Transition phenomena for the maximum of a random walk with small drift

Johannes Kugler

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

eingereicht am 9. September 2014 von
Johannes Marian Kugler
aus Mannheim

Transition phenomena for the maximum of a random walk with small drift

Johannes Kugler

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

eingereicht am 9. September 2014 von
Johannes Marian Kugler
aus Mannheim

Erster Berichterstatter: Prof. Dr. Vitali Wachtel

Zweiter/Externer Berichterstatter: Prof. Dr. Ron Doney

Tag der Disputation: 11.11.2014

Eidesstattliche Versicherung:

Hiermit erkläre ich an Eidesstatt, dass diese Dissertation von mir selbstständig, unabhängig und ohne unerlaubte Beihilfe angefertigt ist.

Ich habe diese Dissertation in der gegenwärtigen oder einer ähnlichen Fassung an keiner anderen Fakultät eingereicht.

München, 14.04.2015

Johannes Kugler

Acknowledgement

The help I received from my supervisor, Vitali Wachtel, cannot be overestimated. I appreciate his support during my doctoral studies and I am grateful for his unfailing interest in my own and our joint research and his numerous advices.

I am also thankful to my co-author Denis Denisov for our fruitful joint work in Manchester and Munich. I would also like to express my gratitude to Seva Shneer for his help with various questions and the warm welcome during my stay in Edinburgh.

I am grateful to my colleagues Kai Beermann, Robert Graf and Andrej Nikonov not only for proofreading my thesis but also for making me have a nice time during my doctoral studies.

Especially, I would like to thank my parents and my grandmother for their financial and personal assistance during my education and my partner Karin Stadler for being supportive in every aspect of my life. Last but not least, i am grateful to Fabian Petersen for proofreading the thesis.

This work was partially supported by a grant of the Deutsche Forschungsgemeinschaft and I am grateful to them for making this thesis possible.

Abstract

This thesis examines the tail behaviour of the maximum $M^{(a)}$ of a random walk with negative drift $-a$. It consists of four chapters.

Chapter 1 contains an introduction to heavy tailed and subexponential distributions. It also comprises a brief introduction to the theory of random walks and a survey of some known results concerning the maximum of a random walk $S^{(a)}$.

It is known that, for fixed x, the probability $\mathbf{P}(M^{(a)} > x/a)$ is exponential as $a \to 0$ (heavy traffic asymptotics) and, for (strong) subexponential distributions, $\mathbf{P}(M^{(a)} > x)$ decays according the the integrated tail as $x \to \infty$ for fixed a (heavy tail asymptotics). The second Chapter presents a link between these two asymptotics. In particular, the behaviour of the probability $\mathbf{P}(M^{(a)} > x)$ is studied as $a \to 0$ for x such that $x \to \infty$ as $a \to 0$ and the regions of x for which the heavy traffic asymptotics and the heavy tail asymptotics hold are identified. Furthermore, the distributions for which an intermediate zone between these two limits exists are identified and the exact limit in this zone is provided. The approach in this chapter is not based on a representation via geometric sums, like most of the results on the behaviour of $\mathbf{P}(M^{(a)} > x)$ are. Instead, martingale arguments and inequalities are used.

Chapter 3 contains non-asymptotic results on the maximum of a random walk. Namely, it comprises computable upper bounds for the probability $\mathbf{P}(M^{(a)} > x)$ for fixed a and x in different settings of power moment existences. As it is usual for deriving upper bounds, these upper bounds are attained by truncation of summands. The approach used for the truncation is to split the time axis by stopping times into intervals of random but finite length and then choose a level of truncation on each interval. Hereby one can reduce the problem of finding upper bounds for $M^{(a)}$ to the problem of finding upper bounds for $M^{(a)}_{\tau_z} = \max_{n \le \tau_z} S^{(a)}_n$, where $\tau_z = \min\{n \ge 1 : S^{(a)}_n < -z\}$, $z > 0$. Additionally, the obtained inequalities are tested in the heavy traffic and heavy tail regime for regular varying tails and it is shown that they are asymptotically precise in this case.

The fourth Chapter deals with the case of a family of $\Delta^{(a)}$-latticed random walks and provides a local version of the heavy traffic asymptotics for the probability $\mathbf{P}(M^{(a)} = \Delta^{(a)} x)$ for x such that $x \to \infty$ and $ax = O(1)$ as $a \to 0$. This local limit theorem follows from a representation of $\mathbf{P}(M^{(a)} = \Delta^{(a)} x)$ via a geometric sum and a uniform renewal theorem, which is also proved in this chapter.

Zusammenfassung

Diese Dissertation untersucht das Tail-Verhalten des Maximums $M^{(a)}$ einer Irrfahrt mit negativem Drift $-a$. Sie besteht aus vier Kapiteln.

Kapitel 1 gibt eine kurze Einführung in die Theorie von Heavy Tail und subexponentiellen Verteilungen. Es enthält auch eine Einführung in die Theorie von Irrfahrten und einen Überblick über einige bekannte Resultate bezüglich dem Maximum einer Irrfahrt $S^{(a)}$.

Es ist bekannt, dass die Wahrscheinlichkeit $\mathbf{P}(M^{(a)} > x/a)$ für $a \to 0$ exponentiell in x ist (Heavy Traffic Asymptotik). Für $x \to \infty$ und festes a hingegen fällt $\mathbf{P}(M^{(a)} > x)$ im Fall von (stark) subexponentiellen Zuwächsen gemäß dem integrierten Tail ab (Heavy Tail Asymptotik). Das zweite Kapitel verbindet diese zwei Resultate. Darin wird das asymptotische Verhalten der Wahrscheinlichkeit $\mathbf{P}(M^{(a)} > x)$ für $a \to 0$ untersucht, falls x so ist, dass $x \to \infty$ für $a \to 0$. Des Weiteren werden die Regionen von x identifiziert, für welche die Heavy Traffic Asymptotik bzw. die Heavy Tail Asymptotik gelten. Darüber hinaus wird hergeleitet, für welche Verteilungen eine weitere Region existiert, in der weder die Heavy Traffic Asymptotik noch die Heavy Tail Asymptotik gilt und der exakte Grenzwert in dieser Region wird aufgezeigt. Der Ansatz in diesem Kapitel basiert nicht auf der Darstellung des Maximums durch eine geometrische Summe, wie die meisten Resultate zum asymptotischen Verhalten von $\mathbf{P}(M^{(a)} > x)$. Stattdessen werden Martingalargumente und Ungleichungen benutzt.

Kapitel 3 enthält nicht-asymptotische Resultate bezüglich des Maximums einer Irrfahrt. Es beinhaltet konkrete obere Schranken für die Wahrscheinlichkeit $\mathbf{P}(M^{(a)} > x)$ für fixe a und x unter verschiedenen Momentannahmen. Diese obere Schranken werden durch Abschneiden von Summanden hergeleitet. Genauer wird die Zeitachse mittel Stopp- zeiten in Intervalle von zufälliger aber endlicher Länge zerlegt und dann auf jedem Intervall ein Abschneidungsniveau gewählt. Dadurch kann man aus oberen Schranken für $M^{(a)}_{\tau_z} = \max_{n \le \tau_z} S_n$, wobei $\tau_z = \min\{n \ge 1 : S^{(a)}_n < -z\}$, $z > 0$, auch obere Schranken für $M^{(a)}$ herleiten. Zusätzlich werden die hergeleiteten Ungleichungen im Fall von regulär variierenden Verteilungen in den Heavy Traffic und Heavy Tail Regionen getestet und es wird gezeigt, dass diese dort asymptotisch präzise sind.

Das vierte Kapitel beschäftigt sich mit dem Fall von Irrfahrten auf einem Gitter mit Gitterabstand $\Delta^{(a)}$. Es wird eine lokale Version der Heavy Traffic Asymptotik für die Wahrscheinlichkeit $\mathbf{P}(M^{(a)} = \Delta^{(a)} x)$ bewiesen, falls x so ist, dass $x \to \infty$ und $ax = O(1)$ für $a \to 0$. Diese Asymptotik folgt aus einer Darstellung von $\mathbf{P}(M^{(a)} = \Delta^{(a)} x)$ mittels einer geometrischen Summe und einem uniformen Erneuerungstheorem, welches auch bewiesen wird.

Preface

This thesis consists of an introduction and four closely related chapters which deal with the study of the asymptotic and non-asymptotic behaviour of the maximum of a random walk.

Chapter 1 serves as a joint introduction to the chapters 2 - 4. In Chapter 2 the asymptotical tail behaviour of the maximum of a random walk is fully described, Chapter 3 presents various upper bounds for the tail of the maximum and Chapter 4 comprises a local limit theorem for the maximum of a random walk.

Inhaltsverzeichnis

Introduction

The total maximum $M^{(a)}$ of a random walk $S^{(a)}$ with drift $-a$ plays a crucial role in a number of applications. For example, its distribution coincides with the stationary distribution of the queue-length in a $G/G/1$ queue. Another important application comes from insurance mathematics: Under some special restrictions on the increments $X_i^{(a)}$ the quantity $\mathbf{P}(M^{(a)} > x)$ is equal to the ruin probability in the so-called renewal arrivals model.

The asymptotic tail behaviour of the maximum of a random walk has been studied extensively in the literature. The first result goes back to Cramér and Lundberg (see, for example, Asmussen [3]). They considered light-tailed increments $X_i^{(a)}$ with fixed $a > 0$ and showed that if the so-called Cramér condition $\mathbf{E}[e^{h_0 X_1^{(a)}}] = 1$ is fulfilled for some $h_0 > 0$ and $\mathbf{E}[X_1^{(a)} e^{h_0 X_1^{(a)}}] < \infty$, then there exists a constant $c_0 \in (0, 1)$ such that

$$\mathbf{P}(M^{(a)} > x) \sim c_0 e^{-h_0 x} \quad \text{as } x \to \infty. \tag{0.1}$$

If the Cramér condition is not fulfilled, that means the distribution F of the increments $X^{(a)}$ is heavy tailed, the most classical result for the asymptotics of $M^{(a)}$ is due to Veraverbeke [48] (see also Embrechts et al. [21]): Denote by $F(\cdot)$ the distribution function of $X_1^{(a)}$ and let $\overline{F}(\cdot) = 1 - F(\cdot)$. Suppose the integrated tail $\overline{F}^I(\cdot) = \min\{1, \int_\cdot^\infty \overline{F}(u)du\}$ is subexponential. Then, for fixed a,

$$\mathbf{P}(M^{(a)} > x) \sim \frac{1}{a}\overline{F}^I(x) \quad \text{as } x \to \infty. \tag{0.2}$$

The assumption of the integrated tail being subexponential is not equivalent to the assumption that the tail distribution is subexponential, see Klüppelberg [34]. But Klüppelberg [34] has shown that if F is strong subexponential, then the integrated tail is subexponential and it is known that strong subexponential distributions form a large subclass of heavy tailed distributions.

Let us consider the case $a \to 0$. For fixed $a > 0$ the random walk $S^{(a)}$ drifts to $-\infty$ and the total maximum $M^{(a)}$ is finite almost surely. However, as $a \to 0$, $M^{(a)} \to \infty$ in probability. From this fact arises the natural question how fast $M^{(a)}$ grows as $a \to 0$. The studies on this question were initiated by Kingman [33], who considered the case when $|X^{(a)}|$ has an exponential moment, and proved that for fixed x,

$$\mathbf{P}(M^{(a)} > x/a) \sim e^{-2x/\sigma^2} \quad \text{as } a \to 0, \tag{0.3}$$

where $\sigma^2 = \mathbf{Var}(X^{(0)})$ denotes the variance of the increments in the case of zero drift. Prohorov [45] extended this result to the case that the increments have finite variance. Kingman and Prohorov had a motivation for examining $M^{(a)}$ that comes from queueing theory: As mentioned above, it is well known that the stationary distribution of the waiting time of a customer in a single-server first-come-first-served queue coincides with the distribution of the maximum of a corresponding random walk. In the context of queueing theory, the limit $a \to 0$ means that the traffic load tends to 1. Thus, the question on the distribution of $M^{(a)}$ may be seen as the question on the growth rate of a stationary waiting-time distribution in a $G/G/1$ queue. This is one of the most important questions in queueing theory and is usually referred to as heavy traffic analysis.

One can see that (0.3) has a form similar to (0.1). Indeed, if the Cramér condition holds, then, it is known that $h_0 \to 0$ as $a \to 0$ and in the limit (0.1) becomes (0.3), see e.g Asmussen [2]. In the special case that the increments are normal distributed with expected value $-a$ and variance σ^2, one has $\mathbf{E}[e^{hX^{(a)}}] = e^{-ha+h^2\sigma^2/2}$ and therefore immediately attains $h_0 = 2a/\sigma^2$.

Now suppose we let $a \to 0$ and $x \to \infty$ simultaneously in the subexponential case. If $a \to 0$ much ßlower"than $x \to \infty$, the probability $\mathbf{P}(M^{(a)} > x)$ should still behave like in the integrated tail approximation (0.2). On the other hand, if $a \to 0$ much "faster"than $x \to \infty$, the heavy traffic approximation from (0.3) should still be precise. This fact raises the interesting mathematical question what "fasteränd ßlower"mean in this context and how the exponential asymptotics turns into the integrated tail asymptotics. In particular, it is of interest whether there exists a transition point, at which the transition from (0.3) to (0.2) takes place. Or, otherwise, whether there is a region in which neither the heavy traffic nor the heavy tail asymptotics holds and what the asymptotical behaviour of $\mathbf{P}(M^{(a)} > x)$ will be like in this region. As shown in this thesis, answers to these questions depend on the distributions of the increments of the random walk.

Blanchet and Lam [7] (see also Blanchet and Glynn [6]) extended the heavy traffic asymptotics (0.3) to the case when $x \to \infty$. They have shown that if $x \to \infty$ sufficiently slow as $a \to 0$,

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} \quad \text{as } a \to 0, \tag{0.4}$$

where θ_a is the solution to the equation

$$\mathbf{E}\left[e^{\theta_a X^{(a)}}; X^{(a)} \le 1/a\right] = 1. \tag{0.5}$$

By the Taylor expansion one can see (also refer to Blanchet and Glynn [6]) that θ_a allows an expansion of the form $\theta_a = 2a/\sigma^2 + C_2 a^2 + \ldots + C_k a^k$, where $C_i, i \in 2, \ldots, k$, are suitable constants. This expansion is valid up to the order of the moment existence of $X^{(a)}$ and the constants C_i can be defined using these moments.

Another remarkable result, which covers various subexponential distributions (including regular varying and Weibullian), is also contained in Blanchet and Lam [7]. They have recently found a uniform, explicit representation for the probability $\mathbf{P}(M^{(a)} > x)$, which consists of the exponential term from Kingman's asymptotics, the integrated tail term and a convolution-type integral of a negative binomial sum.

The reason why all these results only work in the setting of an $M/G/1$ queue is that their approach is based on the representation of $M^{(a)}$ as a geometric sum of independent random variables:

$$\mathbf{P}(M^{(a)} > x) = \sum_{k=0}^{\infty} q(1-q)^k \mathbf{P}(\chi_1^+ + \chi_2^+ \cdots + \chi_k^+ > x), \tag{0.6}$$

where $\{\chi_l^+\}$ are independent random variables and $q = \mathbf{P}(M^{(a)} = 0)$. The main difficulty in this approach is the fact that one has to know the distribution of χ_l^+ and the parameter

q. In the setting of a $M/G/1$ queue the value q is known and it remains to find good estimates for the probability $\mathbf{P}(\chi_l^+ > x)$ and in the case of a $M/M/1$ queue both values, q and $\mathbf{P}(\chi_l^+ > x)$, can be calculated. However, in general one has to obtain appropriate estimates for q and $\mathbf{P}(\chi_l^+ > x)$. Therefore the approach via the representation as a geometric sum may be unsuitable for general distributions of the increments (which corresponds to the case of a $G/G/1$ queue).

Chapter 1 of this thesis comprises a brief introduction to heavy tailed and subexponential distributions and states known results on the maximum of random walks. In Chapter 2 we show that for a subclass of heavy tailed distributions (which includes regular varying and lognormal distributions) one has, uniform in $x \geq 0$,

$$\mathbf{P}(M^{(a)} > x) \sim e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\} \quad \text{as } a \to 0, \tag{0.7}$$

where $x(a)$ denotes a value for which both terms on the right hand side of the latter relation are of the same order. For another subclass of heavy tailed distributions (including Weibullian and semi-exponential distributions) one has, uniform in $x \geq 0$,

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x \geq x(a) - C\ln(1/a)/\theta_a\}}{a(1-\gamma g(x)/(\theta_a x))^2} \quad \text{as } a \to 0 \tag{0.8}$$

with $\gamma \in (0,1)$ and a large constant C. The exact value of γ depends on the tail distribution $\overline{F}(\cdot)$ and the exact dependence of x and a. One has $\limsup_{a\to 0} g(x)/(\theta_a x) = 0$ for all x such that $x \gg x(a)$ as $a \to 0$ and $\liminf_{a\to 0} g(x)/(\theta_a x) > 0$ for all x such that x and $x(a)$ are of the same order. To prove (0.7) and (0.8) we do not use the approach via geometric sums from (0.6), instead we use an approach which relies on martingale methods. Appearance of martingales is due to the equation $(M^{(a)} + X^{(a)})^+ \overset{d}{=} M^{(a)}$. In [20], Denisov and Wachtel also used a martingale technique to reestablish (0.2) for long-tailed distributions. Another important contribution of Chapter 2 (see also Chapter 3) treats the case that the increments are regular varying of index $r > 2$. It is shown that in this case there exists a sharp transition point

$$x_{RV}(a) \approx \frac{\sigma^2(r-2)}{2}\frac{1}{a}\ln\frac{1}{a}. \tag{0.9}$$

This means that, for values x above the critical value, the heavy tail approximation holds and under this value the heavy traffic asymptotics is valid. Furthermore, $x_{RV}(a)$ is a value for which the two terms on the right hand side of (0.7) are of the same order. This generalizes a result from Olvera-Cravioto, Blanchet and Glynn [43], who derived this critical value in the setting of a $M/G/1$ queue. If the increments possess a Weibull distribution, that is $\overline{F}(x) = e^{-x^\gamma}$ with $\gamma \in (0,1)$, one could believe that there is still a sharp transition point. Equating the right hand sides of (0.2) and (0.4), one guesses the critical point would be

$$x_W(a) \approx \left(\frac{1}{\theta_a}\right)^{1/(1-\gamma)} - \frac{2}{\theta_a(1-\gamma)}\ln\frac{\sqrt{2/(\gamma\sigma^2)}}{\theta_a}.$$

In [44], Olvera-Cravioto and Glynn conjectured that for Weibull type distributions there is a third region in which neither the heavy traffic nor the integrated tail asymptotic is valid only if $1/2 < \gamma < 1$. However, we show that this is not the case and in fact this third region exists for a larger region, that is for all $\gamma \in (0,1)$. This third region turns out to be the region in which the integrated tail term is at least of the same order as the exponential term on the right hand side of (0.8) and in which $x/x_W(a) = O(1)$.

In (0.7) and (0.8) we fully describe the asymptotical behaviour of $\mathbf{P}(M^{(a)} > x)$ as $a \to 0$ uniform in x. However, for applications in insurance mathematics and queueing theory, it is also of great interest to have non-asymptotical approximations for the distribution of $M^{(a)}$. Especially, it is important to have computable upper bounds for the probability $\mathbf{P}(M^{(a)} > x)$ if a and x are fixed values and this is what Chapter 3 is about. The most classical result in this field is the Lundberg inequality, which states that if the Cramér condition holds for $h_0 > 0$, one has

$$\mathbf{P}(M^{(a)} > x) \le e^{-h_0 x}$$

for all fixed $a, x > 0$. Because of (0.1), the error in the latter inequality is only of constant order and the Lundberg inequality is therefore quite precise.

If the Cramér condition is not fulfilled, upper bounds for $\mathbf{P}(M^{(a)} > x)$ have been derived by Kalashnikov [31] and by Richards [46]. The approach in these papers is again based on the representation (0.6) of $M^{(a)}$ as a geometric sum of independent random variables. Our approach is different. We split the time axis into intervals of finite but random length and choose a level of truncation on these intervals. This gives

$$\mathbf{P}(M^{(a)} > x) \le \sum_{j=0}^{\infty} \mathbf{P}(M_{\tau_z}^{(a)} > x + jz),$$

where $\tau_z = \min\{k \ge 1 : S_k < -z\}$ with arbitrary $0 \le z \le x$. This formula allows us to obtain upper bounds for $M^{(a)}$ from upper bounds for $M_{\tau_z}^{(a)}$. In the case of finite and infinite variance we get upper bounds for the probability $\mathbf{P}(M_{\tau_z}^{(a)} > x)$ by a martingale construction and therefore by the latter formula upper bounds for the probability $\mathbf{P}(M^{(a)} > x)$: Fix some $\theta \in (0,1)$, $\varepsilon > 0$, $\alpha \in (0,1)$ and let $\beta = 1 - \alpha$. Define $A_t := \mathbf{E}[|X^{(a)}|^t]$, $A_{t,+} := \mathbf{E}[(X^{(a)})^t; X^{(a)} > 0] > 0$,

$$c_1 := \frac{3A_t^{1/\theta}\theta^{-(t-1)/\theta}}{(t-1)a^{1/\theta-1}}, \quad c_2 := \frac{3A_{t,+}^{1/\theta}\theta^{-(t-1)/\theta}}{(t-1)\beta^{1/\theta}a^{1/\theta-1}},$$

$$\psi_3(x) := \frac{a\theta^{t-1}x^{t-1}}{A_t}, \quad \psi_4(x) := \frac{\beta\theta^{t-1}ax^{t-1}}{A_{t,+}}.$$

Assume that $A_t < \infty$ for some $t \in (1,2]$. Then, for every a, x satisfying $x^{t-1} \ge \theta^{1-t}(e^\theta - 1)A_t a^{-1}$ and $x \ge z(t-1)\theta^{-1}$, we have

$$\mathbf{P}(M^{(a)} > x) \le c_1 \frac{\mathbf{E}[\tau_z]}{z} \ln\left(1 + \psi_3(x)\right) x^{-(t-1)/\theta} + \left(1 + \psi_3(x)^{-1/\theta}\right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z}\overline{F}^I(\theta x) + \mathbf{P}(X^{(a)} > \theta x)\right).$$

Assume that $\mathbf{Var}(X^{(a)}) < \infty$ and $0 < A_{t,+} < \infty$ for some $t > 2$. Then, for every a, x, z satisfying the conditions $2\alpha e^{-\varepsilon}\theta a x \geq \mathbf{E}[(X^{(a)})^2] \ln(1 + \beta\theta^{t-1} a x^{t-1}/A_{t,+})$, $x^{t-1} \geq \theta^{1-t}(e^\theta - 1)A_{t,+}\beta^{-1}a^{-1}$ and $x \geq z(t-1)\theta^{-1}$, we have

$$\mathbf{P}(M^{(a)} > x) \leq c_2 \frac{\mathbf{E}[\tau_z]}{z} \ln\left(1 + \psi_4(x)\right) x^{-(t-1)/\theta} + \left(1 + \psi_4(x)^{-1/\theta}\right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z} \overline{F}^I(\theta x) + \mathbf{P}(X^{(a)} > \theta x)\right).$$

In Chapter 3 we also test our inequalities in the heavy traffic and heavy tail regime and show that they are asymptotically precise. Particularly we consider the case of regular varying tail-distributions and reestablish the results (0.3) for x under the critical value from (0.9) and (0.2) above this critical value. This means we reestablish the result (0.7) in the case that x is not asymptotically equivalent to $x_{RV}(a)$. Furthermore, we use our bounds on M_{τ_z} to obtain a result on the asymptotics of $M^{(a)}$ in the case of infinite variance. Namely, we show that if $\mathbf{E}[(\min\{0, X^{(0)}\})^2] < \infty$, (0.2) is still valid above some critical value for regular varying distributions with index $r \in (1, 2]$.

