E_\varepsilon^{(2)}(u) = \Big\{ T_i^{(j)} + H_i^{(j)} < u,\ r_j H_i^{(j)} + \Delta(T_i^{(j)}) + \Big(\Delta(u) - \Delta(T_i^{(j)} + H_i^{(j)})\Big) \\ > (1+\varepsilon)u \text{ for some } (i,j) \in \{1,2,\dots\} \times J_+(\varepsilon) \Big\}.$$

Therefore, (3.38) will follow once we show that

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^{(i)} \cap A_\tau(u))}{u\overline{F}(u)} = 0 \tag{3.40}$$

for $i = 1, 2$. Since the arguments for $i = 1$ and $i = 2$ are very similar, we only prove (3.40) for $i = 1$.

Let λ be a positive number satisfying

$$\Big(1 + \frac{1}{\lambda}\Big)\tau \le \frac{\varepsilon\overline{\mu}}{4} \min_{j=1,\dots,K} \Big(\frac{1}{\pi_j r_j}\Big), \tag{3.41}$$

and write

$$E_\varepsilon^{(1)} = E_\varepsilon^{(11)} \cup E_\varepsilon^{(12)}, \tag{3.42}$$

where

$$E_\varepsilon^{(11)}(u) = \Big\{T_i^{(j)} + H_i^{(j)} \ge u,\ r_j H_i^{(j)} + \Delta\big(T_i^{(j)}\big) > \big(T_i^{(j)} + H_i^{(j)}\big)(1+\varepsilon)$$
$$\text{for some } i \le \lambda u \text{ and } j \in J_+(\varepsilon)\Big\}$$

and

$$E_\varepsilon^{(12)}(u) = \Big\{T_i^{(j)} + H_i^{(j)} \ge u,\ r_j H_i^{(j)} + \Delta\big(T_i^{(j)}\big) > \big(T_i^{(j)} + H_i^{(j)}\big)(1+\varepsilon)$$
$$\text{for some } i > \lambda u \text{ and } j \in J_+(\varepsilon)\Big\}.$$

Let us introduce some notation. For a state j let $I_1^{(j)}, I_1^{(j)}, \dots$ be the times between subsequent returns of the underlying Markov chain to state j. Similarly, let $I_1^{(j_0,j)}$ be the time of the first visit to state j starting at the initial state j_0. Note that

$$P\Big(E_\varepsilon^{(11)} \cap A_\tau(u)\Big)$$
$$\le \sum_{j \in J_+(\varepsilon)} \Bigg[P\Big(\Big\{I_1^{(j_0,j)} \ge \tfrac{1}{2}u\big(1-\tau(1+\lambda)\big)\Big\} \cap A_\tau(u)\Big) \tag{3.43}$$
$$+ \sum_{i \le \lambda u} P\Big(\Big\{I_1^{(j)} + \cdots + I_{i-1}^{(j)} \ge \tfrac{1}{2}u\big(1-\tau(1+\lambda)\big)\Big\} \cap A_\tau(u)\Big)\Bigg].$$

Let $N_1^{(j)}, N_1^{(j)}, \dots$ be the numbers of steps it takes the underlying Markov chain to return to state j and let $N_1^{(j_0,j)}$ be the number of steps it takes the underlying Markov chain to visit the state j starting at the initial state j_0. Since the Markov chain is finite and irreducible, the random variables $N_1^{(j)}$ and $N_1^{(j_0,j)}$ have exponentially decaying tails. Note, further, that, by Lemma 3.6, for every $i \le \lambda u$,

$$P\Big(\Big\{I_1^{(j)} + \cdots + I_{i-1}^{(j)} \ge \tfrac{1}{2}u\big(1-\tau(1+\lambda)\big)\Big\} \cap A_\tau(u)\Big)$$
$$\le P\left(\sum_{k=1}^{N_1^{(j)}+\cdots+N_{i-1}^{(j)}} \big(H_k^* \wedge \tau(1+\lambda)u\big) \ge \tfrac{1}{2}u\big(1-\tau(1+\lambda)\big)\right), \tag{3.44}$$

where $H_1^*, H_2^*, \dots$ are iid random variables independent of $N_k^{(j)}$, $k \ge 1$, with distribution F^* (described in that lemma). An exponential Markov inequality immediately tells us that there is a $\theta_1 > 0$ such that, for all $i \le \lambda u$,

$$P\Big(N_1^{(j)} + \cdots + N_{i-1}^{(j)} > 2\lambda u E N_1^{(j)}\Big) \le e^{-\theta_1 \lambda u}. \tag{3.45}$$

Furthermore,

$$P\left(\sum_{k\le 2\lambda u EN_1^{(j)}} \big(H_k^* \wedge \tau(1+\lambda)u\big) \ge \frac{1}{2}u\big(1-\tau(1+\lambda)\big)\right) \tag{3.46}$$
$$\le P\left(\sum_{k\le 2\lambda u EN_1^{(j)}} \Big(\big(H_k^* \wedge \tau(1+\lambda)u\big) - E\big(H_k^* \wedge \tau(1+\lambda)u\big)\Big) > \frac{u}{8}\right)$$

as long as

$$\lambda \le \frac{1}{2EN_1^{(j)}} \wedge 1 \quad \text{and} \quad \tau \le \frac{1}{8}.$$