The fourth Chapter of this thesis deals with the question whether there exists a local version of the heavy traffic asymptotics (0.3) if the increments possess an aperiodic $\Delta^{(a)}$-lattice distribution. We consider the case when $x \to \infty$ with $ax = O(1)$ as $a \to 0$ and show that

$$\mathbf{P}(M^{(a)} = \Delta^{(a)} x) \sim \frac{2a\Delta^{(0)}}{\sigma^2} \exp\left\{-\frac{2ax\Delta^{(0)}}{\sigma^2}\right\} \quad \text{as } a \to 0.$$

This result follows from a representation of $\mathbf{P}(M^{(a)} = \Delta^{(a)} x)$ as a geometric sum and the application of a uniform renewal theorem which is also derived in this chapter.

1 Preliminaries

This chapter contains an overview of the used notation and a brief introduction to the theory of random walks including a survey of some known results on the maximum of a random walk.

1.1 Notation

In this thesis the following notation is used:

- By $\searrow$ and $\nearrow$ we mean (weakly) decreasing and (weakly) increasing respectively.
- The symbol $\sim$ is used to denote äsymptotically equivalent". Thus, for two non-negative functions f and g and a constant $a \in [-\infty, \infty]$,

$$f(x) \sim g(x) \text{ as } x \to a \Longleftrightarrow \lim_{x\to a} f(x)/g(x) = 1.$$

- The symbols o, O and $\asymp$ are used to indicate "having smaller order", "not having larger orderänd "having the same order"respectively. That is for non-negative functions f and g and a constant $a \in [-\infty, \infty]$,

$$f(x) = o(g(x)) \text{ as } x \to a \Leftrightarrow \lim_{x\to a} \frac{f(x)}{g(x)} = 0,$$
$$f(x) = O(g(x)) \text{ as } x \to a \Leftrightarrow \limsup_{x\to a} \frac{f(x)}{g(x)} < \infty,$$
$$f(x) \asymp g(x) \text{ as } x \to a \Leftrightarrow \liminf_{x\to a} \frac{f(x)}{g(x)} > 0 \quad \text{and} \quad \limsup_{x\to a} \frac{f(x)}{g(x)} < \infty.$$

- The symbols $\gg$ and $\ll$ are used to indicate "having larger orderänd "having smaller order"respectively. For non-negative functions f and g and a constant $a \in [-\infty, \infty]$,

$$f(x) \gg g(x) \text{ as } x \to a \Leftrightarrow g(x) = o(f(x)) \text{ as } x \to a \Leftrightarrow \lim_{x\to a} \frac{f(x)}{g(x)} = \infty,$$
$$f(x) \ll g(x) \text{ as } x \to a \Leftrightarrow f(x) = o(g(x)) \text{ as } x \to a \Leftrightarrow \lim_{x\to a} \frac{f(x)}{g(x)} = 0.$$

- The symbols $\succ$ and $\prec$ are used to refine the notation from $\gg$ and $\ll$. In particular, for two non-negative functions f and g and a constant $a \in [-\infty, \infty]$,

$$f(x) \succ g(x) \text{ as } x \to a \Leftrightarrow f(x) - g(x) \to \infty \text{ as } x \to a,$$
$$f(x) \prec g(x) \text{ as } x \to a \Leftrightarrow f(x) - g(x) \to -\infty \text{ as } x \to a.$$

- The sign $\stackrel{d}{=}$ is used to denote "have the same distribution". That is for two random variables X, Y on $(\Omega, \mathcal{F}, \mathbf{P})$,

$$X \stackrel{d}{=} Y \Leftrightarrow \mathbf{P}(X \in A) = \mathbf{P}(Y \in A) \quad \text{for all } A \in \mathcal{F}.$$

- For two non-negative functions $f(x,y)$ and $g(x,y)$ and a constant $a \in [-\infty, \infty]$,

$$f(x,y) = o(g(x,y)) \text{ as } x \to a \text{ uniform in } y \Leftrightarrow \lim_{x\to a} \sup_{y} \frac{|f(x,y)|}{|g(x,y)|} = 0.$$

- For a distribution function F denote by $\overline{F}$ the (right-)tail distribution function, i.e. $\overline{F}(x) = 1 - F(x)$.
- $\overline{F}^I(\cdot)$ is the integrated tail defined by $\overline{F}^I(x) = \min\left\{1, \int_x^\infty \overline{F}(u)du\right\}$, $x \geq 0$, and $F^I(\cdot)$ is defined by $F^I(x) = 1 - \overline{F}^I(x)$, $x \geq 0$.
- F^{*n} stands for the n-fold convolution of F with itself and the corresponding right tail is defined as $\overline{F^{*n}}(x) = 1 - F^{*n}(x)$.
- The indicator function is displayed by $\mathbf{1}\{A\}$ and is 1 if A holds and 0 if it does not.
- For a random variable X and the expectation $\mathbf{E}$, the expression $\mathbf{E}[X; A]$ is used to abbreviate $\mathbf{E}[X\mathbf{1}\{A\}]$.
- For and random variable X and $a, b \in \mathbb{R}$ with $a < b$, we use the convention $\mathbf{E}[X; X \in [b,a]] = -\mathbf{E}[X; X \in [a,b]]$
- The term $M/G/1$ queue is used to denote a first in first out (FIFO) queue with a markovian (exponential) interarrival distribution, general but independent service time distribution and 1 server. Accordingly, a $G/G/1$ queue is used to denote a FIFO queue with a general but independent interarrival distribution.
- For a random variable X the positive part of X is defined as $X^+ := X\mathbf{1}\{X \geq 0\}$ and the negative part is defined as $X^- := -X\mathbf{1}\{X \leq 0\}$.

Additional notation is introduced during the course of this chapter.

In the following, we give some definitions and state some well known results on heavy tailed and subexponential distributions, mostly collected from [27]. For an extensive introduction to heavy tailed distributions, see e.g. [11].

1.2 Heavy tailed distributions

Consider a random variable X on $\mathbb{R}$ with distribution F.

Definition 1.1. The distribution F is said to have *right-unbounded support* if $\overline{F}(x) > 0$ for all $x > 0$.

Definition 1.2. The distribution F is called *(right-)heavy tailed* if

$$\int_{-\infty}^{\infty} e^{\lambda x} dF(x) = \infty \quad \text{for all } \lambda > 0,$$

that is, if and only if F does not possess any positive exponential moment. Otherwise the distribution F is called light-tailed.

If the distribution F is concentrated on the positive real axis $\mathbb{R}_0^+$, the latter definition immediately implies that for a light-tailed distribution all power moments are finite.

Definition 1.3. A function $f(\cdot) \geq 0$ is called *heavy tailed* if

$$\limsup_{x\to\infty} e^{\lambda x} f(x) = \infty \quad \text{for all } \lambda > 0.$$

It is known that a distribution F is heavy tailed if and only if its tail distribution function $\overline{F}(\cdot)$ is heavy tailed. Hence, for a distribution F to be heavy tailed is a tail-property.

1.3 Subexponential distributions

For various results one requires stronger regularity conditions than the requirement that the distribution is heavy tailed. Recall that for all independent, non-negative random variables $X_1, X_2, \ldots, X_n$ with distribution F, as $x \to \infty$,

$$\begin{aligned} \overline{F^{*n}}(x) &= \mathbf{P}(X_1 + \ldots + X_n > x) \geq \mathbf{P}(\max\{X_1, \ldots, X_n\} > x) \\ &= 1 - (1 - \mathbf{P}(X_1 > x))^n \sim n\mathbf{P}(X_1 > x) = n\overline{F}(x). \end{aligned} \tag{1.1}$$

This motivates the definition of a so-called subexponential distribution.

Definition 1.4. We say a distribution F on $\mathbb{R}_0^+$ with unbounded-support is *subexponential*, and write $F \in \mathcal{S}$, if

$$\overline{F^{*2}}(x) \sim 2\overline{F}(x) \quad \text{as } x \to \infty. \tag{1.2}$$

Subexponential distributions were introduced independently in [14] and [15] and it is known that all subexponential distributions are heavy tailed, see e.g. [23]. Embrechts and Hawkes [22] have shown that the property (1.2) can be generalized to an arbitrary number n instead of 2. To be more specific, the following characterization of subexponential distributions is valid.

Proposition 1.5. *A distribution F on $\mathbb{R}_0^+$ with unbounded support is subexponential if and only if, for an arbitrary $n \geq 2$,*

$$\overline{F^{*n}}(x) \sim n\overline{F}(x) \quad \textit{as } x \to \infty. \tag{1.3}$$

Furthermore, (1.1) and (1.3) imply that F is subexponential if and only if, for all $n \geq 2$,

$$\mathbf{P}(X_1 + \ldots + X_n > x) \sim \mathbf{P}(\max\{X_1, \ldots, X_n\} > x) \quad \text{as } x \to \infty. \tag{1.4}$$

Relation (1.4) can be interpreted as follows: if a random walk exceeds a large level x this is with a probability close to 1 due to the fact that one increment of the random walk exceeds this level. Therefore this result is usually referred to as the "principle of a single big jump".

Subexponential distributions can also be defined on the whole real line. To do so let us introduce another class of distribution functions.

Definition 1.6. A distribution on $\mathbb{R}$ is said to be *long-tailed* if it has right-unbounded support and, for every fixed value $y > 0$,

$$\overline{F}(x+y) \sim \overline{F}(x) \quad \text{as } x \to \infty. \tag{1.5}$$

Every long-tailed distribution is also heavy tailed, see Lemma 2.17 and Theorem 2.6 in [27]. However, (1.5) implies a degree of smoothness and not all heavy tailed distribution function possess this smoothness, see the example after Lemma 2.17 in [27]. It is known that a distribution F on $\mathbb{R}$ with unbounded support is long-tailed with $\overline{F^{*2}}(x) \sim 2\overline{F}(x)$ as $x \to \infty$ if and only if $F^+ := F(x)\mathbf{1}\{x \geq 0\}$ is subexponential. This allows us to extend the definition of subexponentiality to the whole real line.

Definition 1.7. We say a distribution F on $\mathbb{R}$ is called *(whole-line) subexponential*, and write $F \in \mathcal{S}_{\mathbb{R}}$, if F is long-tailed and

$$\overline{F^{*2}}(x) \sim 2\overline{F}(x) \quad \text{as } x \to \infty. \tag{1.6}$$

1.4 Strong subexponential distributions

Consider a random variable X with support $\mathbb{R}$ and distribution F. In applications like random walk theory, queueing theory, risk theory and renewal theory, and especially in the following chapters it is an important question whether $\overline{F} \in \mathcal{S}$ implies that $\overline{F}^I \in \mathcal{S}$. In general this is not the case, see for example Chapter 3.8. in [27]. However, Klüppelberg [34] has shown that those distribution functions for which the latter is true form a large subclass of $\mathcal{S}$, which we will call strong subexponential distributions.

Definition 1.8. Suppose $\mu := \mathbf{E}[X; X \geq 0] < \infty$. A distribution function F with right-unbounded support belongs to the class $\mathcal{S}^*$ of strong subexponential distribution functions if

$$\int_0^x \overline{F}(x-y)\overline{F}(y)dy \sim 2\mu\overline{F}(x) \quad \text{as } x \to \infty. \tag{1.7}$$

This definition can be motivated as follows: For all distributions F on $\mathbb{R}$,

$$\int_{x/2}^x \overline{F}(x-y)\overline{F}(y)dy = \int_0^{x/2} \overline{F}(w)\overline{F}(x-w)dw,$$

where we substituted $w = x - y$. Therefore,

$$\int_0^x \overline{F}(x-y)\overline{F}(y)dy = 2\int_0^{x/2} \overline{F}(x-y)\overline{F}(y)dy \geq 2\overline{F}(x)\int_0^{x/2} \overline{F}(y)dy$$

and consequently, for all distributions with right-unbounded support,

$$\liminf_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-y)\overline{F}(y)dy \geq 2\mu.$$

If F is heavy tailed, one can even show (see for example Lemma 4 in [26]) that

$$\liminf_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-y)\overline{F}(y)dy = 2\mu. \tag{1.8}$$

The observation (1.8) provides that a distribution is strong subexponential if it is heavy tailed and sufficiently regular that the limit

$$\lim_{x\to\infty} \frac{1}{\overline{F}(x)} \int_0^x \overline{F}(x-y)\overline{F}(y)dy$$

exists.

As shown in Klüppelberg [34], every strong subexponential distribution on $\mathbb{R}$ is whole-line subexponential and for every strong subexponential distribution the integrated tail distribution is subexponential, that is

$$\mathcal{S}^* \subset \mathcal{S}_{\mathbb{R}}, \tag{1.9}$$

$$F \in \mathcal{S}^* \Rightarrow F^I \in \mathcal{S}. \tag{1.10}$$

In the same paper (see also [27]) an example of a distribution for which $F^I \in \mathcal{S}$ but $F \notin \mathcal{S}^*$ is given. This means the converse of (1.10) is not true. However, it is known that all subexponential (or even heavy tailed) distributions that are likely to be encountered in practice are strong subexponential.

1.5 An important subclass of subexponential distributions

In this chapter we will consider distributions with right-unbounded support and $\overline{F}(x) \sim e^{-g(x)}$ as $x \to \infty$, where g is a positive function. We assume the existence of values $0 < \gamma < 1$ and $x_0 > 0$ (that may depend on γ) such that

$$\frac{g(x)}{x^\gamma} \searrow, \quad x \geq x_0 = x_0(\gamma). \tag{1.11}$$

Let

$$\gamma^* := \inf\{\gamma > 0 : \exists x_0 = x_0(\gamma) : g(x)/x^\gamma \text{ is decreasing for all } x \geq x_0\}.$$

Furthermore, suppose that $\mathbf{E}[|X|^{1+\varepsilon}] < \infty$ for some $\varepsilon > 0$ and if $\gamma^* > 0$ also assume that $e^{-\delta g(x)}$ is integrable over $\mathbb{R}_0^+$ for all $\delta > 0$.

Let us show that this class of distributions is a subclass of the strong subexponential distributions. In order to do so, we want to use a criterion from Theorem 3.30 of [27]: Let F be a long-tailed distribution on $\mathbb{R}$. Assume there exists some $\widehat{\gamma} < 1$ and $A < \infty$ such that

$$g(x) - g(x-y) \leq \widehat{\gamma} g(y) + A \tag{1.12}$$

for all $x > 0$ and $y \in [0, x/2]$. If the function $e^{-(1-\widehat{\gamma})g(x)}$ is integrable over $\mathbb{R}^+$, then $F \in \mathcal{S}^*$. First, consider $\gamma^* = 0$. Then, one can choose $\varepsilon_2 > 0$ arbitrary close to 0 such that

$$\frac{g(x)}{x^{\varepsilon_2}} \searrow, \quad x \geq x_0 = x_0(\varepsilon_2).$$

Hence, for $x \geq 2x_0$,

$$\frac{g(x)}{x^{\varepsilon_2}} \leq \frac{g(x-w)}{(x-w)^{\varepsilon_2}}, \quad 0 \leq w \leq x/2, \tag{1.13}$$

and consequently

$$\frac{g(x-w)}{g(x)} \geq \left(1 - \frac{w}{x}\right)^{\varepsilon_2}, \quad 0 \leq w \leq x/2,$$

for $x \geq 2x_0$. By using

$$(1-z)^{\varepsilon_2} \geq 1 - \varepsilon_2 z - (\varepsilon_2(1-\varepsilon_2))2^{1-\varepsilon_2}z^2 \geq 1 - 2\varepsilon_2 z, \quad 0 \leq z \leq 1/2, \tag{1.14}$$

one obtains, for $x \geq 2x_0$,

$$g(x) - g(x-w) \leq 2\varepsilon_2 w \frac{g(x)}{x}, \quad 0 \leq w \leq x/2. \tag{1.15}$$

Since $g(x)/x$ is decreasing for $x \geq x_0$,

$$\frac{g(x)}{x} \leq \frac{g(w)}{w}, \quad x_0 \leq w \leq x, \tag{1.16}$$

and thus, for $x \geq 2x_0$,

$$g(x) - g(x-w) \leq 2\varepsilon_2 g(w), \quad 0 \leq w \leq x/2.$$

As a consequence, (1.12) is fulfilled for $A = \sup_{u \leq 2x_0} g(u)$ and $\widehat{\gamma} = 2\varepsilon_2$. On the other side, (1.15) implies that for all $y > 0$ fixed

$$\frac{\overline{F}(x+y)}{\overline{F}(x)} \to 1 \quad \text{as } x \to \infty.$$

The existence of a moment of order $1+\varepsilon$ ensures $g(x) \geq (1+\varepsilon)\ln x$ for x large enough and consequently, for ε_2 sufficiently close to 0,

$$e^{-(1-2\varepsilon_2)g(x)} \leq x^{-(1-2\varepsilon_2)(1+\varepsilon)} = o(x^{-1-\varepsilon/2}) \quad \text{as } x \to \infty.$$

Thus, the distribution is strong subexponential due to the above mentioned criterion.

Now, suppose that $\gamma^* > 0$. Then, for all $\varepsilon_2 > 0$ such that $0 < \gamma^* + \varepsilon_2 < 1$ there exists some x_0 such that for $x \geq 2x_0$,

$$\frac{g(x-w)}{g(x)} \geq \left(1 - \frac{w}{x}\right)^{\gamma^*+\varepsilon_2}, \quad 0 \leq w \leq x/2.$$

Furthermore, one sees by the Taylor expansion that for all $0 \leq z \leq 1/2$

$$\begin{aligned}(1-z)^{\gamma^*+\varepsilon_2} &\geq 1 - (\gamma^*+\varepsilon_2)z - (\gamma^*+\varepsilon_2)(1-(\gamma^*+\varepsilon_2))2^{1-(\gamma^*+\varepsilon_2)}z^2 \\ &\geq 1 - (1-(1-(\gamma^*+\varepsilon_2))^2)z.\end{aligned}$$

Therefore, for $x \geq 2x_0$,

$$g(x) - g(x - w) \leq \alpha w \frac{g(x)}{x}, \quad 0 \leq w \leq x/2, \tag{1.17}$$

with $\alpha = 1 - (1 - (\gamma^* + \varepsilon_2))^2 < 1$. The latter inequality combined with (1.16) gives

$$g(x) - g(x - w) \leq \alpha g(w), \quad 0 \leq w \leq x/2, \tag{1.18}$$

for all $x \geq 2x_0$. Hence, (1.12) is fulfilled for $\widehat{\gamma} = \gamma^* + \varepsilon_2$ and $A = \sup_{u \leq 2x_0} g(u)$. Furthermore, by making use of (1.17) and the fact that $g(x)/x$ is decreasing for x large enough, we obtain

$$\frac{\overline{F}(x + y)}{\overline{F}(x)} \to 1 \quad \text{as } x \to \infty$$

for all fixed $y > 0$. On the other side, $e^{-\alpha g(x)}$ is integrable over $\mathbb{R}_0^+$ due to our assumptions and therefore the distribution is strong subexponential by the cited criterion.

1.6 Examples of heavy tailed and (strong) subexponential distributions

Let us discuss a number of examples of heavy tailed distributions. It is known that all these distributions are also strong subexponential and one sees that all these distributions are also contained in the class from Chapter 1.5. In the following chapters we will only be concerned about the subexponentiality of the right tail of the distributions. Therefore we will mostly give examples of distributions on the positive half-line.

Example 1.9. *Pareto distribution on* $\mathbb{R}_0^+$. The tail distribution function of the Pareto distribution is given by

$$\overline{F}(x) = \left(\frac{\kappa}{x + \kappa}\right)^{\alpha} \tag{1.19}$$

with some $\kappa > 0$ and $\alpha > 0$. Clearly, $\overline{F}(x) \sim \kappa^{\alpha} x^{-\alpha}$ as $x \to \infty$ For this reason the Pareto distributions are also referred to as *power law distributions*. For Pareto distributions, all moments of order β exist if and only if $\beta < \alpha$.

Example 1.10. *Distributions with regular varying tails.* Consider a distribution on the positive half-line and let $x_0 \geq 0$. A positive measurable function $L(\cdot)$ defined on $[x_0, \infty)$ is called slowly varying if

$$\frac{L(tx)}{L(x)} \to 1 \quad \text{for all } t > 0. \tag{1.20}$$

Examples for slowly varying function are e.g. logarithm-type functions. The tail distribution function of a distribution with regular varying tails is given by

$$\overline{F}(x) = L(x) x^{-\alpha} \tag{1.21}$$

with index $\alpha > 0$. A distributions on $\mathbb{R}_0^+$ with regular varying tails possess moments of order β if $\beta < \alpha$. Whether the moments of order $\beta = \alpha$ exists, depends on the slowly varying function and moments of order $\beta > \alpha$ do not exist.

Example 1.11. *Lognormal distribution on* $\mathbb{R}_0^+$. This distribution is given by the density

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \tag{1.22}$$

with parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. All power moments of the lognormal distribution are finite and with the l'Hôspital rule one sees that

$$\overline{F}(x) \sim \frac{\sigma^2}{\sqrt{2\pi\sigma^2}\ln x} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) \quad \text{as } x \to \infty.$$

Example 1.12. *Weibull distribution on* $\mathbb{R}_0^+$. The tail function $\overline{F}$ has the form

$$\overline{F}(x) = e^{-x^\gamma} \tag{1.23}$$

with parameter $0 < \gamma < 1$. All power moments of the Weibull distribution are finite.

Example 1.13. *Semi-exponential distribution on* $\mathbb{R}_0^+$. The tail function $\overline{F}$ has the form

$$\overline{F}(x) = e^{-x^\gamma L(x)} \tag{1.24}$$

with parameter $0 < \gamma < 1$ and a slowly varying function $L(x)$. All power moments of Semi-exponential distributions are obviously finite. Semi-exponential distributions were introduced in [9] and in the same article it was shown that all semi-exponential distributions are subexponential.

One can verify that all the distributions in the latter examples are in the subclass of subexponential distributions from chapter 1.5 and therefore also (strong) subexponential and heavy tailed.

For further examples see [27].

1.7 Model and motivation for studying the maximum of a random walk

In this section we give a brief introduction to the model we use in this thesis and motivate why it is important to study the maximum of a random walk.

Denote by $\{S_n^{(a)}, n \geq 0\}, a \in [0, a_0]$ with $a_0 > 0$, a family of random walks with increments $X_i^{(a)}$ and starting point zero, that is,

$$S_0^{(a)} := 0,\ S_n^{(a)} := \sum_{i=1}^{n} X_i^{(a)},\ n \geq 1.$$

We shall assume that, for every fixed a, the random variables $X_1^{(a)}, X_2^{(a)}, \ldots$ are independent copies of a random variable $X^{(a)}$ with distribution function F and negative drift $-a := \mathbf{E}[X^{(a)}] < 0$. Denote by $M^{(a)} = \sup_{k \geq 0} S_k^{(a)}$ the maximum of the corresponding random walk. The maximum plays an important role in a number of applications. For example, its distribution coincides with the steady state waiting time of a $G/G/1$ queue (see Chapter 1.7.1) and can be interpreted as the ruin probability in the so-called renewal arrivals model (see Chapter 1.7.2).

1.7.1 Duality of random walks and queues

One of the main motivations for studying the random variable $M^{(a)}$ originates from queueing theory, since its distribution coincides with the stationary distribution of the queue-length in a $G/G/1$ queue. This was first shown by Kiefer and Wolfowitz [32].

To explain this duality, let us introduce some notation and results collected from Asmussen [2]. Consider a $G/G/1$ queue with customers numbered as $n = 0, 1, \dots$ and assume that customer 0 arrives at time $t = 0$, finds an empty queue, and his service starts immediately. Let T_n be the interarrival time between the n-th customer and the $(n+1)$-th customer and denote the service time of the n-th customer by U_n. W_n is used to denote the waiting time of the n-th customer, i.e. the time between his arrival and beginning of service. Let $\lambda := \mathbf{E}[T_n] \in (0, \infty)$, $b := \mathbf{E}[U_n] \in (0, \infty)$ and denote by $\rho := b/\lambda$ the so-called traffic intensity.

Lemma 1.14. *For $\rho < 1$ there exists a limiting steady state waiting time W_∞ such that the distribution of W_n converges to that of W_∞ in the total variation norm, that is*

$$\sup_A |\mathbf{P}(W_n \in A) - \mathbf{P}(W_\infty \in A)| \to 0 \quad \text{as } n \to \infty.$$

We want to show that the distribution of W_∞ from the latter lemma coincides with the maximum of a properly defined random walk and therefore assume $\rho < 1$ or equivalently $-a := b - \lambda < 0$.

Beweis. The proof uses tightness arguments and can be found in Theorem 2.2 of Chapter 12.2 in [2]. □

Put $Z_n^{(a)} := U_n - T_n$, where the superscript indicates $\mathbf{E}[Z_n^{(a)}] = -a < 0$. Then, for all $n \geq 0$,

$$W_{n+1} = (W_n + Z_n^{(a)})^+. \tag{1.25}$$

Define a random walk by $S_0^{(a)} := 0$ and for $n \geq 1$, $S_n^{(a)} := \sum_{k=1}^n Z_k^{(a)}$. Denote by $M_n^{(a)} := \max_{0 \leq k \leq n} S_k^{(a)}$ the maximum up to time n and by $M := \max_{k \geq 0} S_k^{(a)}$ the total maximum.

Lemma 1.15.

$$W_n = \max\{S_n^{(a)}, S_n^{(a)} - S_1^{(a)}, \dots, S_n^{(a)} - S_{n-1}^{(a)}, 0\} \stackrel{d}{=} M_n^{(a)}. \tag{1.26}$$

Beweis. By (1.25), the increments of (W_n) are at least those of S_n:

$$W_n - W_{n-k} \geq S_n^{(a)} - S_{n-k}^{(a)}, \quad 0 \leq k \leq n.$$

Choosing $k = n$ and using $W_0 = 0$ gives $W_n \geq S_n^{(a)}$ and, by virtue of $W_{n-k} \geq 0$, we get $W_n \geq S_n^{(a)} - S_{n-k}^{(a)}$ for all $k = 0, 1, \dots, n$. Therefore,

$$W_n \geq \max\{S_n^{(a)}, S_n^{(a)} - S_1^{(a)}, \dots, S_n^{(a)} - S_{n-1}^{(a)}, 0\}.$$

For the converse, it is sufficient to show that $W_n = S_n^{(a)} - S_{n-k}^{(a)}$ for some $k = 0, 1, \dots, n$. If $S_k^{(a)} \geq 0$ for all $0 \leq k \leq n$, (1.25) gives $W_n = S_n^{(a)}$. On the other side, if there exists some $1 \leq k \leq n$ such that $S_k^{(a)} < 0$, (1.25) implies that $W_l = 0$ for some $0 \leq l \leq n$. Letting k be the last such l we conclude again by (1.25) that $W_n = S_n^{(a)} - S_{n-k}^{(a)}$.

The second equality of (1.26) is trivial. □

Lemma 1.16. *$M^{(a)} < \infty$ a.s. and*

$$W_n \xrightarrow{w} M^{(a)} \quad \text{as } n \to \infty.$$

Beweis. From the law of large numbers, $S_n^{(a)}/n \xrightarrow{a.s.} -a$ and therefore $S_n^{(a)} \to -\infty$. This ensures the finiteness of $M^{(a)}$. The random variable $\theta := \max\{k \geq 0 : S_k^{(a)} > 0\}$ is finite a.s. and $\mathbf{P}(M^{(a)} \neq M_n^{(a)}) \leq \mathbf{P}(\theta > n) \to 0$ as $n \to \infty$. Thus, Lemma 1.15 gives $W_n \xrightarrow{w} M^{(a)}$. □

Combining the latter results we see that

$$W_\infty \stackrel{d}{=} M^{(a)},$$

so the steady state waiting time can be interpreted as the maximum of an appropriately defined random walk.

In queueing theory heavy tailed distributions appear for example in the modeling of data traffic in communication networks. In this situation, statistical evidence has been found that exponential tail decay is not compatible with the empirical observations, see e.g. Adler et al. [1].

1.7.2 Duality of random walks and ruin probabilities

Another important application comes from insurance mathematics: Under some restrictions on $X^{(a)}$ the quantity $\mathbf{P}(M^{(a)} > x)$ is equal to the ruin probability in the so-called renewal arrivals model.

To outline this duality, let us introduce some notation taken from Asmussen [3]. A risk reserve process $(R_t)_{t \geq 0}$ is a model for the time evolution of the reserves of an insurance company. Denote by $x = R_0$ the initial reserve and suppose the ruin probability $\psi(x)$ is the probability that the reserve ever drops below zero:

$$\psi(x) = \mathbf{P}\left(\inf_{t \geq 0} R_t < 0\right).$$

For mathematical purposes it is often more convenient to work with the claim surplus process $(W_t)_{t \geq 0}$ defined by $W_t = x - R_t$.

Studying ruin probabilities, usually the following setup is used:

- There are only finitely many claims in finite time intervals, which means that the number of arrivals N_t in the time interval $[0, t]$ is finite for all t. Denote the interarrival times by $T_2, T_3, \dots$ and suppose T_1 is the time of the first claim.

- The size of the n-th claim is U_n. The claims are independent of each other and independent of the interarrival times.
- Premiums flow in at rate p, say, per unit time.

Summarizing, we have

$$R_t = x + pt - \sum_{k=1}^{N_t} U_k, \quad W_t = \sum_{k=1}^{N_t} U_k - pt$$

and therefore for the ruin probability

$$\psi(x) = \mathbf{P}\left(\inf_{t\geq 0} R_t < 0\right) = \mathbf{P}\left(\sup_{t\geq 0} W_t > x\right).$$

The case of interest is the nontrivial case of the ruin probability, that is $-a := \mathbf{E}[U_1 - pT_1] < 0$. Put $Y_k^{(a)} := U_k - pT_k$, $k \geq 1$, and

$$S_0^{(a)} := 0, \quad S_n^{(a)} := \sum_{k=1}^{n} Y_k^{(a)}.$$

Due to our assumptions the family $(Y_k^{(a)})_{k\geq 0}$ of random variables is *iid* and since the surplus process $(W_t)_{t\geq 0}$ is monotone decreasing between the claims, the ruin probability is given by

$$\psi(x) = \mathbf{P}\left(\sup_{t\geq 0} W_t > x\right) = \mathbf{P}\left(\max_{k\geq 0} S_k^{(a)} > x\right) = \mathbf{P}(M^{(a)} > x).$$

In actuarial mathematics, there is statistical evidence suggesting that most claim sizes should be modeled as heavy tailed random variables. For a discussion and further references on that see for example Chapter 1.2 of Embrechts, Klüppelberg and Mikosch [23] or Kalashnikov [31].

1.8 Some known results on the maximum of a random walk

We use the setting introduced in chapter 1.7 and state some classical results on the maximum $M^{(a)}$.

The asymptotic tail behaviour of $M^{(a)}$ has been studied extensively in the literature. The first result goes back, apparently, to Cramér and Lundberg (see, for example, Asmussen [2]). If a is fixed,

$$\mathbf{E}[e^{h_0 X^{(a)}}] = 1 \quad \text{for some } h_0 > 0, \tag{1.27}$$

and, in addition, $\mathbf{E}[X^{(a)} e^{h_0 X^{(a)}}] < \infty$, then there exists a constant $c_0 \in (0, 1)$ such that

$$\mathbf{P}(M^{(a)} > x) \sim c_0 e^{-h_0 x} \quad \text{as } x \to \infty. \tag{1.28}$$

If (1.27) is not fulfilled, in other words F is heavy tailed, then one should assume that the distribution of $X^{(a)}$ is regular in some sense. To be more specific, recall that (1.9) states that every strong subexponential distribution is whole-line subexponential and therefore heavy tailed. On the other side, (1.10) implies that the integrated tail $\overline{F}^I(\cdot)$ is subexponential for every strong subexponential distribution function F. In this case, for fixed a, a classical result concerning the maximum of a random walk is the following: If the integrated tail $\overline{F}^I$ is subexponential, then

$$\mathbf{P}(M^{(a)} > x) \sim \frac{1}{a}\overline{F}^I(x) \quad \text{as } x \to \infty. \tag{1.29}$$

This result was proved for regular varying distributions by Callaert and Cohen in [13] and by Cohen [17]. In the present form it was proved by Veraverbeke [48] (see also Embrechts et al. [21]). It is known that (1.29) is valid if and only if $\overline{F}^I$ is subexponential, see e.g. Corollary 6.1 in Embrechts and Veraverbeke [24].

Let us discuss the behaviour of the total maximum as $a \to 0$. This case is interesting in terms of queueing theory as it describes the behaviour of a system in heavy traffic. For all $a > 0$ the total maximum $M^{(a)}$ is finite almost surely. However, $M^{(a)} \to \infty$ in probability as $a \to 0$. From this fact arises the natural question how fast $M^{(a)}$ grows as $a \to 0$. Studies on this question were initiated by Kingman [33], who considered the case when $|X^{(a)}|$ has an exponential moment with $\mathbf{E}[(X^{(a)})^2] \to \mathbf{E}[(X^{(0)})^2] > 0$ as $a \to 0$, and proved that for fixed x,

$$\mathbf{P}(M^{(a)} > x/a) \sim e^{-2x/\sigma^2} \quad \text{as } a \to 0, \tag{1.30}$$

where $\sigma^2 = \mathbf{Var}(X^{(0)})$ denotes the variance of the increments in the case of zero drift. Prohorov [45] extended this result to the case that the increments have finite variance and currently it is known that it is sufficient to assume the Lindeberg-type condition

$$\lim_{a \to 0} \mathbf{E}[(X^{(a)})^2; |X^{(a)}| > K/a] = 0 \quad \text{for all } K > 0$$

to establish (1.30). For an extensive discussion, see e.g. Theorem X.7.1 of Asmussen [2] or equation (21) of [49].

Blanchet and Lam [7] (see also Blanchet and Glynn [6]) generalized (1.30) to the case where x depends on a. In particular, they have shown that if $x = O(1)$ or $x \to \infty$ sufficiently slow as $a \to 0$,

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} \quad \text{as } a \to 0. \tag{1.31}$$

Here, θ_a is the solution to the equation

$$\mathbf{E}\left[e^{\theta_a X^{(a)}}; X^{(a)} \le 1/a\right] = 1. \tag{1.32}$$

By using the Taylor expansion (see also Blanchet and Glynn [6]) one sees that θ_a allows an expansion of the form $\theta_a = 2a/\sigma^2 + C_2 a^2 + \ldots + C_k a^k$, where $C_i, i \in 2, \ldots, k$ are suitable constants. This expansion is valid up to the order of the moment existence of $X^{(a)}$ and the constants C_i can be defined using these moments.

The stated asymptotical results do not necessarily give a good approximation for the probability $\mathbf{P}(M^{(a)} > x)$ for fixed values of a and x. For example, for ßmall"values of x and "large"values of a all the above mentioned approximations are imprecise (for a discussion see Kalashnikov [31]). Therefore it is also of great interest to have non-asymptotic properties of $\mathbf{P}(M^{(a)} > x)$. In the light-tailed case, the first non-asymptotic results on $M^{(a)}$ go back to Cramér and Lundberg (see, for example, Asmussen [3]): If the Cramér condition is fulfilled for some $h_0 > 0$ one has for all $x > 0$ the so-called Lundberg inequality

$$\mathbf{P}(M^{(a)} > x) \le e^{-h_0 x}. \tag{1.33}$$

Because of (1.27) the Lundberg inequality has optimal order and the error is only a constant. The proof of the Lundberg inequality is based on the observation that $\mathbf{E}[e^{h_0 X^{(a)}}] = 1$ implies that the sequence $e^{h_0 S_n^{(a)}}$ is a martingale and therefore $\mathbf{E}[e^{h_0 S_n^{(a)}}] = 1$ for all $n \in \mathbb{N}_0$. Applying Doob's martingale inequality one obtains

$$\mathbf{P}(M_n^{(a)} > x) = \mathbf{P}\left(\sup_{0\le k\le n} e^{h_0 S_k^{(a)}} > e^{h_0 x}\right) \le e^{-h_0 x}\mathbf{E}[e^{h_0 S_n^{(a)}}] = e^{-h_0 x} \tag{1.34}$$

for all $n \in \mathbb{N}_0$. For $A_n = \{M_n^{(a)} > x\}$ and $A = \{M^{(a)} > x\}$ one has $A_n \uparrow A$ as $n \to \infty$ and hence by the σ-continuity of $\mathbf{P}$,

$$\mathbf{P}(M^{(a)} > x) = \lim_{n\to\infty} \mathbf{P}(M_n^{(a)} > x) \le e^{-h_0 x}.$$

The same martingale property allows one to make an exponential change of measure, which is used in the proof of (1.28).

2 Heavy traffic and heavy tails for subexponential distributions

The ideas in this chapter were developed in cooperation with Dr. Denis Denisov during a visit at the University of Manchester and a visit of Dr. Denisov at the University of Munich. The results stated in this chapter mainly correspond to publication [19], only Corollary 2.9 is not contained in this general form in [19]. In the publication it is only stated for the Weibull case and the proof is omitted since it uses similar methods as the proof of Proposition 2.4.

2.1 Introduction and statement of the results

We use the notation and setting introduced in Chapter 1.7. For reasons of simplicity we also assume that $X^{(a)} = X^{(0)} - a$ in this chapter. We will give a short discussion on how this condition can be weakened in Remark 2.10 after presenting the main results. In Chapter 1.7 we mentioned several important results on the asymptotical behaviour of the maximum of a random walk: If a is fixed, the Cramér condition is fulfilled for some $h_0 > 0$ and $\mathbf{E}[X^{(a)} e^{h_0 X^{(a)}}] < \infty$, then there exists some constant $c_0 \in (0, 1)$ such that

$$\mathbf{P}(M^{(a)} > x) \sim c_0 e^{-h_0 x} \quad \text{as } x \to \infty. \tag{2.1}$$

Furthermore, if a is fixed and the integrated right tail $\overline{F}^I(x)$ is subexponential, then

$$\mathbf{P}(M^{(a)} > x) \sim \frac{1}{a} \overline{F}^I(x) \quad \text{as } x \to \infty. \tag{2.2}$$

If, on the other hand, $|X^{(a)}|$ has finite variance σ^2 and x is fixed,

$$\mathbf{P}(M^{(a)} > x/a) \sim e^{-2x/\sigma^2} \quad \text{as } a \to \infty. \tag{2.3}$$

Furthermore, if x is in general not fixed, but $x = O(1)$ or $x \to \infty$ sufficiently slow as $a \to 0$, then

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} \quad \text{as } a \to 0, \tag{2.4}$$

where θ_a is the solution to the equation

$$\mathbf{E}\left[e^{\theta_a X^{(a)}}; X^{(a)} \le 1/a\right] = 1. \tag{2.5}$$

It is known that θ_a allows a Taylor expansion of the form $\theta_a = 2a/\sigma^2 + C_2 a^2 + \ldots + C_k a^k$ with $C_i, i \in 2, \ldots, k$ are suitable constants. This expansion is valid up to the order of the moment existence of $X^{(a)}$ and the constants C_i can be defined using these moments.

One can see that (2.1) has a form similar to (2.4). Indeed, if the Cramér condition holds then letting $a \to 0$ one can see that $h_0 \to 0$ and in the limit (2.1) becomes (2.4), see for example Asmussen [2]. However it is not immediately clear what happens if one lets $a \to 0$ and $x \to \infty$ simultaneously when the Cramér condition does not hold and the distribution of $X^{(a)}$ is subexponential. The problem is the following: For $x \gg 1/a$ the heavy traffic theory predicts an exponential decay, whereas the heavy tail asymptotics predicts a decay according to the integrated tail of the distribution. This fact raises an interesting mathematical issue, how exponential asymptotics turn into the integrated tail asymptotics in the subexponential case. On one hand, if $a \to 0$ much faster than $x \to \infty$, the probability $\mathbf{P}(M^{(a)} > x/a)$ still behaves like in the heavy traffic approximation (2.3) by virtue of (2.4). On the other hand, if $a \to 0$ much slower than $x \to \infty$, the heavy tail approximation from (2.2) should still be valid. In particular, the question is whether there exists a transition point, at which the transition from (1.30) to (1.29) takes place. Or, otherwise, whether there is a third region in which neither the heavy traffic nor the heavy tail asymptotics holds and what the asymptotical behaviour of $\mathbf{P}(M^{(a)} > x)$ will be like in this region. This chapter deals with these questions and as it turns out answers depend on the distribution of the increments of the random walk.

We will consider different distributions and examine whether there exists a sharp transition point for theses distributions. Namely, we will show that if the increments are regular varying of index $r > 2$, then there exists a sharp transition point

$$x_{RV}(a) \approx \frac{\sigma^2(r-2)}{2}\frac{1}{a}\ln\frac{1}{a}. \tag{2.6}$$

and this is a value for which the terms on the right hand side of (2.2) and (2.4) have equal order. This generalizes a result from Olvera-Cravioto, Blanchet and Glynn [43], who derived this critical value in the setting of a $M/G/1$ queue. In Chapter 3 we will even show that if only $\mathbf{E}[(\min\{X^{(a)}, 0\})^2] < \infty$, but the variance is in general not finite, i.e $r \in (1, 2)$, then the heavy tail approximation (1.29) holds above the boundary value $a^{1/(1-r)}$. In the case of Weibull-like tails, that is $\overline{F}(x) = e^{-x^\gamma}$, $\gamma \in (0, 1)$, one could still believe there is a sharp transition point and try to find it by equating (2.2) and (2.4). Then, the critical point would be

$$x_W(a) \approx \left(\frac{1}{\theta_a}\right)^{1/(1-\gamma)} - \frac{2}{\theta_a(1-\gamma)}\ln\frac{\sqrt{2/(\gamma\sigma^2)}}{\theta_a}. \tag{2.7}$$

There are some recent results that coped with the case of Weibull-type distributions, but only in the case of a $M/G/1$ queue. In [44], Olvera-Cravioto and Glynn examine the case of a $M/G/1$ queue and conjecture that for Weibull-type distributions there is a third region in which neither the heavy traffic nor the integrated tail asymptotic is valid if and only if $1/2 < \gamma < 1$. However, we show that this is not the case and surprisingly this third region exists for a larger amount of γ, that is for all $\gamma \in (0, 1)$, see examples below. There is also a remarkable recent result by Blanchet and Lam, which covers various subexponential distributions. To be more specific they consider distributions very similar to the ones introduced in Chapter 1.5 and derive a uniform, explicit representation for the probability $\mathbf{P}(M^{(a)} > x)$, which consists of the exponential term from heavy traffic asymptotics (2.4), the integrated tail term (2.2) and a convolution term. For further discussion of this result see Remark 2.8. The reason why all these results only work in the setting of a $M/G/1$ queue is that their approach is based on the representation of

$M^{(a)}$ as a geometric sum of independent random variables:

$$\mathbf{P}(M^{(a)} > x) = \sum_{k=0}^{\infty} q(1-q)^k \mathbf{P}(\chi_1^+ + \chi_2^+ \cdots + \chi_k^+ > x),$$

where $\{\chi_l^+\}$ are independent random variables and $q = \mathbf{P}(M^{(a)} = 0)$. The main difficulty in this approach is the fact that one has to know the distribution of χ_l^+ and the parameter q. However, q and $\mathbf{P}(\chi_1^+ > x)$ are only known in some special cases. For example, if the left tail of $X^{(a)}$ decays according to an exponential distribution, that is $\mathbf{P}(X^{(a)} < -x) = be^{-\beta x}$ for some $b \in (0,1]$ and $\beta > 0$, the undershoot under 0 is also exponentially distributed and hence one can verify

$$\mathbf{E}[S_{\tau_-^{(a)}}^{(a)}] = -1/\beta, \quad \text{where } \tau_-^{(a)} = \min\{k \geq 1 : S_k^{(a)} < 0\}.$$

Consequently, the known formula $\mathbf{P}(M^{(a)} = 0) = 1/\mathbf{E}[\tau_-^{(a)}]$ and Wald's identity give

$$\mathbf{P}(M^{(a)} = 0) = \frac{1}{\mathbf{E}[\tau_-^{(a)}]} = \frac{a}{-\mathbf{E}[S_{\tau_-^{(a)}}^{(a)}]} = \beta a.$$

This means the value q is known in this case. However, the distribution of the overshoot remains unknown and one has to obtain appropriate estimates for $\mathbf{P}(\chi_1^+ > x)$. This case corresponds to the case of a $M/G/1$ queue. If the right tail of $X^{(a)}$ also decays exponentially (which corresponds to a $M/M/1$ queue) both values q and $\mathbf{P}(\chi_l^+ > x)$ are known and no estimates are required. However, in the general case (which corresponds to the case of a $G/G/1$ queue) the value q and the distribution of χ_1^+ remain unknown and, using the approach via geometric sums, one has to find good estimates for both of them. Therefore, an approach via geometric sums may be unsuitable for general distributions. In the present work we use a different approach that relies on martingale methods. Appearance of martingales is due to the equation $(M^{(a)} + X^{(a)})^+ \stackrel{d}{=} M^{(a)}$.