Applying Lemma 3.7 with $c = \tau(1+\lambda)u$, we immediately conclude that, for some $\theta_2 > 0$,

$$P\left(\sum_{k\le 2\lambda u EN_1^{(j)}} \Big(\big(H_k^* \wedge \tau(1+\lambda)u\big) - E\big(H_k^* \wedge \tau(1+\lambda)u\big)\Big) > \frac{u}{8}\right) \tag{3.47}$$
$$\le \theta_2 u^{-1/8\tau}.$$

Bounding in a similar way the first term under the sum on the right-hand side of (3.43), we immediately conclude from the above that, for all $\lambda > 0$ and $\tau > 0$ small enough and such that (3.41) holds,

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^{(11)} \cap A_\tau(u))}{u\overline{F}(u)} = 0. \tag{3.48}$$

Before proceeding to treat $P(E_\varepsilon^{(12)} \cap A_\tau(u))$, we note, for future use, that the same argument as the one used above to establish (3.48) also shows the following. For a fixed $\tau > 0$, let

$$T_* = \inf\Big\{T_i^{(j)} : j = 1, \dots, K,\ i = 1, 2, \dots \text{ and } H_i^{(j)} > \tau(u+i)\Big\}. \tag{3.49}$$

Then, for all $\tau > 0$ small enough,

$$\lim_{u\to\infty} \frac{P(\sup_{0\le t\le T_*}(\Delta(t) - (1+\varepsilon)t) > u)}{u\overline{F}(u)} = 0. \tag{3.50}$$

We now switch to estimating $P(E_\varepsilon^{(12)} \cap A_\tau(u))$. Note that, for any $\lambda > 0$,

$$
\begin{aligned}
&P\Big(E_\varepsilon^{(12)} \cap A_\tau(u)\Big)\\
&\quad \le \sum_{j\in J_+(\varepsilon)} \sum_{i\ge\lambda u} P\Big(\Big\{T_i^{(j)} + H_i^{(j)} \ge u,\ r_j H_i^{(j)} + \Delta\big(T_i^{(j)}\big)\\
&\qquad\qquad > \big(T_i^{(j)} + H_i^{(j)}\big)(1+\varepsilon)\Big\} \cap A_\tau(u)\Big)\\
&\quad := \sum_{j\in J_+(\varepsilon)} \sum_{i\ge\lambda u} P\Big(E_{\varepsilon,i,j}^{(3)}(u) \cap A_\tau(u)\Big).
\end{aligned}
\tag{3.51}
$$

Let $\mathcal{G}$ be the σ-field generated by the process $(Z_n,\ n \ge 0)$, that is, by the sequence of states the system goes through. Note that, for every $j \in J_+(\varepsilon)$ and $i \ge \lambda u$,

$$
\begin{aligned}
&P\Big(E_{\varepsilon,i,j}^{(3)}(u) \cap A_\tau(u)\Big)\\
&\quad \le P\Big(\Big\{\tau(r_j - 1 - \varepsilon)(u+i) + \Delta\big(T_i^{(j)}\big) - (1+\varepsilon)T_i^{(j)} > 0\Big\} \cap A_\tau(u)\Big)\\
&\quad := P\Big(E_{\varepsilon,i,j,\tau}^{(3)}(u)\Big) = E\Big(P\Big(E_{\varepsilon,i,j,\tau}^{(3)}(u)|\mathcal{G}\Big)\Big).
\end{aligned}
\tag{3.52}
$$

For $l, j = 1, \dots, K$ and $i \ge 1$, let $J_{l,j}(i)$ be the number of visits to state l until the ith visit to state j. Obviously, $J_{l,j}(i)$ is measurable with respect to $\mathcal{G}$ for all l, j, i. Notice that

$$
E\big(J_{l,j}(2) - J_{l,j}(1)\big) = \frac{\pi_l}{\pi_j},
$$

and so, since the Markov chain is finite and irreducible, for any given $\rho > 0$ there is $\theta_3 > 0$ such that

$$
P\left(\left|\frac{J_{l,j}(i)}{i} - \frac{\pi_l}{\pi_j}\right| > \rho\right) \le \theta_3 e^{-i/\theta_3} \tag{3.53}
$$

for all $l, j = 1, \dots, K$ and $i \ge 1$.

For $\rho > 0$ so small that

$$
\sum_{l=1}^K \mu_l\left(\frac{\pi_l}{\pi_j} - \rho\right)(1 + \varepsilon - r_l) \ge \frac{\varepsilon\overline{\mu}}{2\pi_j} \tag{3.54}
$$

for all $j \in J_+(\varepsilon)$, we let

$$
A_{i,j}(\rho) = \left\{w : \left|\frac{J_{l,j}(i)}{i} - \frac{\pi_l}{\pi_j}\right| \le \rho \text{ for all } l = 1, \dots, K\right\} \in \mathcal{G}.
$$

It follows from (3.53) that, for some $\theta_3 = \theta_3(\rho) > 0$,

$$
P\Big(E_{\varepsilon,i,j,\tau}^{(3)}(u)\Big) \le \theta_3 e^{-i/\theta_3} + E\Big(\mathbf{1}\big(A_{i,j}(\rho)\big)P\big(E_{\varepsilon,i,j,\tau}^{(3)}(u)|\mathcal{G}\big)\Big). \tag{3.55}
$$

Denote

$$W_i^{(j)} = \left(\Delta\left(T_i^{(j)}\right) - (1+\varepsilon)T_i^{(j)}\right) - E\left(\Delta\left(T_i^{(j)}\right) - (1+\varepsilon)T_i^{(j)}\right)$$
$$= \left(\Delta\left(T_i^{(j)}\right) - (1+\varepsilon)T_i^{(j)}\right) - \sum_{l=1}^{K} \mu_l J_{l,j}(i)(r_l - 1 - \varepsilon).$$

We then have

$$P\left(E^{(3)}_{\varepsilon,i,j,\tau}(u)|\mathcal{G}\right) = P\left(\left\{\tau(r_j - 1 - \varepsilon)(u+i) + W_i^{(j)} > \sum_{l=1}^{K} \mu_l J_{l,j}(i)(1+\varepsilon - r_l)\right\} \cap A_\tau(u) \Big| \mathcal{G}\right) \tag{3.56}$$
$$\le P\left(\left\{W_i^{(j)} > i\frac{\varepsilon\overline{\mu}}{4\pi_j}\right\} \cap A_\tau(u)\Big|\mathcal{G}\right)$$

by (3.54) and (3.41).

Let S_n be the sojourn time the system spends in the nth state it visits (i.e., in state Z_n); then the total increase of the time Δ during that sojourn is $r_{Z_n} S_n$. Then

$$W_i^{(j)} = \sum_{n=0}^{\sum_{l=1}^{K} J_{l,j}(i)-1} U_n,$$

where

$$U_n = \left((r_{Z_n} - 1 - \varepsilon)S_n\right) - E\left((r_{Z_n} - 1 - \varepsilon)S_n\right), \qquad n \ge 1.$$

We conclude that

$$P\left(\left\{W_i^{(j)} > i\frac{\varepsilon\overline{\mu}}{4\pi_j}\right\} \cap A_\tau(u)\Big|\mathcal{G}\right) \le P\left(\sum_{n=0}^{\sum_{l=1}^{K} J_{l,j}(i)-1} U_n \mathbf{1}\left(S_n \le \tau(u+n)\right) > i\frac{\varepsilon\overline{\mu}}{4\pi_j}\Big|\mathcal{G}\right).$$

Denote

$$U_n^* = U_n \mathbf{1}\left(S_n \le \tau(u+n)\right) - E\left(U_n \mathbf{1}\left(S_n \le \tau(u+n)\right)|\mathcal{G}\right), \qquad n \ge 0,$$

and observe that

$$E\left(U_n \mathbf{1}\left(S_n \le \tau(u+n)\right)|\mathcal{G}\right) \to 0$$

as $u \to 0$ uniformly over $\omega \in A_{i,j}(\rho)$ and $n \ge 0$. Furthermore, $U_0^*, U_1^*, \dots$ are, conditionally on $\mathcal{G}$, independent zero mean random variables, and for some absolute constant $C > 0$, for all $\omega \in A_{i,j}(\rho)$ and $n \ge 0$, $|U_n^*| \le C\tau(u+n)$. For all u large enough, all $\omega \in A_{i,j}(\rho)$ and all $i \ge \lambda u$, we then have

$$P\left(\left\{W_i^{(j)} > i\frac{\varepsilon\overline{\mu}}{4\pi_j}\right\} \cap A_\tau(u)\Big|\mathcal{G}\right) \le P\left(\sum_{n=0}^{\sum_{l=1}^{K} J_{l,j}(i)-1} U_n^* > i\frac{\varepsilon\overline{\mu}}{8\pi_j}\Big|\mathcal{G}\right). \tag{3.57}$$

We are now in a position to apply Lemma 3.7 with $c = C\tau(u + \sum_{l=1}^{K} J_{l,j}(i))$ to conclude that there is a $\theta_3 > 0$ and $\tau_0 > 0$ such that, for all $0 < \tau \le \tau_0$, all $\omega \in A_{i,j}(\rho)$, all $i \ge \lambda u$ and all $u > 0$ large enough,

$$P\left(\left\{W_i^{(j)} > i\frac{\varepsilon\overline{\mu}}{4\pi_j}\right\} \cap A_\tau(u)\middle|\mathcal{G}\right) \le \theta_3 i^{-\tau/\theta_3}. \tag{3.58}$$

Therefore, we conclude by (3.51), (3.52), (3.55), (3.56) and (3.58) that, for all $\lambda > 0$ and $\tau > 0$ small enough,

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^{(12)} \cap A_\tau(u))}{u\overline{F}(u)} = 0. \tag{3.59}$$

Now the statement (3.40) with $i = 1$ follows from (3.48) and (3.59).