Before we state our main result we introduce assumptions on the distribution of $X^{(a)}$. If one writes

$$\overline{F}(x) = e^{-g(x)}, \tag{2.8}$$

the function $g(x) = \ln(-\overline{F}(x))$ is usually called hazard function of the distribution F.

Definition 2.1. Let $\gamma \in [0,1)$. The distribution F belongs to the class K_γ, $\gamma \in (0,1)$, if g is positive, twice differentiable, concave and if for every $\varepsilon_1 > 0$ there exists $x_0 = x_0(\varepsilon_1)$ such that

$$\frac{g(x)}{x^{\gamma+\varepsilon_1}} \searrow, \quad x \geq x_0, \tag{2.9}$$

and

$$\frac{g(x)}{x^{\gamma-\varepsilon_1}} \nearrow, \quad x \geq x_0. \tag{2.10}$$

The distribution F belongs to the class K_0, if g is positive, twice differentiable, concave, if (2.9) holds for $\gamma = 0$ and if there exists some value x_1 such that

$$\frac{g(x)}{\ln x} \nearrow, \quad x \geq x_1. \tag{2.11}$$

By definition, for a distribution being in the class K_γ is a tail property and, as shown in Chapter 1.5, all distributions $F \in K_\gamma$ with $\gamma \in [0, 1)$ are strong subexponential. The class K_γ contains the most popular strong subexponential distribution functions. For example, distributions with regular varying and lognormal-type tails are in K_0 and distributions with Weibull-type tails or semi-exponential tails are in K_γ, where the value γ corresponds to the parameter γ in (1.23) and (1.24). An example of a subexponential distribution that is not contained in any class K_γ, $\gamma \in [0, 1)$, is a distribution on $\mathbb{R}^+$ with $g(x) = x/\ln^\beta x$, $\beta > 0$.

Whenever we write $F \in K_0$, we will also assume $g'(x)x \ln x/g(x) \nearrow$ for x large enough and if $g(x) = O(\ln(x))$ assume that $(-g''(x))/(g'(x))^2$ converges as $x \to \infty$ in this chapter. On the other hand, if $F \in K_\gamma$ with $\gamma \in (0, 1)$, assume that $xg'(x)/((\gamma - \varepsilon_1)g(x)) \nearrow$ for all $\varepsilon_1 > 0$ if x is large enough. These assumptions are no big restriction and, in particular, the conditions can be verified for all strong subexponential distributions introduced in Chapter 1.6. We motivate the assumptions in the following lemma.

Lemma 2.2. *Suppose $F \in K_\gamma$ for some $\gamma \in [0, 1)$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$. If $\gamma = 0$,*

$$g'(x) \geq \frac{g(x)}{x \ln x} \tag{2.12}$$

for x large enough. If $g'(x)x \ln x/g(x) \nearrow$ for x large enough, there exists some $b \in [0, 1/2)$ such that

$$\frac{-g''(x)}{(g'(x))^2} \to b \in [0, 1/2) \quad \text{as } x \to \infty. \tag{2.13}$$

On the other hand, if $\gamma > 0$,

$$\frac{xg'(x)}{g(x)} \in [\gamma - \varepsilon_1, \gamma + \varepsilon_1] \tag{2.14}$$

for all $\varepsilon_1 > 0$ if x is large enough. Furthermore, if $xg'(x)/((\gamma - \varepsilon_1)g(x)) \nearrow$ for all $\varepsilon_1 > 0$ if x is large enough,

$$\frac{g''(x)}{(g'(x))^2} = o(1) \quad \text{as } x \to \infty. \tag{2.15}$$

The only reason we need the additional assumptions is to compare $\overline{F}^I$ with $\overline{F}$. If one assumes that $\overline{F}^I(x) = e^{-g(x)}$ instead of $\overline{F}(x) = e^{-g(x)}$, one can omit these additional assumptions. In the following Lemma we introduce a boundary sequence $x(a)$ that helps to define transition zones from the heavy traffic and the heavy tail asymptotics. If there exists a sharp transtion point, it will turn out that this value will coincides with the transition value.

Lemma 2.3. *Suppose $F \in K_\gamma$, $\gamma \in [0, 1)$ and that $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$. Then, there exists an increasing solution $x(a) \geq 1/a$ to the equation*

$$\theta_a x - g(x) - \ln(a\theta_a) = o(1) \quad \text{as } a \to 0. \tag{2.16}$$

This solution is asymptotically unique in the sense that any other solution $\widetilde{x}(a)$ satisfies $\widetilde{x}(a) = x(a) + o(1/a)$ as $a \to 0$.

For $c > 0$ define,

$$G_c(x) := \begin{cases} 1, & \text{if } x \le 0, \\ \exp\{-(\theta_a + c)x\}, & \text{if } x > 0 \end{cases} \tag{2.17}$$

and for $\delta \in (0,1)$ and $c > 0$ let

$$\widehat{G}_c(x) := \begin{cases} \frac{1}{ac}\overline{F}^I(x), & \text{if } x > \delta x(a), \\ 0, & \text{if } x \le \delta x(a). \end{cases} \tag{2.18}$$

Note that $\widehat{G}_c(x) = 0$ rather than 1 as in [20]. This is due to the fact that we are going to consider the sum of two functions and 1 will come from another term as e^{-2az/σ^2} is approximately 1 for small a and z in a fixed interval. Furthermore, define for $\alpha, c > 0$

$$\widetilde{G}_c(x) := \begin{cases} e^{\alpha}, & \text{if } x \le 0, \\ \exp\{-(\theta_a + c)x\}, & \text{if } x > 0, \end{cases} \tag{2.19}$$

and put $\mu_y := \inf\{k \ge 1 : S_k^{(a)} \ge y\}$.

Proposition 2.4. *Suppose $F \in K_0$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$. Put $c_a := 1/(x(a)\ln(1/a))$ and*

$$Y_n^{(1)} := G_{-c_a}(x - S^{(a)}_{n\wedge\mu_x}) + \widehat{G}_{1-\varepsilon}(x - S^{(a)}_{n\wedge\mu_{x-\delta x(a)}}). \tag{2.20}$$

Then, $(Y_n^{(1)})$ is a non-negative supermartingale for all $a > 0$ and $\delta \in (0,1)$ small enough. Define

$$Y_n^{(2)} := \widetilde{G}_{c_a}(x - S^{(a)}_{n\wedge\mu_x}) + \widehat{G}_{1+\varepsilon}(x - S^{(a)}_{n\wedge\mu_{x-2\delta x(a)}}). \tag{2.21}$$

Then, $(Y_n^{(2)})$ is a non-negative submartingale for all $a > 0$ and $\delta \in (0,1)$ small enough.

Furthermore, for $\delta \in (0,1)$ and $N > 1$, define

$$L^{(N)}(x) := \left(\sum_{k=0}^{\infty}\left(\frac{\gamma g(x)}{\theta_a x}\right)^k\right)^2 \wedge N \tag{2.22}$$

and

$$\widehat{G}'_c(x) := \begin{cases} \frac{L^{(N)}(x)}{ac}\overline{F}^I(x), & \text{if } x > \delta x(a), \\ 0, & \text{if } x \le \delta x(a). \end{cases} \tag{2.23}$$

Proposition 2.5. *Assume $F \in K_\gamma$ for some $\gamma \in (0,1)$ and $\mathbf{E}[|\min\{0, X^{(a)}\}|^{1+1/(1-\widetilde{\gamma})}] < \infty$ for some $\widetilde{\gamma} > \gamma$. Put $c_a := 1/(x(a)\ln(1/a))$ and*

$$\widetilde{Y}_n^{(1)} := G_{-c_a}(x - S^{(a)}_{n\wedge\mu_x}) + \widehat{G}'_{1-\varepsilon}(x - S^{(a)}_{n\wedge\mu_{x-\delta x(a)}}). \tag{2.24}$$

Then, $(\widetilde{Y}_n^{(1)})$ is a non-negative supermartingale for all $a > 0$ and $\delta \in (0,1)$ small enough and N large enough. Define

$$\widetilde{Y}_n^{(2)} := \widetilde{G}_{c_a}(x - S^{(a)}_{n\wedge\mu_x}) + \widehat{G}'_{1+\varepsilon}(x - S^{(a)}_{n\wedge\mu_{x-2\delta x(a)}}). \tag{2.25}$$

Then, $(\widetilde{Y}_n^{(2)})$ *is a non-negative submartingale for all* $a > 0$ *and* $\delta \in (0,1)$ *small enough and* N *large enough.*

Lemma 2.6. *Suppose* $F \in K_\gamma$ *for some* $\gamma \in (0,1)$ *and* $\mathbf{E}[|\min\{0, X^{(a)}\}|^{1+1/(1-\widetilde{\gamma})}] < \infty$ *for some* $\widetilde{\gamma} > \gamma$*. Then,*

$$\limsup_{a\to 0} \frac{g(x)}{\theta_a x} \le 1 \tag{2.26}$$

for all $x \ge \delta x(a)$ *such that* $e^{-\theta_a x} = O(\overline{F}^I(x)/a)$ *as* $a \to 0$*.*

With the super- and submartingales from Propositions 2.4 and 2.5 we can derive subexponential asymptotics for the probability $\mathbf{P}(M^{(a)} > x)$.

Theorem 2.7. *Suppose* $F \in K_0$ *and* $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$*. Then, uniformly in* x*,*

$$\begin{aligned}\mathbf{P}(M^{(a)} > x) &\sim e^{-\theta_a x} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \ge \delta x(a)\} \\ &\sim e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \ge \delta x(a)\} \quad \text{as } a \to 0.\end{aligned} \tag{2.27}$$

On the other hand, if $F \in K_\gamma$*,* $\gamma \in (0,1)$*, and* $\mathbf{E}[|\min\{0, X^{(a)}\}|^{1+1/(1-\widetilde{\gamma})}] < \infty$ *for some* $\widetilde{\gamma} > \gamma$ *one has, uniformly in* x*,*

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} + \frac{L^{(N)}(x)}{a}\overline{F}^I(x)\mathbf{1}\{x \ge \delta x(a)\} \tag{2.28}$$

$$\sim e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x > x(a) - C\ln(1/a)/\theta_a\}}{a(1-\gamma g(x)/(\theta_a x))^2} \quad \text{as } a \to 0 \tag{2.29}$$

with sufficiently large constants $C > 0$ *and* $N \ge 1/(1-\widetilde{\gamma})$*.*

Remark that due to Lemma 2.6 the right hand side in (2.29) is well defined. As shown in the proof of Theorem 2.7, the constant C is such that, for $\delta x(a) \le x < x(a) - C\ln(1/a)/\theta_a$, one has $e^{-\theta_a x} \gg \overline{F}^I(x)/a$. Theorem 2.7 implies that the intermediate region, in which neither the heavy traffic nor the heavy tail asymptotics is valid, appears if and only if $F \in K_\gamma$, $\gamma \in (0,1)$. In this case the right tail of the increments decreases at least as fast as it does for the Weibull distribution, that is

$$\overline{F}(x) \ll e^{-x^{\varepsilon_1}} \quad \text{as } x \to \infty$$

for some $\varepsilon_1 > 0$. Consequently, there is no intermediate region for regular varying and Lognormal-type tail distributions and there is one for Weibull-like and semi-exponential distributions. However, for those x such that

$$\frac{1}{a}\overline{F}^I(x) \asymp e^{-\theta_a x} \quad \text{as } a \to 0$$

there appears a mixing of the two terms in the regular varying and the lognormal case.

Remark 2.8. In the case of a $M/G/1$ queue, which corresponds to the case that the left tail of the increments are exponential, Blanchet and Lam [7] have shown that, as $a \to 0$,

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} + \left(\frac{\overline{F}^I(x)}{a} - \int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y) \right) e^{-\theta_a (x-y)} dF^I(y) \right) \mathbf{1}\left\{ x \geq \frac{1}{a} \right\}. \tag{2.30}$$

One can show with similar techniques as used in the proof of Proposition 2.4 and 2.5, that under our assumptions on the distribution of the increments the integral term from (2.30) is approximately

$$\left(\frac{1}{a(1-\gamma g(x)/(\theta_a x))^2))^2} - 1 \right) \overline{F}^I(x). \tag{2.31}$$

In particular, one can show the following corollary.

Corollary 2.9. *Suppose that $F \in K_\gamma$ for some $\gamma \in [0,1)$. Then, the integral term in the result* (2.30) *is asymptotically negligible in the case $\gamma = 0$ and is asymptotical equivalent to the term from* (2.31) *in the case $\gamma \in (0,1)$. That means the right hand side of* (2.30) *is the same as the right hand side of* (2.27) *in the case $\gamma = 0$ and* (2.29) *in the case $\gamma \in (0,1)$.*

For the proof of Proposition 2.4, Proposition 2.5 and Theorem 2.7 it turns out that it is not necessary to have an exact solution in (2.5), but it is sufficient to have an equation of the form

$$\mathbf{E}\left[e^{\theta_a X^{(a)}}; X^{(a)} \leq 1/a \right] = 1 + o(ac_a), \tag{2.32}$$

where $c_a = o(a)$ reflects the required precision and depends on the distribution of the right tail. A suitable choice will be $c_a = 1/(x(a)\ln(1/a))$. Then, the definition of $x(a)$ implies that $c_a \gg a^k$ for k large enough, that means $k^* := \max\{k \in \{1,2,\dots\} : c_a \ll a^k\}$ is well defined. Then, by considering an asymptotic equation of the form (2.32), one can pick a solution θ_a of the form

$$\theta_a = \frac{2a}{\sigma^2} + C_2 a^2 + \cdots + C_{k^*} a^{k^*}. \tag{2.33}$$

The expansion is valid up to the order of the moment existence of $X^{(a)}$ and the constants $C_2, C_3, \dots, C_{k^*}$ can be defined by expansion and using these moments. For Weibulls with parameter strictly less than 1/2, one has $\theta_a x(a) = 2ax(a)/\sigma^2 + o(1)$, see example 4 in Chapter 1.6, so the result from (2.29) simplifies to

$$\mathbf{P}(M^{(a)} > x) \sim e^{-2ax/\sigma^2} + \frac{\overline{F}^I(x)}{a(1-\gamma g(x)/(\theta_a x))^2} \quad \text{as } a \to 0.$$

For even lighter tails one needs more moments to expand θ_a like in (2.33).

Remark 2.10. The assumption $X^{(a)} = X - a$ is needed only for the expansion of θ_a and can be generalized to the assumption that the moment equivalence $\lim_{a\to 0} \mathbf{E}[(X^{(a)})^k] = \mathbf{E}[(X^{(0)})^k]$ is valid for all $2 \le k \le 1 + 1/(1-\widetilde{\gamma})$. In particular, if the tails decay slower than $e^{-\sqrt{x}}$ we only need to assume $\lim_{a\to 0} \mathbf{E}[(X^{(a)})^2] = \sigma^2$ and $\sup_{a\le a_0} \mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ for some $\varepsilon, a_0 > 0$. For increments of type $X^{(a)} = X^{(0)} - a$ we always have an expansion in powers, so the coefficient will change but still depend only on first cumulants.

2.2 Examples

Let us consider different kinds of distribution functions and outline the result for these distributions given by Theorem 2.7. Especially, we shall see that, depending on the distribution, there may be a mixing area around the transition zone, in which the order of the exponential term and the order of the integrated tail term are the same.

Example 1: distributions with regular varying tails. Suppose the right tail of $X^{(a)}$ is regular varying with index $r > 2$, that is $g(x) = r\ln x - \ln L_1(x)$, where L_1 is some slowly varying function. Then, equating the two terms on the right hand side of (2.27) one can see that the transition point coincides with the critical value from Lemma 2.3. In particular, this value is

$$x_{RV}(a) \approx \frac{(r-2)\sigma^2}{2a} \ln\frac{1}{a}.$$

On the other side, it is known that $\overline{F}^I(x) \sim x^{1-r}L_1(x)/(r-1)$ for all x such that $x \to \infty$ as $a \to 0$ and obviously $\overline{F}^I(x)/a \gg e^{-2ax/\sigma^2}$ for all $x \gg x_{RV}(a)$. Hence, Theorem 2.7 can be rewritten as

$$\mathbf{P}(M^{(a)} > x) \sim e^{-2ax/\sigma^2} + \frac{x^{1-r}L_1(x)}{(r-1)a}\mathbf{1}\{x \ge \delta x(a)\}. \tag{2.34}$$

One can see that $e^{-2ax/\sigma^2} \gg \overline{F}^I(x)/a$ if $\delta x(a) \le x < cx_{RV}(a)(1+o(1))$ with $c < 1$ and that $e^{-2ax/\sigma^2} \ll \overline{F}^I(x)/a$ if $x > cx_{RV}(a)(1+o(1))$ with $c > 1$. In the case $c = 1$, it depends on the exact dependence of x and a and the order of the slowly varying function L_1, whether the exponential term or the integrated tail term dominates or if they even have the same order. Hence, Theorem 2.7 states that

$$\mathbf{P}(M^{(a)} > x) \sim \begin{cases} e^{-2ax/\sigma^2}, & \text{if } \lim_{a\to 0} x/x_{RV}(a) < 1, \\ \frac{1}{a}\overline{F}^I(x), & \text{if } \lim_{a\to 0} x/x_{RV}(a) > 1, \\ e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x), & \text{if } \lim_{a\to 0} x/x_{RV}(a) = 1. \end{cases} \tag{2.35}$$

Let us discuss the region $\lim_{a\to 0} x/x_{RV}(a) = 1$ a little bit more. It is easy to see that $\overline{F}^I(x) \sim \overline{F}^I(x_{RV}(a))$ in this region and for $x = x_{RV}(a) + O(1/a)$, $e^{-2ax/\sigma^2} \asymp e^{-2ax_{RV}(a)/\sigma^2}$. On the other side, $\overline{F}^I(x_{RV}(a))/a \gg e^{-2ax_{RV}(a)/\sigma^2}$ for $L_1(x) \gg (\ln(x))^{r-1}$ and $\overline{F}^I(x_{RV}(a))/a \ll e^{-2ax_{RV}(a)/\sigma^2}$ for $L_1(x) \ll (\ln x)^{r-1}$. Hence, if $x = x_{RV}(a) + O(1/a)$, follows $\overline{F}^I(x)/a \gg e^{-2ax/\sigma^2}$ for $L_1(x) \gg (\ln x)^{r-1}$ and $\overline{F}^I(x)/a \ll e^{-2ax/\sigma^2}$ for $L_1(x) \ll (\ln x)^{r-1}$.

Example 2: distributions with Pareto-like tails. Suppose the right tail of $X^{(a)}$ is Pareto distributed with index $r > 2$, that is $g(x) = r \ln x$. This is the same as in Example 1, but with $L_1(x) \equiv 1$. One can refine the region $x \sim x_{RV}(a)$ from Example 1 and show that the critical value is

$$x_P(a) \approx \frac{(r-2)\sigma^2}{2a} \ln(1/a) + \frac{(r-1)\sigma^2}{2a} \ln(\ln(1/a)) =: x_{P1}(a) + x_{P2}(a).$$

Hence, Theorem 2.7 states that

$$\mathbf{P}(M^{(a)} > x) \sim \begin{cases} e^{-2ax/\sigma^2}, & \text{if } \lim_{a\to 0} \frac{x - x_{P1}(a)}{x_{P2}(a)} < 1, \\ \frac{1}{a}\overline{F}^I(x), & \text{if } \lim_{a\to 0} \frac{x - x_{P1}(a)}{x_{P2}(a)} > 1, \\ e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x), & \text{if } \lim_{a\to 0} \frac{x - x_{P1}(a)}{x_{P2}(a)} = 1. \end{cases} \tag{2.36}$$

In the last case of the latter result there may occur some mixing between the two terms, for example if $x = \frac{(r-2)\sigma^2}{2a} \ln(1/a) + \frac{(r-1)\sigma^2}{2a} \ln(\ln(1/a)) + o(1/a)$, then

$$\mathbf{P}(M^{(a)} > x) \sim \left(1 + (r-1)\left(\frac{(r-2)\sigma^2}{2}\right)^{r-1}\right) \frac{1}{a}\overline{F}^I(x)$$

Example 3: distributions with lognormal-type tails. Let $g(x) = r \ln^\beta x$ with $\beta > 1$ and $r > 0$ such that $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$. Then, one can verify

$$\overline{F}^I(x) \sim \frac{x}{\beta(\ln x)^{\beta-1}} e^{-r \ln^\beta x}$$

and by equating the integrated tail term and the exponential term one can calculate the critical value

$$x_{LN}(a) \approx \frac{r\sigma^2}{2a} \ln^\beta(1/\theta_a).$$

Hence, Theorem 2.7 states that

$$\mathbf{P}(M^{(a)} > x) \sim \begin{cases} e^{-2ax/\sigma^2}, & \text{if } \lim_{a\to 0} x/x_{LN}(a) < 1, \\ \frac{1}{a}\overline{F}^I(x), & \text{if } \lim_{a\to 0} x/x_{LN}(a) > 1, \\ e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x), & \text{if } \lim_{a\to 0} x/x_{LN}(a) = 1. \end{cases} \tag{2.37}$$

In the region $x \sim x_{LN}(a)$, one can see that for $x = r \ln^\beta(1/\theta_a)/\theta_a$, $\overline{F}^I(x)/a \gg e^{-\theta_a x}$ if $\beta \in (1, 2)$, and $\overline{F}^I(x)/a \ll e^{-\theta_a x}$ if $\beta \geq 2$.