Next for $\tau > 0$ define the event

$$\begin{aligned} B_\tau(u) = \Big\{&H_i^{(j)} > \tau(u+i) \\ &\text{for at least two different pairs } (i, j) \in \{1, 2, \ldots\} \times \{1, \ldots, K\}\Big\}. \end{aligned} \tag{3.60}$$

It is an immediate consequence of Lemma 3.8 that, for any $\tau > 0$,

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^{+}(u) \cap B_\tau(u))}{u\overline{F}(u)} = 0. \tag{3.61}$$

Therefore, it follows from (3.38) and (3.61) that to establish (3.36) it is enough to prove that

$$\begin{aligned} &\lim_{\tau\downarrow 0}\limsup_{u\to\infty} \frac{P(E_\varepsilon^{+}(u) \cap C_\tau(u))}{u\overline{F}(u)} \\ &\quad \le \frac{1}{\varepsilon^\gamma\overline{\mu}} \sum_{j\in J_+(\varepsilon)} \theta_j\pi_j(r_j - 1 - \varepsilon)(r_j - 1)^{\gamma-1}, \end{aligned} \tag{3.62}$$

where

$$\begin{aligned} C_\tau(u) = \Big\{&H_i^{(j)} > \tau(u+i) \text{ for exactly one pair} \\ &(i, j) \in \{1, 2, \ldots\} \times \{1, \ldots, K\}\Big\}. \end{aligned} \tag{3.63}$$

To this end, let

$$\begin{aligned} C_\tau^{+}(u) = \Big\{&H_i^{(j)} > \tau(u+i) \text{ for exactly one pair} \\ &(i, j) \in \{1, 2, \ldots\} \times \{1, \ldots, K\}, \text{ and the corresponding } j \in J_+(\varepsilon)\Big\} \\ \subset{}& C_\tau(u). \end{aligned} \tag{3.64}$$

We will first check that, for all $\tau > 0$ small enough,

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^+(u) \cap (C_\tau(u) \setminus C_\tau^+(u)))}{u\overline{F}(u)} = 0. \tag{3.65}$$

Let T_*, H_* and r_* be, correspondingly, the starting time, the length of the holding time satisfying $H_i^{(j)} > \tau(u+i)$ and the corresponding rate r_j; see also (3.49). Note that on the event $C_\tau(u)$ these are well-defined random variables. Moreover, for some $0 < \varepsilon' < \varepsilon$, $r_* < 1 + \varepsilon'$ on the event $C_\tau(u) \setminus C_\tau^+(u)$. Consider the events

$$S_1(u) = \{T_* > u\}, \qquad S_2(u) = \left\{T_* + H_* \le \frac{u}{2}\right\},$$
$$S_3(u) = \left\{T_* \le u,\ T_* + H_* > \frac{u}{2}\right\}.$$

Notice that replacing any holding time $H_i^{(j)}$ with $j \notin J_+(\varepsilon)$ and such that $T_i^{(j)} > u$ by $\min(H_i^{(j)}, \tau(u+i))$ cannot take a realization in $E_\varepsilon^+(u)$ to the complement of this event. Therefore, replacing H_* with $\min(H_*, \tau(u+i))$, we can use the same argument as that used in the proof of (3.48) to see that, for any $\tau > 0$, $\delta_1 > 0$ and $\delta_2 > 0$,

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^+(u) \cap (C_\tau(u) \setminus C_\tau^+(u)) \cap S_1(u))}{u\overline{F}(u)} = 0, \tag{3.66}$$

and that same argument we used to prove (3.48) also gives us that

$$\lim_{u\to\infty} \frac{P(E_\varepsilon^+(u) \cap (C_\tau(u) \setminus C_\tau^+(u)) \cap S_3(u))}{u\overline{F}(u)} = 0. \tag{3.67}$$

Now, with $0 < \varepsilon' < \varepsilon$ as above, write

$$\begin{aligned}
&P\big(E_\varepsilon^+(u) \cap \big(C_\tau(u) \setminus C_\tau^+(u)\big) \cap S_2(u)\big) \\
&\quad\le P\Big(E_\varepsilon^+(u) \cap \big(C_\tau(u) \setminus C_\tau^+(u)\big) \cap S_2(u) \\
&\qquad\qquad \cap \big\{\Delta(T_* + H_*) \le (1+\varepsilon)(T_* + H_*) + (\varepsilon - \varepsilon')u\big\}\Big) \\
&\qquad + P\Big(E_\varepsilon^+(u) \cap \big(C_\tau(u) \setminus C_\tau^+(u)\big) \cap S_2(u) \\
&\qquad\qquad \cap \big\{\Delta(T_* + H_*) > (1+\varepsilon)(T_* + H_*) + (\varepsilon - \varepsilon')u\big\}\Big) \\
&\quad:= P\big(D_1(u)\big) + P\big(D_2(u)\big).
\end{aligned}$$