Example 4: distributions with Weibull-like tails. Suppose the right tail of $X^{(a)}$ possesses a Weibull distribution, that is $g(x) = x^\gamma$ with $\gamma \in (0, 1)$. Then, one can easily see for example by substitution and using asymptotical properties of the incomplete gamma function, that

$$\overline{F}^I(x) \sim \frac{1}{\gamma} x^{1-\gamma} \overline{F}(x).$$

By equating the exponential and the integrated tail term, a critical value is

$$x_W(a) = \left(\frac{1}{\theta_a}\right)^{1/(1-\gamma)} - \frac{2}{\theta_a(1-\gamma)} \ln \frac{\sqrt{2/(\gamma\sigma^2)}}{\theta_a}. \tag{2.38}$$

By the definition of $x_W(a)$,

$$e^{-\theta_a x_W(a)} \sim \frac{1}{a}\overline{F}^I(x_W(a)) \sim \frac{1}{a} x_W(a)^{1-\gamma} e^{-x_W(a)^\gamma} \sim \frac{2}{\sigma^2} e^{-x_W(a)^\gamma} \tag{2.39}$$

and for $z \ll x_W(a)^{1-\gamma/2}$ one has

$$e^{-(x_W(a)-z)^\gamma} \sim e^{-x_W(a)^\gamma + \gamma z / x_W(a)^{1-\gamma}}. \tag{2.40}$$

With these results, one can easily see that

$$e^{-\theta_a x} \gg \frac{1}{a}\overline{F}^I(x) \tag{2.41}$$

for all $x \leq x_W(a) - z$ with $1/\theta_a \ll z \ll x_W(a)$ and

$$e^{-\theta_a x} \ll \frac{1}{a}\overline{F}^I(x) \tag{2.42}$$

for all $x \geq x_W(a) + z$ with $1/\theta_a \ll z \ll x_W(a)$. The relations (2.39) and (2.40) imply that for $x = x_W(a) + K/\theta_a + o(1/a)$, where K is a fixed constant,

$$e^{-\theta_a x} \sim \frac{e^{-K}}{a}\overline{F}^I(x_W(a)) \sim \frac{e^{-K(1-\gamma)}}{a}\overline{F}^I(x). \tag{2.43}$$

Analogously, for $x = x_W(a) - K/\theta_a + o(1/a)$,

$$\frac{1}{a}\overline{F}^I(x) \sim e^{-K(1-\gamma)} e^{-\theta_a x}. \tag{2.44}$$

Furthermore, $g(x)/(\theta_a x) \sim 1$ for $x \gg x_W(a)$ and $g(x)/(\theta_a x) \sim 1/(1-\gamma)$ for $x \sim x_W(a)$. Combining all the results from above, we arrive at

$$\mathbf{P}(M^{(a)} > x) \sim \begin{cases} e^{-\theta_a x}, & \text{if } x \prec x_W(a) - 1/\theta_a, \\ \left(1 + \frac{e^{-K(1-\gamma)}}{(1-\gamma)^2}\right) e^{-\theta_a x}, & \text{if } x = x_W(a) - K/\theta_a + o(1/a), K > 0, \\ \left(e^{-K(1-\gamma)} + \frac{1}{(1-\gamma)^2}\right) \frac{\overline{F}^I(x)}{a}, & \text{if } x = x_W(a) + K/\theta_a + o(1/a), K > 0, \\ \frac{\overline{F}^I(x)}{(a(1-g(x)/(\theta_a x))^2)}, & \text{if } x \succ x_W(a) + 1/\theta_a, x \not\gg x_W(a), \\ \frac{1}{a}\overline{F}^I(x), & \text{if } x \gg x_W(a). \end{cases} \tag{2.45}$$

Example 5: Semi-exponential distributions. Semiexponential distributions are distributions for which the right tail $\overline{F}(x)$ has the form

$$\overline{F}(x) = e^{-x^\gamma L_1(x)}, \tag{2.46}$$

where $\gamma \in (0,1)$ and L_1 is a slowly varying function. Assume that L_1 is differentiable. The rule of l'Hôspital gives

$$\overline{F}^I(x) \sim \frac{\overline{F}(x)}{g'(x)} \sim \frac{\overline{F}(x)}{\gamma x^{\gamma-1}L_1(x) + x^\gamma L_1'(x)}$$

in this case. From this, one can see that for all L_1 such that $L_1(x) \to \infty$ as $x \to \infty$,

$$e^{-\theta_a x} \gg \frac{1}{a}\overline{F}^I(x)$$

for x such that $x = O((1/\theta_a)^{1/(1-\gamma)})$ and

$$e^{-\theta_a x} \ll \frac{1}{a}\overline{F}^I(x)$$

for $x \gg (1/\theta_a)^C$ with $C > 1/(1-\gamma)$. The latter can be verified by use of the so called Potter bounds (see e.g. Theorem 1.5.6 in [5]), which state the regular varying functions grow slower than any power. Hence, the critical value is $(1/\theta_a)^{1/(1-\gamma)}$ and Theorem 2.7 states that

$$\mathbf{P}(M^{(a)} > x) \sim \begin{cases} e^{-\theta_a x}, & \text{if } x = O((1/\theta_a)^{1/(1-\gamma)}), \\ \frac{1}{a}\overline{F}^I(x), & \text{if } x \geq (1/\theta_a)^{\delta+1/(1-\gamma)}, \\ e^{-\theta_a x} + \frac{\overline{F}^I(x)}{a(1-g(x)/(\theta_a x))^2}, & \text{otherwise,} \end{cases} \tag{2.47}$$

for an arbitrary $\delta > 0$.

With the latter examples we perform comparison between the distributions for which a transition zone exists with the ones from large deviations. For a good survey of the results in large deviations, see for example [39]. The probability $\mathbf{P}(S_n > x)$ behaves as the tail of the normal distribution $\overline{\Phi}(x/\sqrt{n})$ below some threshold series (c_n) and as $n\overline{F}(x)$ above another threshold series (d_n).

Remark 2.11. Just like in large deviation theory, there is no transition zone for regular varying distribution functions and for lognormal-type distribution functions with index $\beta \leq 2$. However, unlike in large deviations, there is still no transition zone for lognormal-type distribution functions with index $\beta > 2$. If the tail possesses a Weibull-type distribution with index $\gamma \in (0,1)$ there is a transition zone just like in large deviations, but the value $\gamma = 1/2$ is not a threshold in the result (2.29) as it is in large deviations.

2.3 Proofs

2.3.1 Proof of Lemma 2.2

Let us first consider the case $\gamma = 0$. The assumption (2.11) implies

$$\left(\frac{g(x)}{\ln x}\right)' \geq 0$$

and an easy calculation shows that this is equivalent to

$$g'(x) \geq \frac{g(x)}{x \ln x}.$$

Furthermore, by virtue of $g'(x)x\ln x/g(x) \nearrow$,

$$\left(\frac{g'(x)x\ln x}{g(x)}\right)' = \frac{g''(x)x\ln x}{g(x)} + \frac{g'(x)(1+\ln x)}{g(x)} - \frac{(g'(x))^2 x\ln x}{(g(x))^2} \geq 0.$$

The concavity implies $g''(x) \leq 0$, therefore the latter inequality gives $g'(x)(1+\ln x) \geq -g''(x)x\ln x$ and consequently we obtain by regarding (2.12) that

$$\frac{-g''(x)}{(g'(x))^2} \leq \frac{1+\ln x}{xg'(x)\ln x} \leq \frac{1+\ln x}{g(x)}.$$

Due to (2.11) the term on the right hand side of the latter inequality is decreasing. Because of $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$, one has $g(x) \geq (2+\varepsilon)\ln x$ and therefore the term on the right hand side of the latter inequality is bounded by $1/2$. Due to (2.11) the term is also decreasing. This immediately implies (2.13) with $b = 0$ for all g such that $g(x) \gg \ln(x)$ as $x \to \infty$. In the case $g(x) = O(\ln(x))$ as $x \to \infty$, the relation (2.13) follows directly from our asssumptions.

Now, consider the case $\gamma > 0$. The condition (2.9) gives

$$\left(\frac{g(x)}{x^{\gamma+\varepsilon_1}}\right)' \leq 0$$

for all $\varepsilon_1 > 0$ if x is large enough and an easy calculation shows that this is equivalent to

$$g'(x) \leq (\gamma + \varepsilon_1)\frac{g(x)}{x}.$$

Using (2.10) instead of (2.9), one can show in the same way that, for all $\varepsilon_1 > 0$ such that $\gamma - \varepsilon_1 > 0$, one has

$$g'(x) \geq (\gamma - \varepsilon_1)\frac{g(x)}{x}$$

if x is large enough. The condition $xg'(x)/((\gamma-\varepsilon_1)g(x)) \nearrow$ implies that, for all $\varepsilon_1 > 0$ and x large enough,

$$\left(\frac{xg'(x)}{(\gamma-\varepsilon_1)g(x)}\right)' = \frac{xg''(x)}{(\gamma-\varepsilon_1)g(x)} + \frac{g'(x)}{(\gamma-\varepsilon_1)g(x)} - \frac{(g'(x))^2}{(\gamma-\varepsilon_1)(g(x))^2} \geq 0.$$

Thus, the concavity of g gives $g'(x) \geq -xg''(x)$ and consequently, by (2.14),

$$\frac{-g''(x)}{(g'(x))^2} \leq \frac{1}{xg'(x)} \leq \frac{1}{(\gamma-\varepsilon_1)g(x)} = o(1) \quad \text{as } x \to \infty.$$

2.3.2 Proof of Lemma 2.3

We will prove that there exists an $x(a)$ such that we have exact equality $\theta_a x - g(x) - \ln(a\theta_a) = 0$. The latter follows from the continuity of $g(x)$. Indeed, on one hand for $x = 1/\theta_a$, (2.33) gives

$$\begin{aligned}\theta_a x - g(x) - \ln(a\theta_a) &= 2\ln(1/a) - g(1/a) + O(1)\\ &\leq 2\ln(1/a) - (2+\varepsilon)\ln(1/a) + O(1) < 0\end{aligned}$$

for a small enough. Here we used that the existence of $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}]$ implies $g(x) \geq (2+\varepsilon)\ln x$ for x large enough. On the other hand, (2.9) implies that we have $g(x) = o(x^{\gamma+\varepsilon_1})$ for all ε_1 such that $0 < \gamma < \gamma + \varepsilon_1 < 1$. Hence, for $x = (1/\theta_a)^C$,

$$\theta_a x - g(x) - \ln(a\theta_a) \sim \theta_a^{1-C} - \theta_a^{-C(\gamma+\varepsilon_1)} > 0$$

provided $C > (1-\gamma_0-\varepsilon_1)^{-1}$.

Moreover, the function $x(a)$ is monotone increasing in a. Indeed, using $\theta_a = 2a/\sigma^2 + o(a)$, we attain

$$(\theta_a x - g(x) + \ln a)'_a = (2/\sigma^2 + o(1))x + 1/a \geq 0.$$

This means that as a decreases, $\theta_a x - g(x) + \ln a$ decreases. Since $x \gg g(x)$ as $x \to \infty$ this implies that the solution $x(a)$ to $\theta_a x - g(x) + \ln a = 0$ increases.

Suppose that $\widetilde{x}(a)$ is another solution to (2.16). One can easily see that, for $g(x) = (2+\varepsilon)\ln(x)$, (2.16) implies

$$\liminf_{a\to 0} \frac{\theta_a x(a)}{\ln(1/a)} \geq \varepsilon > 0 \tag{2.48}$$

and since $g(x) \geq (2+\varepsilon)\ln x$ the latter holds for all g such that $F \in K_\gamma$, $\gamma \in [0,1)$. By regarding (2.48) and the definition of $x(a)$ from (2.16),

$$\frac{g(x(a))}{\theta_a x(a)} = 1 + \frac{\ln(1/(a\theta_a))}{\theta_a x(a)} + o(1). \tag{2.49}$$

This means that there exists a constant C_1 such that $g(x(a))/(\theta_a x(a)) \leq C_1$. To show that $\widetilde{x}(a) = x(a) + o(1/a)$ let us first consider the case $\gamma = 0$. For $\gamma = 0$, inequality (1.15) states that, for all $\varepsilon_1 > 0$,

$$g(x) - g(x-w) \leq 2\varepsilon_1 w \frac{g(x)}{x}, \quad 0 \leq w \leq x/2. \tag{2.50}$$

Furthermore, $g(x) \ll x$ implies $x(a) \sim \widetilde{x}(a)$ and therefore $|x(a) - \widetilde{x}(a)| \ll x(a)$. Hence, (2.50) gives that for all $\varepsilon_1 > 0$,

$$\begin{aligned}|\widetilde{x}(a) - x(a)| &= \left|\frac{g(\widetilde{x}(a)) - g(x(a))}{\theta_a} + o(1/\theta_a)\right|\\ &\leq 2\varepsilon_1 |\widetilde{x}(a) - x(a)| \frac{g(x(a))}{\theta_a x(a)} + o(1/\theta_a) \leq 2\varepsilon_1 C_1 |\widetilde{x}(a) - x(a)| + o(1/\theta_a)\end{aligned}$$

and for $\varepsilon_1 < 1/(2C_1)$ this implies $\widetilde{x}(a) = x(a) + o(1/a)$.

Now suppose $\gamma \in (0,1)$. In this case, (2.10) ensures the existence of $\varepsilon_2 > 0$ such that $g(x) \gg x^{\varepsilon_2}$. By the definition of $x(a)$,

$$x(a) = \frac{g(x(a))}{\theta_a} - \frac{1}{\theta_a} \ln \frac{1}{a\theta_a} + o\left(\frac{1}{\theta_a}\right). \tag{2.51}$$

Therefore, there exists some $\varepsilon_2 > 0$ such that

$$x(a) \gg \left(\frac{1}{\theta_a}\right)^{1+\varepsilon_2} \tag{2.52}$$

and hence, by (2.49),

$$\frac{g(x(a))}{\theta_a x(a)} = 1 + o(1). \tag{2.53}$$

Combining this result with (1.17) and $|x(a) - \widetilde{x}(a)| \ll x(a)$, we see that one can choose $0 < \alpha < 1$ with

$$\begin{aligned} |\widetilde{x}(a) - x(a)| &= \left| \frac{g(\widetilde{x}(a)) - g(x(a))}{\theta_a} + o(1/\theta_a) \right| \le \alpha |\widetilde{x}(a) - x(a)| \frac{g(x(a))}{\theta_a x(a)} + o(1/\theta_a) \\ &\le \alpha |\widetilde{x}(a) - x(a)| + o(1/\theta_a) + o(|\widetilde{x}(a) - x(a)|). \end{aligned}$$

Consequently, we obtain $\widetilde{x}(a) = x(a) + o(1/a)$.

2.3.3 Proof of Proposition 2.4

During the whole proof we assume a to be sufficiently small, even if not explicitly mentioned. The supermartingale property for $(Y_n^{(1)})$ is equivalent to the following two inequalities:

$$\begin{aligned} &\mathbf{E}[G_{-c_a}(x - y - X^{(a)})] + \mathbf{E}[\widehat{G}_{1-\varepsilon}(x - y - X^{(a)})] \\ &\qquad \le G_{-c_a}(x-y) + \widehat{G}_{1-\varepsilon}(x-y), \quad y \le x - \delta x(a), \end{aligned} \tag{2.54}$$

and

$$\mathbf{E}[G_{-c_a}(x - y - X^{(a)})] \le G_{-c_a}(x-y), \quad y \in (x - \delta x(a), x]. \tag{2.55}$$

Put $t := x - y$ and remark that $x(a)$ does not depend on x, but only on a. Then, (2.54) is equivalent to

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}_{1-\varepsilon}(t - X^{(a)})] \le G_{-c_a}(t) + \widehat{G}_{1-\varepsilon}(t), \quad t \ge \delta t(a), \tag{2.56}$$

where we wrote $t(a)$ instead of $x(a)$ due to the change of variables. In addition, (2.55) is equivalent to

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] \le G_{-c_a}(t), \quad t \in [0, \delta t(a)). \tag{2.57}$$

Let us bound the expectation on the left side of the latter inequality. Put $\kappa_a = \theta_a - c_a$ for brevity, then

$$\begin{aligned}
&\mathbf{E}[G_{-c_a}(t - X^{(a)})] \\
&\le e^{-\kappa_a t}\mathbf{P}(X^{(a)} \le -1/\kappa_a) + e^{-\kappa_a t}\mathbf{E}\left[e^{\kappa_a X^{(a)}}; X^{(a)} \in (-1/\kappa_a, 1/\kappa_a]\right] \\
&\qquad\qquad + e^{-\kappa_a t}\mathbf{E}\left[e^{\kappa_a X^{(a)}}; X^{(a)} \in (1/\kappa_a, t]\right] + \overline{F}(t) \\
&=: G_1 + G_2 + G_3 + G_4. \qquad (2.58)
\end{aligned}$$

By virtue of (2.48), $x(a) \ge 1/a$ and therefore

$$c_a \le \frac{a}{\ln(1/a)} = o(a) \quad \text{as } a \to 0. \qquad (2.59)$$

Recalling that $\theta_a = 2a/\sigma^2 + o(a)$ and using the estimate $e^x \le 1 + x + x^2$, which is valid for $|x| \le 1$, one obtains

$$e^{\kappa_a X^{(a)}} = e^{\theta_a X^{(a)} - c_a X^{(a)}} \le e^{\theta_a X^{(a)}}\left(1 - c_a X^{(a)} + (c_a X^{(a)})^2\right).$$

for $X^{(a)} \in [-1/\kappa_a, 1/\kappa_a]$. Using $c_a = o(a)$ again, we get

$$\begin{aligned}
G_2 &\le e^{-\kappa_a t}\mathbf{E}\left[e^{\theta_a X^{(a)}}\left(1 - c_a X^{(a)} + (c_a X^{(a)})^2\right); X^{(a)} \in (-1/\kappa_a, 1/\kappa_a]\right] \\
&\le e^{-\kappa_a t}\mathbf{E}\left[e^{\theta_a X^{(a)}}\left(1 - c_a X^{(a)}\right); X^{(a)} \in (-1/\kappa_a, 1/\kappa_a]\right] + \sigma^2 e^2 c_a^2 e^{-\kappa_a t}.
\end{aligned}$$

Furthermore, since $\theta_a = 2a/\sigma^2 + o(a)$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$,

$$\begin{aligned}
&\mathbf{E}\left[X^{(a)} e^{\theta_a X^{(a)}}; X^{(a)} \in [-1/\kappa_a, 1/\kappa_a]\right] \\
&= \mathbf{E}\left[X^{(a)}; X^{(a)} \in [-1/\kappa_a, 1/\kappa_a]\right] + \theta_a\mathbf{E}\left[(X^{(a)})^2; X^{(a)} \in [-1/\kappa_a, 1/\kappa_a]\right] + o(a) \\
&= -a + \sigma^2\theta_a + o(a) = a + o(a).
\end{aligned}$$

One can easily see that (2.9) implies $g(t) \ll t^{\varepsilon_2}$ as $t \to \infty$ for all $\varepsilon_2 > 0$, and for this reason (2.51) gives

$$t(a) = x(a) \ll \left(\frac{1}{\theta_a}\right)^{1+\varepsilon_2} \qquad (2.60)$$

and hence

$$c_a \gg a^{1+\varepsilon_2} \quad \text{as } a \to 0 \qquad (2.61)$$

for each $\varepsilon_2 > 0$. Therefore, by the assumption $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$,

$$\mathbf{P}(|X^{(a)}| > 1/\kappa_a) = o(ac_a). \qquad (2.62)$$

and, as a consequence,

$$\mathbf{E}\left[e^{\theta_a X^{(a)}}; X^{(a)} \in (1/\kappa_a, 1/a]\right] = o(ac_a).$$

Combining the latter calculations with the definition of θ_a from (2.5) and the relation (2.59), we obtain

$$\begin{aligned} G_2 &\le e^{-\kappa_a t} - c_a e^{-\kappa_a t}\mathbf{E}\left[X^{(a)}e^{\theta_a X^{(a)}}; X^{(a)} \in [-1/\kappa_a, 1/\kappa_a]\right] + o(ac_a e^{-\kappa_a t}) \\ &= e^{-\kappa_a t} - ac_a e^{-\kappa_a t} + o(ac_a e^{-\kappa_a t}). \end{aligned}$$

Next, integrating by parts,

$$\begin{aligned} G_3 &= e^{-\kappa_a t}\int_{1/\kappa_a}^{t} e^{\kappa_a y}\mathbf{P}(X^{(a)} \in dy) \\ &= -\overline{F}(t) + e^{1-\kappa_a t}\overline{F}(1/\kappa_a) + \kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy. \end{aligned} \tag{2.63}$$

By plugging the latter results into (2.58) and using (2.62) to bound G_1 and $\overline{F}(1/\kappa_a)$, we get

$$\begin{aligned} &\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \\ &\qquad \le -ac_a e^{-\kappa_a t} + \kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy + o(ac_a e^{-\kappa_a t}) \end{aligned} \tag{2.64}$$

for all $t \ge 0$. If $0 \le t \le 1/\kappa_a$, the integral term does not give any positive contribution, we attain

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \le -ac_a e^{-\kappa_a t} + o(ac_a e^{-\kappa_a t}) \tag{2.65}$$

and the right hand side is negative for a small enough. If, on the other hand, $t > 1/\kappa_a$, we calculate the integral in (2.64) and in order to do so we consider different cases. First, consider $1/\kappa_a < t \le t(a) - C\ln(1/a)/\theta_a$ with a positive constant C such that $C < \varepsilon/4$, which is possible because of (2.48). We have

$$\begin{aligned} &\int_{1/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy \le \int_{1/\kappa_a}^{t(a)-C\ln(1/a)/\theta_a} e^{\theta_a y - g(y) - c_a y}dy \\ &\le \left(e^{\theta_a(t(a)-C\ln(1/a)/\theta_a) - g(t(a)-C\ln(1/a)/\theta_a)} + e^{\theta_a/\kappa_a - g(1/\kappa_a)}\right)\int_{1/\kappa_a}^{t(a)} e^{-c_a y}dy. \end{aligned} \tag{2.66}$$

Here we used the fact that $\theta_a y - g(y)$ is convex and takes its maximum at one of the edges of the interval $[1/\kappa_a, t(a) - C\ln(1/a)/\theta_a]$. This is true since $\theta_a y - g(y)$ is increasing for y such that $g'(y) < \theta_a$ and decreasing for y such that $g'(y) > \theta_a$ and g is concave. Furthermore,

$$\int_{1/\kappa_a}^{t(a)} e^{-c_a y}dy = \frac{1}{c_a}\left(e^{-c_a/\kappa_a} - e^{-c_a t(a)}\right) \le \frac{1}{c_a}. \tag{2.67}$$

Let us calculate the first term on the right hand side of (2.66). Recall that our choice of C and (2.48) give $t(a)/2 > C\ln(1/a)/\theta_a$. Hence, due to (2.50) and the definition of $t(a)$,

$$\begin{aligned}
&e^{\theta_a(t(a)-C\ln(1/a)/\theta_a)-g(t(a)-C\ln(1/a)/\theta_a)}\\
&\le e^{\theta_a t(a)-C\ln(1/a)-g(t(a))+2\varepsilon_1 C\ln(1/a)g(t(a))/(\theta_a t(a))}\\
&= e^{\theta_a t(a)-g(t(a))}e^{-C\ln(1/a)(1-2\varepsilon_1 g(t(a))/(\theta_a t(a)))}\\
&\sim a\theta_a e^{-C\ln(1/a)(1-2\varepsilon_1 g(t(a))/(\theta_a t(a)))}. \qquad (2.68)
\end{aligned}$$

By virtue of (2.48) and (2.49) there exists a constant C' such that

$$g(t(a))/(\theta_a t(a)) \le C'. \qquad (2.69)$$

Plugging this result into (2.68), we attain

$$e^{\theta_a(t(a)-C\ln(1/a)/\theta_a)-g(t(a)-C\ln(1/a)/\theta_a)} \le (1+o(1))\theta_a a^{1+C(1-2\varepsilon_1 C')}. \qquad (2.70)$$

Since ε_1 was arbitrary one can choose $\varepsilon_1 < 1/(2C')$ and thus, the relations (2.61) and (2.70) combined with $\theta_a = 2a/\sigma^2 + o(a)$ imply

$$e^{\theta_a(t(a)-C\ln(1/a)/\theta_a)-g(t(a)-C\ln(1/a)/\theta_a)} = O(a^{2+C(1-2\varepsilon_1 C')}) = o(c_a^2). \qquad (2.71)$$

On the other hand, from $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$, $\kappa_a = 2a/\sigma^2 + o(a)$ and (2.61) follows that

$$e^{-g(1/\kappa_a)} = \overline{F}(1/\kappa_a) = o(a^{2+\varepsilon}) = o(c_a^2). \qquad (2.72)$$

Plugging the results from (2.67), (2.71) and (2.72) into (2.66), we finally obtain

$$\kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy = o(ac_a e^{-\kappa_a t}) \qquad (2.73)$$

for $t \le t(a) - C\ln(1/a)/\theta_a$.

Next, consider the case $t > t(a) - C\ln(1/a)/\theta_a$. In this case we split the integral from (2.64) into two parts:

$$\int_{1/\kappa_a}^{t} = \int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} + \int_{t-C_1\ln t/\kappa_a}^{t}. \qquad (2.74)$$

with a constant C_1 to be chosen later. The first integral can be estimated similar to the case $t \le t(a) - C\ln(1/a)/\theta_a$. By (2.67) and (2.72),

$$\begin{aligned}
&\int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} e^{\kappa_a y}\overline{F}(y)dy = \int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} e^{\kappa_a y-g(y)}dy\\
&\le te^{\kappa_a(t-C_1\ln t/\kappa_a)-g(t-C_1\ln t/\kappa_a)} + e^{\theta_a/\kappa_a-g(1/\kappa_a)}\int_{1/\kappa_a}^{t} e^{-c_a y}dy\\
&\le t^{1-C_1}e^{\kappa_a t-g(t-C_1\ln t/\kappa_a)} + o(c_a). \qquad (2.75)
\end{aligned}$$

Here we used that $\kappa_a y - g(y) = y(\kappa_a - g(y)/y)$ is convex and takes its maximum at one of the edges. By virtue of $g(t)/t \searrow$ and $C\ln(1/a)/\theta_a \le t(a)/2$,

$$\frac{g(t)}{t} \le \frac{g(t(a) - C\ln(1/a)/\theta_a)}{t(a) - C\ln(1/a)/\theta_a} \le 2\frac{g(t(a))}{t(a)}$$

uniform in $t > t(a) - C\ln(1/a)/\theta_a$. The latter result plus (2.69) gives

$$\frac{g(t)}{\kappa_a t} \sim \frac{g(t)}{\theta_a t} = O(1) \tag{2.76}$$

uniform in $t > t(a) - C\ln(1/a)/\theta_a$. Let us consider t such that $t(a) - C\ln(1/a)/\theta_a \le t \le 4t(a)/\varepsilon$. In this case, for C_1 sufficiently small one has $C_1 \ln t/\kappa_a \le t/2$ and therefore (2.50) and (2.76) imply that there exists a constant C_2 such that

$$g(t) - g(t - C_1 \ln t/\kappa_a) \le 2\varepsilon_1 C_1 \ln t \frac{g(t)}{\kappa_a t} \le 2\varepsilon_1 C_1 C_2 \ln t.$$

Consequently,

$$e^{-g(t - C\ln t/\kappa_a)} \le t^{2\varepsilon_1 C_1 C_2}\overline{F}(t). \tag{2.77}$$

Since $\varepsilon_1 > 0$ can be chosen arbitrary small, one can choose ε_1 such that $2\varepsilon_1 C_2 < 1$. Furthermore, by (2.60),

$$\kappa_a t^{1-C_1(1-2\varepsilon_1 C_2)} = O(\kappa_a t(a)^{1-C_1(1-2\varepsilon_1 C_2)}) = o(1)$$

and consequently, by plugging (2.77) into (2.75), we obtain

$$\begin{aligned}\kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t - C_1\ln t/\kappa_a} e^{\kappa_a y}\overline{F}(y)dy &\le \kappa_a t^{1-C_1(1-2\varepsilon_1 C_2)}\overline{F}(t) + o(ac_a e^{-\kappa_a t})\\ &= o\left(\overline{F}(t)\right) + o(ac_a e^{-\kappa_a t}).\end{aligned} \tag{2.78}$$

Now, suppose t is such that $t \ge 4t(a)/\varepsilon$. In this case, by virtue of (2.48) and $\theta_a = 2a/\sigma^2 + o(a)$,

$$\begin{aligned}\frac{t}{2} - C_1\frac{\ln t}{\theta_a} \ge \frac{2t(a)}{\varepsilon} - C_1\frac{\ln(4t(a)/\varepsilon)}{\theta_a} &\ge (1 + o(1))\left(\frac{2\ln(1/a)}{\theta_a} - C_1\frac{\ln(4\ln(1/a)/\theta_a)}{\theta_a}\right)\\ &= (1 + o(1))\frac{\ln(1/a)}{\theta_a}(2 - C_1) \ge 0\end{aligned}$$

for all $C_1 < 2$. Therefore, by proceeding analogously to the latter calculations, we obtain for $t > 4t(a)/\varepsilon$, $1 < C_1 < 2$ and ε_1 small enough

$$\begin{aligned}\kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t - C_1\ln t/\kappa_a} e^{\kappa_a y}\overline{F}(y)dy &\le \kappa_a t^{1-C_1(1-2\varepsilon_1 C_2)}\overline{F}(t) + o(ac_a e^{-\kappa_a t})\\ &= o\left(\overline{F}(t)\right) + o(ac_a e^{-\kappa_a t}).\end{aligned} \tag{2.79}$$

Let us examine the second integral from the right hand side of (2.74). For ε_1 small enough, $2\varepsilon_1 g(t)/(\theta_a t) < 1$. Hence, by applying (2.50) with C_1 defined as above and ε_1 small enough, we attain

$$\begin{aligned}
&\kappa_a e^{-\kappa_a t} \int_{t-C_1\ln(1/a)/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy = \kappa_a \int_0^{C_1\ln(1/a)/\kappa_a} e^{-\kappa_a w - g(t-w)}dw \\
&\le \kappa_a e^{-g(t)} \int_0^{C_1\ln(1/a)/\kappa_a} e^{-\kappa_a w(1-2\varepsilon_1 g(t)/(\kappa_a t))}dw \le \frac{\overline{F}(t)}{1-2\varepsilon_1 g(t)/(\kappa_a t)} \\
&\sim \frac{\overline{F}(t)}{1-2\varepsilon_1 g(t)/(\theta_a t)} \qquad (2.80)
\end{aligned}$$

for all $t > t(a) - C\ln(1/a)/\theta_a$. By plugging (2.73), (2.78), (2.79) and (2.80) into (2.64) we obtain

$$\begin{aligned}
&\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \\
&\le -ac_a e^{-\kappa_a t} + \frac{\overline{F}(t)\mathbf{1}\{t \ge t(a) - C\ln(1/a)/\theta_a\}}{1-2\varepsilon_1 g(t)/(\theta_a t)} + o(ac_a e^{-\kappa_a t}) \\
&\qquad + o(\overline{F}(t))\mathbf{1}\{t \ge t(a) - C\ln(1/a))/\theta_a\}. \qquad (2.81)
\end{aligned}$$

The indicator function after the o-term shall mean that this o-term only appears if the condition of the indicator function is fulfilled.

Let us show that the latter inequality implies (2.57). One easily sees that (2.48) and the fact that $t(a)$ increases with the order of g imply that for all $0 < \delta < 1$ and our choice of C,

$$\delta t(a) < t(a) - C\ln(1/a)/\theta_a.$$

Consequently, for $0 \le t < \delta t(a)$ and a small enough,

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \le -ac_a e^{-\kappa_a t} + o(ac_a e^{-\kappa_a t}) \le 0.$$

It remains to show (2.56) and to do so we need to examine $\widehat{G}_{1-\varepsilon}$. By the definition of $\widehat{G}_{1-\varepsilon}$,

$$\begin{aligned}
a(1-\varepsilon)\mathbf{E}[\widehat{G}_{1-\varepsilon}(t - X^{(a)})] &= \int_{-\infty}^{t-\delta t(a)} F(dz)\overline{F}^I(t-z) \\
&= \left(\int_0^{t-\delta t(a)} + \int_{-\infty}^0\right) F(dz)\overline{F}^I(t-z).
\end{aligned}$$

Integrating the first integral by parts, we obtain

$$\begin{aligned}
&\int_0^{t-\delta t(a)} F(dz)\overline{F}^I(t-z)\\
&= \overline{F}(0)\overline{F}^I(t) - \overline{F}(t-\delta t(a))\overline{F}^I(\delta t(a)) + \int_0^{t-\delta t(a)} \overline{F}(z)\overline{F}(t-z)dz\\
&= \overline{F}(0)\overline{F}^I(t) - \overline{F}(t-\delta t(a))\overline{F}^I(\delta t(a)) + \int_0^{t/2} \overline{F}(z)\overline{F}(t-z)dz\\
&\qquad + \int_{\delta t(a)}^{t/2} \overline{F}(z)\overline{F}(t-z)dz
\end{aligned}$$

and by integrating the second integral by parts,

$$\int_{-\infty}^{0} F(dz)\overline{F}^I(t-z) = F(0)\overline{F}^I(t) - \int_{-\infty}^{0} \overline{F}(t-z)F(z)dz.$$

Combining the above identities, we get

$$\begin{aligned}
&a(1-\varepsilon)\mathbf{E}[\widehat{G}_{1-\varepsilon}(t-X^{(a)})]\\
&= \overline{F}^I(t) - \overline{F}(t-\delta t(a))\overline{F}^I(\delta t(a)) + \int_0^{t/2} \overline{F}(z)\overline{F}(t-z)dz\\
&\qquad + \int_{\delta t(a)}^{t/2} \overline{F}(z)\overline{F}(t-z)dz - \int_{-\infty}^{0} \overline{F}(t-z)F(z)dz.
\end{aligned}$$

Hence, for every $t \geq \delta t(a)$,

$$\begin{aligned}
&\mathbf{E}[\widehat{G}_{1-\varepsilon}(t-X^{(a)})] - \widehat{G}_{1-\varepsilon}(t)\\
&= \frac{\overline{F}(t)}{a(1-\varepsilon)}\left(-\frac{\overline{F}(t-\delta t(a))}{\overline{F}(t)}\overline{F}^I(\delta t(a)) + \int_0^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz\right.\\
&\qquad \left. + \int_{\delta t(a)}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz - \int_{-\infty}^{0} F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz\right). \qquad (2.82)
\end{aligned}$$

Consider $\nu = \nu(t)$ such that $(t/g(t))^{1-\delta_1} \ll \nu \ll t/g(t)$ with a small constant $\delta_1 > 0$. We will see later what small means in this context. By (2.50),

$$\begin{aligned}
\int_0^{\nu} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz &\leq \int_0^{\nu} \overline{F}(z)\exp\left\{2\varepsilon_1 z\frac{g(t)}{t}\right\}dz\\
&= \int_0^{\nu} \overline{F}(z)dz + 2\varepsilon_1\frac{g(t)}{t}\int_0^{\nu} z\overline{F}(z)dz + o\left(\frac{g(t)}{t}\right), \qquad (2.83)
\end{aligned}$$

where we used Taylor approximation and the assumption $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ with $\varepsilon > 0$ in the last equation. Due to our assumptions the function g is concave and increasing, consequently we have

$$g(z) - g(\nu) \geq g(t-\nu) - g(t-z)$$

for all $z \in (\nu, t/2]$. Hence,

$$\int_{\nu}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \leq t\exp\{g(t)-g(\nu)-g(t-\nu)\}$$
$$\leq t\exp\{-g(\nu)+\varepsilon_1\nu g(t)/t\} \sim t\exp\{-g(\nu)\}, \tag{2.84}$$

where we again used (2.50) and that $\nu \ll t/g(t)$. Furthermore, we have $g(t) \geq (2+\varepsilon)\ln t$ because of $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$. Hence, by (2.84) and recalling $g(t) \ll t^{\varepsilon_2}$ for all $\varepsilon_2 > 0$,

$$\int_{\nu}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \leq (1+o(1))t\nu^{-(2+\varepsilon)} = o\left(\frac{g(t)}{t}\right) \tag{2.85}$$

for $\nu \gg (t/g(t))^{1-\delta_1}$ with δ_1 small enough. In the case $t/2 \leq \delta t(a)$ one has

$$\int_{\delta t(a)}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \leq 0 \tag{2.86}$$

and if $\delta t(a) \leq t/2$ one can show in analogy to (2.83) that

$$\int_{\delta t(a)}^{\nu} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \leq \int_{\delta t(a)}^{\nu} \overline{F}(z)dz + 2\varepsilon_1\frac{g(t)}{t}\int_{\delta t(a)}^{\nu} z\overline{F}(z)dz \tag{2.87}$$

for $\delta t(a) \leq \nu$. Furthermore, because of $t(a) \gg 1/a$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$,

$$\int_{\delta t(a)}^{\nu} \overline{F}(z)dz = o(a) \quad \text{and} \quad \int_{\delta t(a)}^{\nu} z\overline{F}(z)dz = o(1). \tag{2.88}$$

Consequently, (2.87), (2.88) and (2.85) give

$$\int_{\delta t(a)}^{\nu} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz = o(a) + o\left(\frac{g(t)}{t}\right). \tag{2.89}$$

On the other hand, (2.50) implies

$$g(t-z) - g(t) \leq -2\varepsilon_1 z\frac{g(t-z)}{t-z} \leq -2\varepsilon_1 z\frac{g(t)}{t}$$

for all $z < 0$ such that $-z \leq \nu$. Thus, since $\nu \ll t/g(t)$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$,

$$\int_{-\infty}^{0} F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \geq \int_{-\nu}^{0} F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \geq \int_{-\nu}^{0} F(z)\exp\left\{2\varepsilon_1 z\frac{g(t)}{t}\right\}dz$$
$$= \int_{-\nu}^{0} F(z)dz + 2\varepsilon_1\frac{g(t)}{t}\int_{-\nu}^{0} zF(z)dz + o\left(\frac{g(t)}{t}\right). \tag{2.90}$$

Plugging (2.83), (2.85), (2.86), (2.89) and (2.90) into (2.82) with $(t/g(t))^{1-\delta_1} \ll \nu \ll t/g(t)$, where δ_1 is small enough, we obtain

$$\begin{aligned}&\mathbf{E}[\widehat{G}_{1-\varepsilon}(t-X^{(a)})]-\widehat{G}_{1-\varepsilon}(t)\\&\leq \frac{\overline{F}(t)}{a(1-\varepsilon)}\Bigg[\left(\int_0^{\nu}\overline{F}(z)dz-\int_{-\nu}^{0}F(z)dz\right)\\&\qquad+2\varepsilon_1\frac{g(t)}{t}\left(\int_0^{\nu}z\overline{F}(z)dz-\int_{-\nu}^{0}zF(z)dz\right)+o(a)+o\left(\frac{g(t)}{t}\right)\Bigg].\end{aligned} \tag{2.91}$$

Choose δ_1 so small that

$$\int_{\nu}^{\infty}\overline{F}(z)dz=o\left(\frac{g(t)}{t}\right)\quad\text{and}\quad\int_{-\infty}^{-\nu}F(z)dz=o\left(\frac{g(t)}{t}\right).$$

Then,

$$\int_0^{\nu}\overline{F}(z)dz-\int_{-\nu}^{0}F(z)dz=-a+o\left(\frac{g(t)}{t}\right) \tag{2.92}$$

and, by Fubini's theorem,

$$\int_0^{\nu}z\overline{F}(z)dz-\int_{-\nu}^{0}zF(z)dz=\frac{\sigma^2}{2}+o(1). \tag{2.93}$$

Hence, for $t\geq\delta t(a)$,

$$\begin{aligned}\mathbf{E}[\widehat{G}_{1-\varepsilon}(t-X^{(a)})]-\widehat{G}_{1-\varepsilon}(t)&\leq\frac{\overline{F}(t)}{a(1-\varepsilon)}\left(-a+2\varepsilon_1\frac{\sigma^2g(t)}{2t}+o(a)+o\left(\frac{g(t)}{t}\right)\right)\\&=\frac{\overline{F}(t)}{1-\varepsilon}\left(-1+2\varepsilon_1\frac{g(t)}{\theta_at}+o(1)+o\left(\frac{g(t)}{\theta_at}\right)\right),\end{aligned}$$

where we used $\theta_a\sim 2a/\sigma^2$. Since g is increasing and $g(t)/t$ is decreasing, the relation (2.76) implies that for all $t\geq\delta t(a)$

$$\frac{g(t)}{\theta_at}\leq\frac{g(\delta t(a))}{\theta_a\delta t(a)}\leq\frac{g(t(a))}{\theta_a\delta t(a)}=O(1) \tag{2.94}$$

and we finally obtain

$$\mathbf{E}[\widehat{G}_{1-\varepsilon}(t-X^{(a)})]-\widehat{G}_{1-\varepsilon}(t)\leq\frac{\overline{F}(t)}{1-\varepsilon}\left(-1+2\varepsilon_1\frac{g(t)}{\theta_at}+o(1)\right) \tag{2.95}$$

for $t\geq\delta t(a)$ and an arbitrary small $\varepsilon_1>0$.

With the results from (2.81) and (2.95) we can show (2.56). For $\delta t(a)\leq t\leq t(a)-C\ln(1/a)/\theta_a$, (2.81) and (2.95) give

$$\begin{aligned}&\mathbf{E}[G_{-c_a}(t-X^{(a)})]+\mathbf{E}[\widehat{G}_{1-\varepsilon}(t-X^{(a)})]-G_{-c_a}(t)-\widehat{G}_{1-\varepsilon}(t)\\&\leq -ac_ae^{-\kappa_at}+\frac{\overline{F}(t)}{1-\varepsilon}\left(-1+2\varepsilon_1\frac{g(t)}{\theta_at}\right)+o(ac_ae^{-\kappa_at})+o(\overline{F}(t)).\end{aligned}$$

For ε_1 small enough, (2.94) ensures that

$$\frac{2\varepsilon_1 g(t)}{\theta_a t} < 1 \tag{2.96}$$

uniform in $\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$. Hence,

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}_{1-\varepsilon}(t - X^{(a)})] - G_{-c_a}(t) - \widehat{G}_{1-\varepsilon}(t) \le 0$$

uniform in $\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$. Now, consider $t > t(a) - C\ln(1/a)/\theta_a$. Then, again by virtue of (2.81) and (2.95),

$$\begin{aligned}
&\mathbf{E}[G_{-c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}_{1-\varepsilon}(t - X^{(a)})] - G_{-c_a}(t) - \widehat{G}_{1-\varepsilon}(t) \\
&\quad \le -ac_a e^{-\kappa_a t} + \frac{\overline{F}(t)}{1-\varepsilon}\left(-1 + 2\varepsilon_1 \frac{g(t)}{\theta_a t} + \frac{(1-\varepsilon)}{1 - 2\varepsilon_1 g(t)/(\theta_a t)}\right) + o(ac_a e^{-\kappa_a t}) + o(\overline{F}(t)).
\end{aligned} \tag{2.97}$$

One can easily infer from (2.94) that

$$1 - \varepsilon < \left(1 - 2\varepsilon_1 \frac{g(t)}{\theta_a t}\right)^2$$

for $\varepsilon_1 > 0$ sufficiently small and by plugging this result into (2.97), we obtain

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}_{1-\varepsilon}(t - X^{(a)})] - G_{-c_a}(t) - \widehat{G}_{1-\varepsilon}(t) \le 0$$

uniform in $t > t(a) - C\ln(1/a)/\theta_a$. Summing up all the above results this means that $Y_n^{(1)}$ is a non-negative supermartingale.

Now, let us show that $(Y_n^{(2)})$ is a submartingale. Therefore it is sufficient to show that

$$\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}_{1+\varepsilon}(t - X^{(a)})] \ge \widetilde{G}_{c_a}(t) + \widehat{G}_{1+\varepsilon}(t), \quad t \ge 2\delta t(a), \tag{2.98}$$

and

$$\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] \ge \widetilde{G}_{c_a}(t), \quad t \in [0, 2\delta t(a)). \tag{2.99}$$

Let us first examine $\widehat{G}_{1+\varepsilon}$ for $t \ge 2\delta t(a)$. Due to (2.82),

$$\begin{aligned}
&\mathbf{E}[\widehat{G}_{1+\varepsilon}(t - X^{(a)})] - \widehat{G}_{1+\varepsilon}(t) \\
&\ge \frac{\overline{F}(t)}{a(1+\varepsilon)}\left(-\frac{\overline{F}(t - \delta t(a))}{\overline{F}(t)}\overline{F}^I(\delta t(a)) + \int_0^\nu \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz - \int_{-\infty}^0 F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz\right)
\end{aligned} \tag{2.100}$$

for $\nu = \nu(t)$ such that $(t/g(t))^{1-\delta_1} \ll \nu \ll t/g(t)$ where δ_1 is small. By using that $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ and choosing $\delta_1 > 0$ sufficiently small, we attain

$$\begin{aligned}
\int_{-\infty}^{-\nu} F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz &\le \int_{-\infty}^{-\nu} F(z)dz \le \nu^{-(1+\varepsilon)}\int_{-\infty}^{-\nu} |z|^{1+\varepsilon}F(z)dz \\
&= o(\nu^{-(1+\varepsilon)}) = o\left(\frac{g(t)}{t}\right).
\end{aligned} \tag{2.101}$$

For $t \geq 2\delta t(a)$, the inequalities (2.50) and (1.16) give

$$\frac{\overline{F}(t-\delta t(a))}{\overline{F}(t)} \leq e^{2\varepsilon_1 \delta t(a) g(t)/t} \leq e^{2\varepsilon_1 g(\delta t(a))}$$

for every $\varepsilon_1 > 0$. Furthermore, l'Hôspital's rule and (2.13) give

$$\frac{(1/g'(t))\overline{F}(t)}{\overline{F}^I(t)} \sim 1 + \frac{g''(t)}{(g'(t))^2} \sim 1 - b \quad \text{as } t \to \infty.$$

with $1 - b \in (1/2, 1]$. Hence,

$$\overline{F}^I(t) \sim \frac{\overline{F}(t)}{(1-b)g'(t)} \quad \text{as } t \to \infty. \tag{2.102}$$

Combining the latter results with (2.12), one attains

$$\frac{\overline{F}(t-\delta t(a))}{\overline{F}(t)} \overline{F}^I(\delta t(a)) \leq (1+o(1)) \frac{\delta t(a) \ln(\delta t(a))}{(1-b) g(\delta t(a))} e^{-(1-2\varepsilon_1) g(\delta t(a))}$$

and, by using $g(\delta t(a)) \geq (2+\varepsilon)\ln(\delta t(a))$ and $t(a) \geq 1/a$,

$$\frac{\overline{F}(t-\delta t(a))}{\overline{F}(t)} \overline{F}^I(\delta t(a)) \leq (1+o(1))(\delta t(a))^{1-(1-2\varepsilon_1)(2+\varepsilon)} = o(a) \tag{2.103}$$

for ε_1 small enough. Furthermore, since $\overline{F}(t-z) \geq \overline{F}(t)$ for all $z \geq 0$ and $\overline{F}(t-z) \leq \overline{F}(t)$ for all $z \leq 0$,

$$\int_0^{\nu} \overline{F}(z) \frac{\overline{F}(t-z)}{\overline{F}(t)} dz - \int_{-\nu}^{0} F(z) \frac{\overline{F}(t-z)}{\overline{F}(t)} dz \geq -a + o\left(\frac{g(t)}{t}\right), \tag{2.104}$$

where we used (2.88) and (2.92). Plugging the results from (2.101), (2.103) and (2.104) into (2.100), we obtain

$$\begin{aligned} \mathbf{E}[\widehat{G}_{1+\varepsilon}(t-X^{(a)})] - \widehat{G}_{1+\varepsilon}(t) &\geq \frac{\overline{F}(t)}{a(1+\varepsilon)} \left(-a + o(a) + o\left(\frac{g(t)}{t}\right)\right) \\ &\geq \frac{\overline{F}(t)}{1+\varepsilon} \left(-1 + o(1) + o\left(\frac{g(t)}{\theta_a t}\right)\right) = -\frac{\overline{F}(t)}{1+\varepsilon} + o(\overline{F}(t)), \end{aligned} \tag{2.105}$$

where we used (2.94) in the last equality.