Observe that, on the event $D_2(u)$, $\Delta(T_*) - (1+\varepsilon)T_* > (\varepsilon - \varepsilon')u$, and so it follows from (3.50) that, for all $\tau > 0$ small enough,

$$\lim_{u\to\infty} \frac{P(D_2(u))}{u\overline{F}(u)} = 0.$$

On the other hand, on the event $D_1(u)$, we have

$$\sup_{t\geq u/2}\big(\Delta(t+T_*+H_*)-\Delta(T_*+H_*)-(1+\varepsilon')t\big)>0,$$

and then the strong Markov property and the argument leading to (3.67) give us

$$\lim_{u\to\infty}\frac{P(D_1(u))}{u\overline{F}(u)}=0$$

for all $\tau>0$ small enough. In conclusion, for all $\tau>0$ small enough,

$$\lim_{u\to\infty}\frac{P(E_\varepsilon^+(u)\cap(C_\tau(u)\setminus C_\tau^+(u))\cap S_2(u))}{u\overline{F}(u)}=0, \tag{3.68}$$

and now (3.65) follows from (3.66)–(3.68).

Summing up, to finish the proof of the theorem, we need to show that

$$\begin{aligned}&\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(E_\varepsilon^+(u)\cap C_\tau^+(u))}{u\overline{F}(u)}\\ &\qquad\leq\frac{1}{\varepsilon^\gamma\overline{\mu}}\sum_{j\in J_+(\varepsilon)}\theta_j\pi_j(r_j-1-\varepsilon)(r_j-1)^{\gamma-1}.\end{aligned} \tag{3.69}$$

Let T_*, H_* and r_* be, as above, the starting time, the length of the holding time satisfying $H_i^{(j)}>\tau(u+i)$ and the corresponding rate r_j, and let I_* be the corresponding i $(=1,2,\ldots)$. For $\delta>0$ write

$$\begin{aligned}P\big(E_\varepsilon^+(u)\cap C_\tau^+(u)\big)=\ &P\big(E_\varepsilon^+(u)\cap C_\tau^+(u)\cap\{T_*+H_*\geq u(1-\delta)\}\big)\\ &+P\big(E_\varepsilon^+(u)\cap C_\tau^+(u)\cap\{T_*+H_*<u(1-\delta)\}\big)\\ :=\ &P\big(D_3(u)\big)+P\big(D_4(u)\big).\end{aligned} \tag{3.70}$$

We have

$$\begin{aligned}&P\big(D_3(u)\big)\\ &\quad\leq\sum_{j\in J_+(\varepsilon)}\sum_{i=1}^\infty P\left(H_i^{(j)}>\max\left(u(1-\delta)-i\left(\frac{\overline{\mu}}{\pi_j}+\delta\right),i\,\frac{\varepsilon(\overline{\mu}/\pi_j)-\varepsilon\delta-2\delta}{r_j-1-\varepsilon}\right)\right)\\ &\qquad+P\left(D_3(u)\cap\left\{H_i^{(j)}\leq u(1-\delta)-i\left(\frac{\overline{\mu}}{\pi_j}+\delta\right)\right.\right.\\ &\qquad\qquad\left.\left.\text{for all } j\in J_+(\varepsilon)\text{ and } i=1,2,\ldots\right\}\right)\end{aligned} \tag{3.71}$$

$$+ P\Big(D_3(u) \cap \Big\{H_i^{(j)} \le i\frac{\varepsilon(\overline{\mu}/\pi_j) - \varepsilon\delta - 2\delta}{r_j - 1 - \varepsilon} \text{ for all } j \in J_+(\varepsilon) \text{ and } i = 1, 2, \dots\Big\}\Big).$$

Now, a computation entirely analogous to the one in (3.22)–(3.24) gives us

$$\lim_{\delta\downarrow 0}\limsup_{u\to\infty}\Big(\sum_{j\in J_+(\varepsilon)}\sum_{i=1}^{\infty} P\Big(H_i^{(j)} > \max\Big(u(1-\delta) - i(\overline{\mu}/\pi_j + \delta), i\big(\varepsilon(\overline{\mu}/\pi_j) - \varepsilon\delta - 2\delta\big)/(r_j - 1 - \varepsilon)\Big)\Big)\Big)\frac{1}{u\overline{F}(u)} \tag{3.72}$$
$$\le \frac{1}{\overline{\mu}(\gamma-1)}\sum_{j\in J_+(\varepsilon)}\theta_j\pi_j\Big(\Big(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\Big)^{-\gamma} - 1\Big).$$

We now check that the last two probabilities on the right-hand side of (3.71) are, asymptotically, small. Denote the events measured by these two probabilities by $D_{31}(u)$ and $D_{32}(u)$, correspondingly. Observe that, for a $\lambda > 0$,

$$P\big(D_{31}(u)\big) \le P\Big(\bigcup_{j\in J_+(\varepsilon)}\bigcup_{i>\lambda u}\Big\{\Delta\big(T_i^{(j)}\big) > i\Big(\frac{\overline{\mu}}{\pi_j} + \delta\Big)\Big\} \cap A_{i,j;\tau(u)}\Big) + P\Big(\bigcup_{j\in J_+(\varepsilon)}\bigcup_{i\le\lambda u}\{H_i^{(j)} + T_i^{(j)} > u(1-\delta)\}\Big) \tag{3.73}$$
$$:= P\big(D_{311}(u)\big) + P\big(D_{312}(u)\big),$$

where, for $j = 1, \dots, K$ and $i = 1, 2, \dots,$

$$A_{i,j;\tau(u)} = \Big\{H_l^{(k)} \le \tau(u + l) \text{ for all pairs } (l, k) \in \{1, 2, \dots\} \times \{1, \dots, K\} \text{ such that } T_l^{(k)} < T_i^{(j)}\Big\}.$$

Now the same application of Lemma 3.7 as we used to prove (3.40) shows that

$$\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(D_{311}(u))}{u\overline{F}(u)} = 0 \tag{3.74}$$

for all $\delta > 0$ and $\lambda > 0$. On the other hand, it is clear that

$$\lim_{\lambda\downarrow 0}\limsup_{u\to\infty}\frac{P(D_{312}(u))}{u\overline{F}(u)} = 0 \tag{3.75}$$

for all $\delta > 0$. Therefore, we conclude that

$$\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(D_{31}(u))}{u\overline{F}(u)} = 0 \tag{3.76}$$

for all $\delta > 0$. Furthermore,

$$\begin{aligned} P\big(D_{32}(u)\big) = {} & P\Big(D_{32}(u) \cap \big\{r_* + \Delta(T_*) > (1+\varepsilon')(H_* + T_*)\big\}\Big) \\ & + P\Big(D_{32}(u) \cap \big\{r_* + \Delta(T_*) \le (1+\varepsilon')(H_* + T_*)\big\}\Big) \\ := {} & P\big(D_{321}(u)\big) + P\big(D_{322}(u)\big), \end{aligned} \tag{3.77}$$

where we recall that $0 < \varepsilon' < \varepsilon$ is such that $r_j < 1 + \varepsilon'$ for all $j \notin J_+(\varepsilon)$. Now, for a $\lambda > 0$,

$$P\big(D_{321}(u)\big) \le P\big(D_{311}(u)\big) + P\big(D_{312}(u)\big) + P\Bigg(\bigcup_{j\in J_+(\varepsilon)}\bigcup_{i>\lambda u}\bigg\{T_i^{(j)} \le i\bigg(\frac{\overline{\mu}}{\pi_j} - \delta\bigg)\bigg\}\Bigg),$$

and so it follows from (3.74), (3.75) and Lemma 3.5 that

$$\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(D_{321}(u))}{u\overline{F}(u)} = 0 \tag{3.78}$$

for all $\delta > 0$. Furthermore,

$$\begin{aligned} P\big(D_{322}(u)\big) = {} & P\Bigg(D_{322}(u) \cap \bigg\{\frac{\Delta(t)}{t} > 1 + \varepsilon \text{ for some } u < t < T_*\bigg\}\Bigg) \\ & + P\Bigg(D_{322}(u) \cap \bigg\{\frac{\Delta(t)}{t}\ 1 + \varepsilon \text{ for some } t > H_* + T_*\bigg\}\Bigg). \end{aligned}$$

Since time T_* is the beginning of the only holding time $H_i^{(j)} > \tau(u+i)$, the first probability on the right-hand side above describes a situation of the type (3.38), and so the same argument gives us

$$\lim_{u\to\infty}\frac{P(D_{322}(u) \cap \{\Delta(t)/t > 1 + \varepsilon' \text{ for some } u < t < T_*\})}{u\overline{F}(u)} = 0 \tag{3.79}$$

for all $\delta > 0$ and $\tau > 0$ small enough. Finally, by the strong Markov property,

$$\begin{aligned} & P\Bigg(D_{322}(u) \cap \bigg\{\frac{\Delta(t)}{t}\ 1 + \varepsilon \text{ for some } t > H_* + T_*\bigg\}\Bigg) \\ & \quad \le P\Bigg(\sup_{0\le t\le T_*}\big(\Delta(t) - (1+\varepsilon)t\big) > u(1-\delta)(\varepsilon - \varepsilon')\Bigg), \end{aligned}$$

and so it follows from (3.79) and (3.50) that

$$\lim_{u\to\infty}\frac{P(D_{322}(u))}{u\overline{F}(u)} = 0 \tag{3.80}$$

for all $\delta > 0$ and $\tau > 0$ small enough and, hence, by (3.76) and (3.80) we get

$$\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(D_{32}(u))}{u\overline{F}(u)} = 0 \tag{3.81}$$

for all $\delta > 0$. Now we conclude by (3.72), (3.76) and (3.81) that

$$\begin{aligned}&\lim_{\delta\downarrow 0}\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(D_3(u))}{u\overline{F}(u)}\\ &\qquad\le \frac{1}{\overline{\mu}(\gamma-1)}\sum_{j\in J_+(\varepsilon)}\theta_j\pi_j\left(\left(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\right)^{-\gamma}-1\right).\end{aligned} \tag{3.82}$$