Now, let us examine $\widetilde{G}_{c_a}$. Put $\lambda_a = \theta_a + c_a$, then

$$\begin{aligned} &\mathbf{E}[\widetilde{G}_{c_a}(t-X^{(a)})] \\ &= e^{-\lambda_a t}\left(\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \leq 1/a\right] + \mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (1/a, t]\right]\right) + e^{\alpha}\overline{F}(t). \end{aligned} \tag{2.106}$$

In the case $t < 1/a$ the expectation on the interval $(1/a, t]$ is used to denote the negative expectation on the interval $[t, 1/a)$. Using the bound $e^x \geq 1 + x$, the definition of θ_a, $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ and the relation $\theta_a = 2a/\sigma^2 + o(a)$, we obtain

$$\begin{aligned}
&\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \leq 1/a\right] \\
&\geq \mathbf{E}\left[e^{\theta_a X^{(a)}}; X^{(a)} \leq 1/a\right] + c_a \mathbf{E}\left[X^{(a)} e^{\theta_a X^{(a)}}; X^{(a)} \leq 1/a\right] \\
&\geq 1 + c_a \mathbf{E}\left[X^{(a)}; X^{(a)} \leq 1/a\right] + \theta_a c_a \mathbf{E}\left[(X^{(a)})^2; X^{(a)} \leq 1/a\right] \\
&\geq 1 - ac_a + \theta_a c_a \sigma^2 + o(ac_a) = 1 + ac_a + o(ac_a).
\end{aligned}$$

Plugging this result into (2.106), one attains

$$\begin{aligned}
&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) \\
&\geq ac_a e^{-\lambda_a t} + e^{-\lambda_a t}\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (1/a, t]\right] + e^{\alpha}\overline{F}(t) + o(ac_a e^{-\lambda_a t}) \qquad (2.107)
\end{aligned}$$

for all $t \geq 0$. Hence, (2.99) holds for $1/a \leq t < 2\delta t(a)$. In the case $0 \leq t \leq 1/a$, the Taylor expansion gives

$$\begin{aligned}
&\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right] \\
&\leq \overline{F}(t) + \lambda_a \mathbf{E}\left[X^{(a)}; X^{(a)} > t\right] + \lambda_a^2 \mathbf{E}\left[(X^{(a)})^2; X^{(a)} > t\right] \\
&\qquad + e^{\lambda_a/a}\lambda_a^3 \mathbf{E}\left[(X^{(a)})^3; X^{(a)} \in (t, 1/a]\right].
\end{aligned}$$

Using $\lambda_a/a = O(1)$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$, one can easily verify

$$e^{\lambda_a/a}\mathbf{E}\left[(X^{(a)})^3; X^{(a)} \in (t, 1/a]\right] = O(a^{-1+\varepsilon})$$

and thus, by (2.61),

$$\begin{aligned}
&\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right] \\
&\leq \overline{F}(t) + \lambda_a \mathbf{E}\left[X^{(a)}; X^{(a)} > t\right] + \lambda_a^2 \mathbf{E}\left[(X^{(a)})^2; X^{(a)} > t\right] + o(ac_a). \qquad (2.108)
\end{aligned}$$

Suppose that $t \geq 0$ is such that $t = O(1)$ as $a \to 0$. Then, the latter inequality gives

$$\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right] \leq \overline{F}(t) + o(1)$$

and by plugging this into (2.107) we attain (2.99). Now suppose $t \to \infty$ as $a \to 0$. Since the second moment is finite, integrating by parts gives

$$\mathbf{E}\left[(X^{(a)})^k; X^{(a)} > t\right] = t^k \overline{F}(t) + k\int_t^\infty u^{k-1}\overline{F}(u)du$$

for $k \in \{1, 2\}$. Therefore, by (2.108),

$$\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right]$$
$$\le (1 + \lambda_a t + \lambda_a^2 t^2)\overline{F}(t) + \lambda_a \overline{F}^I(t) + 2\lambda_a^2 \int_t^\infty u\overline{F}(u)du + o(ac_a), \quad 0 \le t \le 1/a.$$

Using (2.13) and the l'Hôspital rule, we conclude

$$\frac{1}{tg'(t)} = \frac{1/g'(t)}{t} \sim \frac{-g''(x)}{(g'(x))^2} \to b \in [0, 1/2) \quad \text{as } t \to \infty.$$

Consequently, the l'Hôspital rule gives

$$\frac{(t/g'(t))\overline{F}(t)}{\int_t^\infty u\overline{F}(u)du} \sim 1 + \frac{g''(t)}{(g'(t))^2} - \frac{1}{tg'(t)} \sim 1 - 2b \quad \text{as } t \to \infty$$

with $1 - 2b \in (0, 1]$ and hence

$$\int_t^\infty u\overline{F}(u)du \sim \frac{t\overline{F}(t)}{(1-2b)g'(t)} \quad \text{as } t \to \infty. \tag{2.109}$$

On the other hand, (2.12) and $g(t) \ge \ln t$ give

$$g'(t) \ge \frac{g(t)}{t \ln t} \ge \frac{1}{t}$$

and therefore we obtain by regarding (2.102) that

$$\overline{F}^I(t) \le (1 + o(1))\frac{t\overline{F}(t)}{(1-b)} \quad \text{and} \quad \int_t^\infty u\overline{F}(u)du \le (1 + o(1))\frac{t^2\overline{F}(t)}{(1-2b)}.$$

Consequently, because of $c_a = o(a)$,

$$\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right]$$
$$\le \left(1 + \left(1 + \frac{1}{(1-b)}\right)\lambda_a t + \left(1 + \frac{2}{(1-2b)}\right)\lambda_a^2 t^2\right)\overline{F}(t) + o(ac_a) + o(\overline{F}(t)). \tag{2.110}$$

Now suppose that $t \ll 1/a$. Then, $\theta_a t = o(1)$ and hence

$$\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right] \le \overline{F}(t) + o(ac_a) + o(\overline{F}(t)), \tag{2.111}$$

which, combined with (2.107), immediately implies (2.99). If t is such that $t \asymp 1/a$ with $t \le 1/a$, relation (2.62) gives $\overline{F}(t) = o(ac_a)$ and consequently by virtue of $\lambda_a = 2a/\sigma^2 + o(a)$

$$\mathbf{E}\left[e^{\lambda_a X^{(a)}}; X^{(a)} \in (t, 1/a]\right] = o(ac_a). \tag{2.112}$$

That means (2.99) is also true for $t \asymp 1/a$ with $t \le 1/a$. Combining the latter results, we conclude that (2.99) is true for all $t \le 1/a$.

Now, let us consider t such that $t > 2\delta t(a)$. In analogy to (2.64) one can derive the following inequality from (2.107) by integration by parts:

$$\begin{aligned}&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] - \widetilde{G}_{c_a}(t)\\ &\ge ac_a e^{-\lambda_a t} + \lambda_a e^{-\lambda_a t}\int_{1/\lambda_a}^{t} e^{\lambda_a u}\overline{F}(u)du + o(ac_a e^{-\lambda_a t}).\end{aligned} \tag{2.113}$$

Consider t such that $2\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$. In this case we infer from the latter inequality and (2.105) that

$$\begin{aligned}&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}_{1+\varepsilon}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) - \widehat{G}_{1+\varepsilon}(t)\\ &\ge -\frac{\overline{F}(t)}{1+\varepsilon} + ac_a e^{-\lambda_a t} + o(\overline{F}(t)) + o(ac_a e^{-\lambda_a t}).\end{aligned}$$

To infer from the latter result that (2.98) holds in this case it is sufficient to show that

$$\overline{F}(t) \ll ac_a e^{-\lambda_a t} \quad \text{as } a \to 0 \tag{2.114}$$

for $2\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$. To do so let us first show that the latter holds for $t = t(a) - C\ln(1/a)/\theta_a$. By (2.70),

$$e^{-g(t(a)-C\ln(1/a)/\theta_a)} \le (1+o(1))\theta_a a^{1+C(1-2\varepsilon_2 C')} e^{-\theta_a(t(a)-C\ln(1/a)/\theta_a)}$$

for $\varepsilon_2, C' > 0$. Hence, by (2.61) and $\theta_a = 2a/\sigma^2 + o(a)$,

$$e^{-g(t(a)-C\ln(1/a)/\theta_a)} \ll ac_a e^{-\theta_a(t(a)-C\ln(1/a)/\theta_a)}, \tag{2.115}$$

which is exactly (2.114) for $t = t(a) - C\ln(1/a)/\theta_a$. Let us show that

$$g(t) - \lambda_a t \searrow \quad \text{for } 2\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a, \tag{2.116}$$

which is equivalent to $g'(t) \le \lambda_a$ for $2\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$. By (2.9), $g(x)/x^{\varepsilon_1} \searrow$ for all $\varepsilon_1 > 0$ or equivalently

$$g'(x) \le \varepsilon_1 \frac{g(x)}{x}.$$

Using (2.94) and that $g(x)/x$ is decreasing in x, we obtain

$$g'(x) \le \varepsilon_1 \frac{g(x)}{x} \le \varepsilon_1 \frac{g(2\delta t(a))}{2\delta t(a)} \le \theta_a \le \lambda_a$$

for ε_1 small enough. Finally, by virtue of (2.116), the definition of c_a, (2.48) and (2.115) we see that (2.114) is true: For $2\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$,

$$\begin{aligned}\overline{F}(t)e^{\lambda_a t} &= e^{\lambda_a t - g(t)} \le e^{\lambda_a(t(a)-C\ln(1/a)/\theta_a) - g(t(a)-C\ln(1/a)/\theta_a)}\\ &\sim e^{\theta_a(t(a)-C\ln(1/a)/\theta_a) - g(t(a)-C\ln(1/a)/\theta_a)} \ll ac_a.\end{aligned} \tag{2.117}$$

It remains to consider $t > t(a) - C\ln(1/a)\theta_a$. In this case, (2.113) gives

$$\begin{aligned}&\mathbf{E}[\widetilde{G}_{c_a}(t-X^{(a)})] - \widetilde{G}_{c_a}(t)\\&\geq ac_a e^{-\lambda_a t} + \lambda_a e^{-\lambda_a t}\int_{t-C_1\ln(1/a)/\theta_a}^{t} e^{\lambda_a u}\overline{F}(u)du + o(ac_a e^{-\lambda_a t})\end{aligned} \tag{2.118}$$

with $C_1 > 0$ defined like in the proof that $(Y_n^{(1)})$ is a supermartingale. By the monotonicity of g,

$$\begin{aligned}&\lambda_a e^{-\lambda_a t}\int_{t-C_1\ln(1/a)/\theta_a}^{t} e^{\lambda_a u}\overline{F}(u)du = \lambda_a\int_0^{C_1\ln(1/a)/\theta_a} e^{-\lambda_a w - g(t-w)}dw\\&\geq \lambda_a\overline{F}(t)\int_0^{C_1\ln(1/a)/\theta_a} e^{-\lambda_a w}dw = \overline{F}(t)\left(1 - e^{-C_1(\lambda_a/\theta_a)\ln(1/a)}\right) = \overline{F}(t) + o(\overline{F}(t))\end{aligned}$$

and therefore

$$\begin{aligned}&\mathbf{E}[\widetilde{G}_{c_a}(t-X^{(a)})] - \widetilde{G}_{c_a}(t)\\&\geq ac_a e^{-\lambda_a t} + \overline{F}(t)\mathbf{1}\{t \geq t(a) - C\ln(1/a)/\theta_a\} + o(ac_a e^{-\lambda_a t}) + o(\overline{F}(t)).\end{aligned}$$

Combining this result with (2.105), we attain

$$\begin{aligned}&\mathbf{E}[\widetilde{G}_{c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}_{1+\varepsilon}(t-X^{(a)})] - \widetilde{G}_{c_a}(t) - \widehat{G}_{1+\varepsilon}(t)\\&\geq -\frac{\overline{F}(t)}{1+\varepsilon} + ac_a e^{-\lambda_a t} + \overline{F}(t)\mathbf{1}\{t \geq t(a) - C\ln(1/a)/\theta_a\} + o(ac_a e^{-\lambda_a t}) + o(\overline{F}(t)).\end{aligned}$$

Hence, (2.98) is also true for $t > t(a) - C\ln(1/a)/\theta_a$.

2.3.4 Proof of Proposition 2.5

The proof goes along the same line as the proof of Proposition 2.4. However, for reasons of completeness, we give the whole proof. During the whole proof we assume a to be sufficiently small, even if not explicitly mentioned.

In analogy to (2.56) and (2.57), one sees that the supermartingale property for $(\widetilde{Y}_n^{(1)})$ is equivalent to

$$\mathbf{E}[G_{-c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}'_{1-\varepsilon}(t-X^{(a)})] \leq G_{-c_a}(t) + \widehat{G}'_{1-\varepsilon}(t), \quad t > \delta t(a), \tag{2.119}$$

and

$$\mathbf{E}[G_{-c_a}(t-X^{(a)})] \leq G_{-c_a}(t), \quad t \in [0, \delta t(a)). \tag{2.120}$$

To examine G_{-c_a} we start from the bound (2.58), which states

$$\begin{aligned}&\mathbf{E}[G_{-c_a}(t-X^{(a)})]\\&\leq e^{-\kappa_a t}\mathbf{P}(X^{(a)} \leq -1/\kappa_a) + e^{-\kappa_a t}\mathbf{E}\left[e^{\kappa_a X^{(a)}}; X^{(a)} \in (-1/\kappa_a, 1/\kappa_a]\right]\\&\qquad + e^{-\kappa_a t}\mathbf{E}\left[e^{\kappa_a X^{(a)}}; X^{(a)} \in (1/\kappa_a, t]\right] + \overline{F}(t)\\&=: G_1 + G_2 + G_3 + G_4.\end{aligned} \tag{2.121}$$

Regarding (2.53) we can infer from (2.9) that, for all $\varepsilon_1 > 0$ with $\gamma + \varepsilon_1 < 1$, there exists a positive constant C_1 such that $\theta_a t(a) \le C_1 t(a)^{\gamma+\varepsilon_1}$ or, equivalently,

$$t(a) \le C_1^{1/(1-(\gamma+\varepsilon_1))} \theta_a^{-1/(1-(\gamma+\varepsilon_1))}. \tag{2.122}$$

Consequently,

$$c_a = \frac{1}{\ln(1/a)t(a)} \ge \frac{\theta_a^{1/(1-(\gamma+\varepsilon_1))}}{\ln(1/a) C_1^{1/(1-(\gamma+\varepsilon_1))}}. \tag{2.123}$$

Suppose without loss of generality that $\widetilde{\gamma} < 1$. Then, for $\varepsilon_1 = (\widetilde{\gamma} - \gamma)/2$, one has $\gamma + \varepsilon_1 \le \widetilde{\gamma} < 1$ and therefore, as $a \to 0$,

$$\kappa_a^{1+1/(1-\widetilde{\gamma})} \sim \theta_a^{1+1/(1-\widetilde{\gamma})} \ll \frac{\theta_a^{1+1/(1-(\gamma+\varepsilon_1))}}{\ln(1/a)}.$$

Hence, the Markov inequality and the assumption $\mathbf{E}[|\min\{0, X^{(a)}\}|^{1+1/(1-\widetilde{\gamma})}] < \infty$ imply that (2.62) is also valid in the case $\gamma > 0$ and one can proceed like in the proof of Proposition 2.4 to verify

$$G_1 + G_2 \le e^{-\kappa_a t} - a c_a e^{-\kappa_a t} + o(a c_a e^{-\kappa_a t}).$$

Therefore, (2.62) and (2.63) give

$$\begin{aligned}\mathbf{E}[G_{-c_a}(t - X^{(a)})] &- G_{-c_a}(t) \\ &\le -a c_a e^{-\kappa_a t} + \kappa_a e^{-\kappa_a t} \int_{1/\kappa_a}^{t} e^{\kappa_a y} \overline{F}(y) dy + o(a c_a e^{-\kappa_a t})\end{aligned} \tag{2.124}$$

for all $t \ge 0$. If $0 \le t \le 1/\kappa_a$, the integral term is non-positive, we attain

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \le -a c_a e^{-\kappa_a t} + o(a c_a e^{-\kappa_a t}), \tag{2.125}$$

and this means the right hand side of the latter inequality is negative for $0 \le t \le 1/\kappa_a$ and a small enough. Now, consider $1/\kappa_a < t \le t(a) - C\ln(1/a)/\theta_a$ with a constant $C > 0$ to be chosen later. One of the main differences of this proof compared to the proof of Proposition 2.4 is that C has to be chosen large in this proof while C was small in the proof of Proposition 2.4. By (2.66) and (2.67),

$$\begin{aligned}&\int_{1/\kappa_a}^{t} e^{\kappa_a y} \overline{F}(y) dy \le \int_{1/\kappa_a}^{t(a)-C\ln(1/a)/\theta_a} e^{\theta_a y - g(y) - c_a y} dy \\ &\le \frac{1}{c_a} e^{\theta_a (t(a) - C\ln(1/a)/\theta_a) - g(t(a) - C\ln(1/a)/\theta_a)} + \frac{1}{c_a} e^{\theta_a/\kappa_a - g(1/\kappa_a)}.\end{aligned} \tag{2.126}$$

Because of (2.9),

$$g(t-w) \ge g(t)\left(1 - \frac{w}{t}\right)^{\gamma+\varepsilon_1}, \quad 0 \le w \le t.$$

Furthermore, for $w \ll t$ as $t \to \infty$, one has

$$\left(1-\frac{w}{t}\right)^{\gamma+\varepsilon_1} = 1-(\gamma+\varepsilon_1)\frac{w}{t}+o\left(\frac{w}{t}\right)$$

and consequently

$$g(t)-g(t-w) \le (\gamma+\varepsilon_1)w\frac{g(t)}{t}+o\left(w\frac{g(t)}{t}\right) \le \widehat{\gamma}w\frac{g(t)}{t}, \quad w \ll t, \tag{2.127}$$

for all $0<\gamma+\varepsilon_1<\widehat{\gamma}<1$. Regarding (2.51) and that $g(u) \gg u^{\varepsilon_2}$ as $u \to \infty$ for some $\varepsilon_2 > 0$, one obtains

$$t(a) \gg \left(\frac{1}{\theta_a}\right)^{1/(1-\varepsilon_2)} \gg \ln(1/a)/\theta_a \tag{2.128}$$

and therefore, by applying the inequality (2.127) and the definition of $t(a)$, one gets

$$\begin{aligned}
&e^{\theta_a(t(a)-C\ln(1/a)/\theta_a)-g(t(a)-C\ln(1/a)/\theta_a)}\\
&\le e^{\theta_a t(a)-C\ln(1/a)-g(t(a))+\widehat{\gamma}C\ln(1/a)g(t(a))/(\theta_a t(a))}\\
&= e^{\theta_a t(a)-g(t(a))}e^{-C\ln(1/a)(1-\widehat{\gamma}g(t(a))/(\theta_a t(a))}\\
&\sim a\theta_a e^{-C\ln(1/a)(1-\widehat{\gamma}g(t(a))/(\theta_a t(a)))}.
\end{aligned} \tag{2.129}$$

By (2.53), we conclude that there exists $\varepsilon_3>0$ such that $\widehat{\gamma}(1+\varepsilon_3)<1$ with

$$\widehat{\gamma}\frac{g(t(a))}{\theta_a t(a)} = \widehat{\gamma}+o(1) \le \widehat{\gamma}(1+\varepsilon_3)$$

and, plugging the latter result into (2.129), we attain

$$e^{\theta_a(t(a)-C\ln(1/a)/\theta_a)-g(t(a)-C\ln(1/a)/\theta_a)} \le (1+o(1))\theta_a a^{1+C(1-\widehat{\gamma}(1+\varepsilon_3))}. \tag{2.130}$$

Consequently, we infer from (2.123), (2.130) and $\theta_a = 2a/\sigma^2+o(a)$ that, for C large enough,

$$e^{\theta_a(t(a)-C\ln(1/a)/\theta_a)-g(t(a)-C\ln(1/a)/\theta_a)} = o(c_a^2). \tag{2.131}$$

On the other hand, by (2.123),

$$\overline{F}(1/\kappa_a) \le e^{-\kappa_a^{-\varepsilon_2}} = o(c_a^2). \tag{2.132}$$

Plugging the results from (2.131) and (2.132) into (2.126), we obtain

$$\kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy = o(ac_a e^{-\kappa_a t}) \tag{2.133}$$

for $t \le t(a)-C\ln(1/a)/\theta_a$ if C is large enough.