Now, a similar decomposition of the event $D_4(u)$ gives us a corresponding upper bound

$$\begin{aligned}&\lim_{\delta\downarrow 0}\lim_{\tau\downarrow 0}\limsup_{u\to\infty}\frac{P(D_4(u))}{u\overline{F}(u)}\\ &\qquad\le \frac{1}{\overline{\mu}}\sum_{j\in J_+(\varepsilon)}\theta_j\pi_j\left[\frac{1}{1+s_j^+(\varepsilon)}\left(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\right)^{-\gamma}\right.\\ &\qquad\qquad\left.-\frac{1}{\gamma-1}\left(\left(\frac{s_j^+(\varepsilon)}{1+s_j^+(\varepsilon)}\right)^{-(\gamma-1)}-1\right)\right],\end{aligned} \tag{3.83}$$

and so the statement (3.69) follows from (3.82) and (3.83). This establishes the asymptotic upper bound and hence completes the proof of the theorem. □

We conclude this section with several statements required in the proof of Theorem 3.1. Some of these statements are well known, and we present them here for completeness. We use the notation introduced earlier in the section. The first lemma shows that the starting times of the sojourns of the underlying process in different states are very unlikely to be much smaller than their means.

LEMMA 3.5. *For every* $j = 1, \dots, K$ *and* $\delta > 0$, *there are positive numbers* $C_1^{(j)}$ *and* $C_2^{(j)}$ *such that, for all* $i \ge 1$,

$$P\left(T_i^{(j)} \le i\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)\right) \le C_1^{(j)}e^{-C_2^{(j)}i} \tag{3.84}$$

and

$$P\left(\Delta(T_i^{(j)}) \le i\left(\frac{\overline{\mu}}{\pi_j}-\delta\right)\right) \le C_1^{(j)}e^{-C_2^{(j)}i}. \tag{3.85}$$

PROOF. Since the proofs of both statements are very similar, we only prove (3.85). It is, of course, enough to consider $i \ge 2$. With the usual notation P_j

and E_j meaning that $j = 1, \dots, K$ is the initial state of the Markov chain, we have, for any $\theta > 0$,

$$\begin{aligned} P&\left(\Delta(T_i^{(j)}) \le i\left(\frac{\overline{\mu}}{\pi_j} - \delta\right)\right) \\ &\le P_j\left(\Delta(T_{i-1}^{(j)}) \le i\left(\frac{\overline{\mu}}{\pi_j} - \delta\right)\right) \\ &\le \exp\left\{\theta i\left(\frac{\overline{\mu}}{\pi_j} - \delta\right)\right\} E_j \exp\{-\theta\Delta(T_{i-1}^{(j)})\} \\ &= \left(\exp\left\{\theta\left(\frac{\overline{\mu}}{\pi_j} - \delta\right)\right\} E_j \exp\{-\theta\Delta(T_1^{(j)})\}\right)^{i-1} \exp\left\{\theta\left(\frac{\overline{\mu}}{\pi_j} - \delta\right)\right\}. \end{aligned}$$

However,

$$E_j\Delta(T_1^{(j)}) = \sum_{k=1}^{K} \frac{\pi_k}{\pi_j}\mu_k r_k = \frac{\overline{\mu}}{\pi_j}$$

by (3.1). Therefore, there is a $\theta > 0$ such that

$$E_j \exp\{-\theta\Delta(T_1^{(j)})\} \le \exp\left\{-\theta\left(\frac{\overline{\mu}}{\pi_j} - \frac{\delta}{2}\right)\right\}$$

and our statement follows with

$$C_1^{(j)} = \exp\left\{\theta\left(\frac{\overline{\mu}}{\pi_j} - \frac{\delta}{2}\right)\right\}, \qquad C_2^{(j)} = \frac{\theta\delta}{2}. \qquad \square$$

The next lemma puts a common stochastic bound on the sojourn random variables.

LEMMA 3.6. *Under the assumption* (3.3) *there is a nonnegative random variable* H^* *with a distribution* F^* *such that* $H^* \overset{\text{st}}{\ge} H_1^{(j)}$ *for every* $j = 1, \dots, K$ *and* $\lim_{x\to\infty} \overline{F}^*(x)/\overline{F}(x) = \theta^*$ *for some* $\theta^* \in (0, \infty)$.

PROOF. Let H be a random variable with distribution F. It follows from (3.3) that for every $j = 1, \dots, K$ there is a $b_j \ge 0$ such that $(\theta_j + 1)H + b_j \overset{\text{st}}{\ge} H_1^{(j)}$. Now set $H^* = (\max(\theta_1, \dots, \theta_K) + 1)H + \max(b_1, \dots, b_K)$. $\square$

The following inequality for sums of independent uniformly bounded zero mean random variables is very useful.

LEMMA 3.7. *Let $Y_1, \dots, Y_k$ be independent zero mean random variables such that, for some $c > 0$, $|Y_n| \le c$ a.s. for $n = 1, \dots, k$. Then, for every $u > 0$,*

$$P\left(\sum_{n=1}^{k} Y_k > u\right) \le \exp\left\{-\frac{u}{2c} \operatorname{arcsinh}\left(cu \Big/ 2 \operatorname{Var}\left(\sum_{n=1}^{k} Y_k\right)\right)\right\}.$$

PROOF. See Prokhorov (1959); see also Petrov (1995), 2.6.1 on page 77 or Lemma A.2 in Mikosch and Samorodnitsky (2000b). □

The next lemma shows that it is very unlikely that two different holding times are both sufficiently long to matter as far as the rate of mixing is concerned.

LEMMA 3.8. *For any $\tau > 0$ the event $B_\tau(u)$ in (3.60) satisfies*

$$\lim_{u \to \infty} \frac{P(B_\tau(u))}{u\overline{F}(u)} = 0.$$

PROOF. This statement is an immediate consequence of Lemma 2.7 in Mikosch and Samorodnitsky (2000b). □

4. Conclusion. Beyond the modeling of insolvency related to non-life insurance, more recently, ruin estimation for general claim processes has become important as a potential methodological tool in the analysis of financial data. As a consequence, more refined insurance-type risk processes are called for. These models will, in particular, have to cater to stochastic intensities driven by exogenous economic factors and, at the same time, allow for heavy-tailed claim amounts. Our results show how classical ruin estimation results are "robust" with respect to stochastic changes away from a constant intensity model. Through various examples, it is shown which intensity models allow for such robustness. Especially the model treated in Section 3 on Markov chain switching models was motivated by practical considerations from finance where underlying market variables may switch randomly between states indicating various levels of economic activity. Our contribution will hopefully add some further understanding to the general class of models that allow for a ruin theoretic analysis.

Though our results were presented with eventual insurance and finance applications in mind, it should be obvious that the same results apply to numerous other areas of applied research, such as teletraffic and Internet data, dam theory and storage models in operations research.

REFERENCES

ADLER, R. (1990). *An Introduction to Continuity Extrema and Related Topics for General Gaussian Processes.* IMS, Hayward, CA.

ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.

ASMUSSEN, S., SCHMIDLI, H. and SCHMIDT, V. (1999). Tail probabilities for non-standard risk and queueing processes with subexponential jumps. *Adv. in Appl. Probab.* **31** 422–447.

BERAN, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall, New York.

BRAVERMAN, M., MIKOSCH, T. and SAMORODNITSKY, G. (2002). The tail behaviour of subadditive functionals acting on Lévy processes. *Ann. Appl. Probab.* **12** 69–100.

BRIYS, E. and DE VARENNE, F. (2001). *Insurance, from Underwriting to Derivatives*. Wiley, New York.

ÇINLAR, E. (1975). *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs, NJ.

CROUHY, M., MARK, R. and GALAI, D. (2000). *Risk Management*. McGraw-Hill, New York.

DOHERTY, N. (2000). *Integrated Risk Management*. McGraw-Hill, New York.

EMBRECHTS, P., GOLDIE, C. and VERAVERBEKE, N. (1979). Subexponentiality and infinite divisibility. *Z. Wahrsch. Verw. Gebiete* **49** 335–347.

EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin.

KAUFMANN, R., GADMER, A. and KLETT, R. (2001). Introduction to dynamic financial analysis. *ASTIN Bulletin* **31** 213–249.

MEDOVA, E. (2000). Measuring risk by extreme values. *Risk* S20–S26.

MIKOSCH, T. and SAMORODNITSKY, G. (2000a). Ruin probability with claims modeled by a stationary ergodic stable process. *Ann. Probab.* **28** 1814–1851.

MIKOSCH, T. and SAMORODNITSKY, G. (2000b). The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.* **10** 1025–1064.

NAGAEV, S. (1979). Large deviations of sums of independent random variables. *Ann. Probab.* **7** 745–789.

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

PROKHOROV, Y. (1959). An extremal problem in probability theory. *Theory Probab. Appl.* **4** 201–204.

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

ROLSKI, R., SCHMIDLI, H., TEUGELS, J. and SCHMIDT, V. (1999). *Stochastic Processes for Insurance and Finance*. Wiley, New York.

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

DEPARTMENT OF MATHEMATICS
ETHZ
CH-8092 ZURICH
SWITZERLAND
E-MAIL: embrechts@math.ethz.ch

SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853
E-MAIL: gennady@orie.cornell.edu