Next, consider the case $t > t(a) - C\ln(1/a)/\theta_a$. We split the integral from (2.124) into two parts

$$\int_{1/\kappa_a}^{t} = \int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} + \int_{t-C_1\ln t/\kappa_a}^{t} \tag{2.134}$$

with a large constant C_1 to be defined later. Note that because of (2.128) we have $t(a) \gg C_1\ln(1/a)/\theta_a$ for any constant $C_1 > 0$. By virtue of (2.132), the inequality (2.75) is also valid in the case $\gamma > 0$ and therefore

$$\int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} e^{\kappa_a y}\overline{F}(y)dy \le t^{1-C_1}e^{\kappa_a t - g(t-C_1\ln t/\kappa_a)} + o(c_a). \tag{2.135}$$

Regarding (2.128), one has $C_1\ln t/\kappa_a \ll t$ for all $C_1 > 0$ and $t > t(a) - C\ln(1/a)/\theta_a$. Hence, inequality (2.127) gives

$$g(t) - g(t - C_1\ln t/\kappa_a) \le \widehat{\gamma}C_1\ln t\frac{g(t)}{\kappa_a t} \tag{2.136}$$

with $\widehat{\gamma} < 1$. Let us show that, uniform in $t > t(a) - C\ln(1/a)/\theta_a$,

$$\frac{g(t)}{\theta_a t} \le 1 + o(1). \tag{2.137}$$

By using (2.128), (2.53) and that $g(t)/t$ is decreasing for t large enough,

$$\frac{g(t)}{\theta_a t} \le \frac{g(t(a) - C\ln(1/a)/\theta_a)}{\theta_a(t(a) - C\ln(1/a)/\theta_a)} \le \frac{g(t(a))}{\theta_a(t(a) - C\ln(1/a)/\theta_a)} \sim \frac{g(t(a))}{\theta_a t(a)} \sim 1 \tag{2.138}$$

uniform in $t > t(a) - C\ln(1/a)/\theta_a$. Combining this result with (2.136), we conclude that there exists some $\varepsilon_4 > 0$ such that $\widehat{\gamma}(1+\varepsilon_4) < 1$ and

$$g(t) - g(t - C_1\ln t/\kappa_a) \le \widehat{\gamma}(1+\varepsilon_4)C_1\ln t.$$

Therefore, by plugging the latter result into (2.135),

$$\kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} e^{\kappa_a y}\overline{F}(y)dy \le \kappa_a t^{1-C_1(1-\widehat{\gamma}(1+\varepsilon_4))}\overline{F}(t) + o(ac_a e^{-\kappa_a t})$$

and the latter inequality implies that, for C_1 large enough,

$$\kappa_a e^{-\kappa_a t}\int_{1/\kappa_a}^{t-C_1\ln t/\kappa_a} e^{\kappa_a y}\overline{F}(y)dy = o(\overline{F}(t)) + o(ac_a e^{-\kappa_a t}). \tag{2.139}$$

Let us examine the second integral from the right hand side of (2.134). Due to (2.127), (2.137) and $c_a = o(a)$,

$$\begin{aligned}
&\kappa_a e^{-\kappa_a t}\int_{t-C_1\ln t/\kappa_a}^{t} e^{\kappa_a y}\overline{F}(y)dy = \kappa_a\int_0^{C_1\ln t/\kappa_a} e^{-\kappa_a w - g(t-w)}dw \\
&\le \kappa_a e^{-g(t)}\int_0^{C_1\ln t/\kappa_a} e^{-\kappa_a w(1-(\gamma+2\varepsilon_1)g(t)/(\kappa_a t))}dw \\
&\le \frac{\overline{F}(t)}{1-(\gamma+2\varepsilon_1)g(t)/(\kappa_a t)} \sim \frac{\overline{F}(t)}{1-(\gamma+2\varepsilon_1)g(t)/(\theta_a t)}
\end{aligned} \tag{2.140}$$

for all $t > t(a) - C\ln(1/a)/\theta_a$ and ε_1 such that $\gamma + 2\varepsilon_1 < 1$. By plugging (2.133), (2.139) and (2.140) into (2.124), we attain

$$\begin{aligned}\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \\ \leq -ac_a e^{-\kappa_a t} + \frac{\overline{F}(t)\mathbf{1}\{t > t(a) - C\ln(1/a)/\theta_a\}}{1 - (\gamma + 2\varepsilon_1)g(t)/(\theta_a t)} + o(ac_a e^{-\kappa_a t}) \\ + o(\overline{F}(t))\mathbf{1}\{t > t(a) - C\ln(1/a))/\theta_a\}. \qquad (2.141)\end{aligned}$$

The indicator function after the o-term is used to denote that this o-term only appears if the condition of the indicator function is fulfilled. Let us show that the latter inequality implies (2.120). By virtue of (2.128), one has $t(a) - C\ln(1/a)/\theta_a \sim t(a) > \delta t(a)$ and consequently

$$\mathbf{E}[G_{-c_a}(t - X^{(a)})] - G_{-c_a}(t) \leq -ac_a e^{-\kappa_a t} + o(ac_a e^{-\kappa_a t}) \leq 0$$

for $0 \leq t < \delta t(a)$.

It remains to show (2.119) for $t \geq \delta t(a)$ and therefore we need to examine $\widehat{G}'_{1-\varepsilon}$. By the definition of $\widehat{G}'_{1-\varepsilon}$,

$$a(1-\varepsilon)\mathbf{E}[\widehat{G}'_{1-\varepsilon}(t - X^{(a)})] = \left(\int_{-\infty}^{0} + \int_{0}^{t-\delta t(a)}\right) F(dz)L^{(N)}(t-z)\overline{F}^I(t-z). \quad (2.142)$$

To examine the integral terms from (2.142) suppose $\nu = \nu(t)$ is such that $(t/g(t))^{1-\delta_1} \ll \nu \ll t/g(t)$ with a small constant $\delta_1 \in (0,1)$. We will see later what small means in this context. Note that $L^{(N)}(\cdot)$ is differentiable for N large enough since $L^{(\infty)}(t) < \infty$ for $t \geq \delta t(a)$ due to (2.94). For the rest of the proof assume without loss of generality that N sufficiently large. By integration by parts,

$$\begin{aligned}\int_0^{\nu} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z) \\ \leq \overline{F}(0)L^{(N)}(t)\overline{F}^I(t) + \int_0^{\nu} \overline{F}(z)L^{(N)}(t-z)\overline{F}(t-z)dz \\ + \int_0^{\nu} \overline{F}(z)(\partial_z L^{(N)}(t-z))\overline{F}^I(t-z)dz. \qquad (2.143)\end{aligned}$$

Since $g(u)$ is increasing, (2.94) ensures that, for $z \in [0,\nu]$,

$$\frac{g(t-z)}{\theta_a(t-z)} \leq \frac{g(t)}{\theta_a(t-\nu)} = \frac{g(t)}{\theta_a t} + o(1) \qquad (2.144)$$

and from this one can easily infer that

$$L^{(N)}(t-z) \leq L^{(N)}(t) + o(1) \qquad (2.145)$$

By additionally regarding (2.14), an easy calculation gives

$$0 < -\partial_t L^{(N)}(t) = O(1/t) \qquad (2.146)$$

and, for $z \in [0, \nu]$,

$$\partial_z L^{(N)}(t-z) \le -\partial_t L^{(N)}(t) + o(-\partial_t L^{(N)}(t)). \tag{2.147}$$

Hence,

$$\begin{aligned}
&\int_0^\nu F(dz) L^{(N)}(t) \overline{F}^I(t-z) \\
&\le \overline{F}(0) L^{(N)}(t) \overline{F}^I(t) + (1+o(1)) \Bigg(L^{(N)}(t) \int_0^\nu \overline{F}(z) \overline{F}(t-z) dz \\
&\qquad\qquad -\partial_t L^{(N)}(t) \int_0^\nu \overline{F}(z) \overline{F}^I(t-z) dz \Bigg). \qquad (2.148)
\end{aligned}$$

On the other hand, integrating by parts gives

$$\begin{aligned}
&\int_{-\infty}^0 F(dz) L^{(N)}(t-z) \overline{F}^I(t-z) \\
&\le F(0) L^{(N)}(t) \overline{F}^I(t) - \int_{-\infty}^0 F(z) L^{(N)}(t-z) \overline{F}(t-z) dz \\
&\qquad - \int_{-\infty}^0 F(z) (\partial_z L^{(N)}(t-z)) \overline{F}^I(t-z) dz. \qquad (2.149)
\end{aligned}$$

The inequalities (2.145) and (2.147) imply that, for $z \in [-\nu, 0]$,

$$L^{(N)}(t-z) \ge L^{(N)}(t) + o(L^{(N)}(t)) \quad \text{and} \quad \partial_z L^{(N)}(t-z) \ge -\partial_t L^{(N)}(t) + o(-\partial_t L^{(N)}(t)).$$

Therefore,

$$\begin{aligned}
\int_{-\infty}^0 F(z) L^{(N)}(t-z) \overline{F}(t-z) dz &\ge \int_{-\nu}^0 F(z) L^{(N)}(t-z) \overline{F}(t-z) dz \\
&\ge (1+o(1)) L^{(N)}(t) \int_{-\nu}^0 F(z) \overline{F}(t-z) dz \qquad (2.150)
\end{aligned}$$

and

$$\begin{aligned}
\int_{-\infty}^0 F(z) (\partial_z L^{(N)}(t-z)) \overline{F}^I(t-z) dz &\ge \int_{-\nu}^0 F(z) (\partial_z L^{(N)}(t-z)) \overline{F}^I(t-z) dz \\
&\ge -(1+o(1)) \partial_z L^{(N)}(t) \int_{-\nu}^0 F(z) \overline{F}^I(t-z) dz. \quad (2.151)
\end{aligned}$$

By combining the results (2.149), (2.150) and (2.151) with (2.148), one attains

$$\begin{aligned}
&\int_{-\infty}^\nu F(dz) L^{(N)}(t) \overline{F}^I(t-z) \\
&\le L^{(N)}(t) \overline{F}^I(t) + (1+o(1)) \Bigg(L^{(N)}(t) \left[\int_0^\nu \overline{F}(z) \overline{F}(t-z) dz - \int_{-\nu}^0 F(z) \overline{F}(t-z) dz \right] \\
&\qquad -\partial_t L^{(N)}(t) \left[\int_0^\nu \overline{F}(z) \overline{F}^I(t-z) dz - \int_{-\nu}^0 F(z) \overline{F}^I(t-z) dz \right] \Bigg). \qquad (2.152)
\end{aligned}$$

By (2.127),

$$\int_0^\nu \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \le \int_0^\nu \overline{F}(z)\exp\left\{(\gamma+2\varepsilon_1)z\frac{g(t)}{t}\right\}dz$$
$$= \int_0^\nu \overline{F}(z)dz + (\gamma+2\varepsilon_1)\frac{g(t)}{t}\int_0^\nu z\overline{F}(z)dz + o\left(\frac{g(t)}{t}\right), \quad (2.153)$$

where we used Taylor approximation and the assumption $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ with $\varepsilon > 0$ in the last equation. Furthermore, (2.127) implies that, for all $z \le 0$ such that $-z \le \nu$,

$$g(t-z) - g(t) \le -(\gamma+\varepsilon_1)z\frac{g(t-z)}{t-z} + o\left(-z\frac{g(t-z)}{t-z}\right)$$
$$\le -(\gamma+2\varepsilon_1)z\frac{g(t)}{t}. \quad (2.154)$$

Thus, since $\nu \ll t/g(t)$ and $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$,

$$\int_{-\nu}^0 F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \ge \int_{-\nu}^0 F(z)\exp\left\{(\gamma+2\varepsilon_1)z\frac{g(t)}{t}\right\}dz$$
$$= \int_{-\nu}^0 F(z)dz + (\gamma+2\varepsilon_1)\frac{g(t)}{t}\int_{-\nu}^0 zF(z)dz + o\left(\frac{g(t)}{t}\right). \quad (2.155)$$

Combining the results from (2.153) and (2.155), we attain

$$\int_0^\nu \overline{F}(z)\overline{F}(t-z)dz - \int_{-\nu}^0 F(z)\overline{F}(t-z)dz$$
$$\le \overline{F}(t)\left(\left\{\int_0^\nu \overline{F}(z)dz - \int_{-\nu}^0 F(z)dz\right\}\right.$$
$$\left. +(\gamma+2\varepsilon_1)\frac{g(t)}{t}\left\{\int_0^\nu z\overline{F}(z)dz - \int_{-\nu}^0 zF(z)dz\right\} + o\left(\frac{g(t)}{t}\right)\right) \quad (2.156)$$

and, by virtue of (2.92), (2.93), $\theta_a = 2a/\sigma^2 + o(a)$ and (2.94), the latter implies

$$\int_0^\nu \overline{F}(z)\overline{F}(t-z)dz - \int_{-\nu}^0 F(z)\overline{F}(t-z)dz$$
$$\le \overline{F}(t)\left(-a + (\gamma+2\varepsilon_1)\frac{\sigma^2 g(t)}{2t} + o\left(\frac{g(t)}{t}\right)\right)$$
$$= -a\overline{F}(t)\left(1 - (\gamma+2\varepsilon_1)\frac{g(t)}{\theta_a t} + o(1)\right). \quad (2.157)$$

Moreover, remark that since $\overline{F}^I(u)$ is decreasing in u,

$$\int_0^\nu \overline{F}(z)\overline{F}^I(t-z)dz - \int_{-\nu}^0 F(z)\overline{F}^I(t-z)dz$$
$$\le \overline{F}^I(t-\nu)\int_0^\nu \overline{F}(z)dz - \overline{F}^I(t+\nu)\int_{-\nu}^0 F(z)dz. \quad (2.158)$$

By using (2.15) instead of (2.13), one can show similar to (2.102) that

$$\overline{F}^I(t) \sim \frac{\overline{F}(t)}{g'(t)}. \tag{2.159}$$

Since (2.14) is valid for all $\varepsilon_1 > 0$, one has $g'(t) \sim \gamma g(t)/t$ and therefore

$$\overline{F}^I(t) \sim \frac{\overline{F}(t)}{g'(t)} \sim \frac{t\overline{F}(t)}{\gamma g(t)}. \tag{2.160}$$

Using the inequality (2.127) and $\nu \ll t/g(t)$, one sees that

$$\overline{F}(t-\nu) \leq \overline{F}(t)e^{\widehat{\gamma}\nu g(t)/t} \sim \overline{F}(t) \tag{2.161}$$

and consequently, by using that $u/g(u)$ is increasing,

$$\overline{F}^I(t-\nu) \sim \frac{(t-\nu)\overline{F}(t-\nu)}{\gamma g(t-\nu)} \leq (1+o(1))\frac{t\overline{F}(t-\nu)}{\gamma g(t)} = (1+o(1))\frac{t\overline{F}(t)}{\gamma g(t)}. \tag{2.162}$$

In analogy to the latter inequality one can show by using (2.154) instead of (2.127) that

$$\overline{F}^I(t+\nu) \geq (1+o(1))\frac{t\overline{F}(t)}{\gamma g(t)} \tag{2.163}$$

and, by plugging (2.162) and (2.163) into (2.158) and regarding (2.94), one obtains

$$\begin{aligned} &\int_0^\nu \overline{F}(z)\overline{F}^I(t-z)dz - \int_{-\nu}^0 F(z)\overline{F}^I(t-z)dz \\ &\leq (1+o(1))\frac{t\overline{F}(t)}{\gamma g(t)}\left(\int_0^\nu \overline{F}(z)dz - \int_{-\nu}^0 F(z)dz\right) = -\frac{at\overline{F}(t)}{\gamma g(t)} + o\left(\frac{at\overline{F}(t)}{g(t)}\right). \end{aligned} \tag{2.164}$$

This means the right hand side of the latter inequality is negative and, by regarding (2.94) and combining the latter results, (2.152) and (2.157), we conclude

$$\begin{aligned} &\int_{-\infty}^\nu F(dz)L^{(N)}(t)\overline{F}^I(t-z) \\ &\leq L^{(N)}(t)\overline{F}^I(t) - L^{(N)}(t)a\overline{F}(t)\left(1-(\gamma+2\varepsilon_1)\frac{g(t)}{\theta_a t} + o(1)\right). \end{aligned} \tag{2.165}$$

Suppose $\nu \leq t - \delta t(a)$. Then, by the definition of $L^{(N)}(\cdot)$,

$$\int_\nu^{t-\delta t(a)} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z) \leq N\int_\nu^{t-\delta t(a)} F(dz)\overline{F}^I(t-z) \tag{2.166}$$

and, by integration by parts,

$$\int_\nu^{t-\delta t(a)} F(dz)\overline{F}^I(t-z) \leq \overline{F}(\nu)\overline{F}^I(t-\nu) + \int_\nu^{t-\delta t(a)} \overline{F}(z)\overline{F}(t-z)dz. \tag{2.167}$$

Let us mention that $\nu \le t/2$ due to the definition of ν. We split the integral on right hand side of the latter inequality as follows:

$$\begin{aligned}\int_{\nu}^{t-\delta t(a)} \overline{F}(z)\overline{F}(t-z)dz &= \int_{\nu}^{t/2} \overline{F}(z)\overline{F}(t-z)dz + \int_{t/2}^{t-\delta t(a)} \overline{F}(z)\overline{F}(t-z)dz \\ &= \int_{\nu}^{t/2} \overline{F}(z)\overline{F}(t-z)dz + \int_{\delta t(a)}^{t/2} \overline{F}(z)\overline{F}(t-z)dz. \qquad (2.168)\end{aligned}$$

Due to our assumptions, the function g is concave and increasing. Consequently, one has

$$g(z) - g(\nu) \ge g(t-\nu) - g(t-z)$$

for all $z \in (\nu, t/2]$ and hence

$$\begin{aligned}\int_{\nu}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz &\le t\exp\{g(t) - g(\nu) - g(t-\nu)\} \\ &\le t\exp\{-g(\nu) + \widehat{\gamma}\nu g(t)/t\} \sim t\exp\{-g(\nu)\},\end{aligned}$$

where we used (2.127) and $\nu \ll t/g(t)$. Since $g(\nu) \gg \nu^{\varepsilon_2}$ for some $\varepsilon_2 > 0$ and $g(t) = o(t^{\gamma+2\varepsilon_1})$ for all $\varepsilon_1 > 0$ such that $\gamma + 2\varepsilon_1 < 1$, we obtain

$$\begin{aligned}\int_{\nu}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz &\le (1+o(1))t\exp\{-g(\nu)\} \le (1+o(1))t\exp\{-\nu^{\varepsilon_2}\} \\ &= o\left(t\exp\{-(t/g(t))^{\varepsilon_2(1-\delta_1)}\}\right) = o\left(\frac{g(t)}{t}\right). \qquad (2.169)\end{aligned}$$

If $t/2 \ge \delta t(a)$ and $\delta t(a) \le \nu$, one sees similar to (2.153) that

$$\begin{aligned}\int_{\delta t(a)}^{t/2} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz &\le \int_{\delta t(a)}^{t/2} \overline{F}(z)dz + (\gamma+2\varepsilon_1)\frac{g(t)}{t}\int_{\delta t(a)}^{t/2} z\overline{F}(z)dz + o\left(\frac{g(t)}{t}\right) \\ &= o(a) + o\left(\frac{g(t)}{t}\right), \qquad (2.170)\end{aligned}$$

where we used (2.88) in the last equality. Plugging (2.169) and (2.170) into (2.168), we get

$$\int_{\nu}^{t-\delta t(a)} \overline{F}(z)\overline{F}(t-z)dz = o(a\overline{F}(t)) + o\left(\frac{g(t)}{t}\overline{F}(t)\right). \qquad (2.171)$$

and finally, by combining the latter result, (2.166) and (2.165) with (2.142), we obtain by regarding (2.94) that

$$\mathbf{E}[\widehat{G}'_{1-\varepsilon}(t - X^{(a)})] - \widehat{G}'_{1-\varepsilon}(t) \le -\frac{L^{(N)}(t)\overline{F}(t)}{1-\varepsilon}\left(1 - \frac{(\gamma+2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right). \qquad (2.172)$$

Combining this bound with the bound from (2.141), we can show (2.119). For $\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$, (2.141) and (2.172) give

$$\begin{aligned}&\mathbf{E}[G_{-c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}'_{1-\varepsilon}(t-X^{(a)})] - G_{-c_a}(t) - \widehat{G}'_{1-\varepsilon}(t)\\&\le -ac_a e^{-\kappa_a t} - \frac{L^{(N)}(t)\overline{F}(t)}{1-\varepsilon}\left(1 - \frac{(\gamma+2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right) + o(ac_a e^{-\kappa_a t})\end{aligned} \tag{2.173}$$

with $0 < \gamma < \gamma + 2\varepsilon_1 < 1$. In analogy to (2.114), one can show by using (2.130) and (2.137) instead of (2.70) and (2.96) respectively, that

$$\overline{F}(t) \ll ac_a e^{-\theta_a t} \quad \text{as } a \to 0 \tag{2.174}$$

for $\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$ with C large enough. Hence, (2.94) and (2.173) give

$$\mathbf{E}[G_{-c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}'_{1-\varepsilon}(t-X^{(a)})] - G_{-c_a}(t) - \widehat{G}'_{1-\varepsilon}(t) \le 0$$

for $\delta t(a) \le t \le t(a) - C\ln(1/a)/\theta_a$. If $t > t(a) - C\ln(1/a)/\theta_a$, (2.141) and (2.172) imply

$$\begin{aligned}&\mathbf{E}[G_{-c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}'_{1-\varepsilon}(t-X^{(a)})] - G_{-c_a}(t) - \widehat{G}'_{1-\varepsilon}(t)\\&\le -ac_a e^{-\kappa_a t} - \frac{L^{(N)}(t)\overline{F}(t)}{1-\varepsilon}\left(1 - \frac{(\gamma+2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right)\\&\qquad + \frac{\overline{F}(t)}{1-(\gamma+2\varepsilon_1)g(t)/(\theta_a t)} + o(ac_a e^{-\kappa_a t}) + o(\overline{F}(t)).\end{aligned}$$

Due to (2.137), one has

$$L^{(N)}(t) = \left(\sum_{k=0}^{\infty}\left(\frac{\gamma g(t)}{\theta_a t}\right)^k\right)^2 \wedge N = \frac{1}{(1-\gamma g(t)/(\theta_a t))^2} \wedge N$$

and therefore, for N large enough and $\varepsilon_1 > 0$ small enough,

$$\frac{L^{(N)}(t)}{1-\varepsilon} = \frac{1}{(1-\varepsilon)(1-\gamma g(t)/(\theta_a t))^2} > \frac{1}{(1-(\gamma+2\varepsilon_1)g(t)/(\theta_a t))^2}.$$

We finally obtain

$$\mathbf{E}[G_{-c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}'_{1-\varepsilon}(t-X^{(a)})] - G_{-c_a}(t) - \widehat{G}'_{1-\varepsilon}(t) \le 0$$

in the case $t > t(a) - C\ln(1/a)/\theta_a$. Summing up all the above results this means that $(\widetilde{Y}_n^{(1)})$ is a non-negative supermartingale.

Let us show that $(\widetilde{Y}_n^{(2)})$ is a submartingale, which is equivalent to

$$\mathbf{E}[\widetilde{G}_{c_a}(t-X^{(a)})] + \mathbf{E}[\widehat{G}'_{1+\varepsilon}(t-X^{(a)})] \ge \widetilde{G}_{c_a}(t) + \widehat{G}'_{1+\varepsilon}(t), \quad t \ge 2\delta t(a), \tag{2.175}$$

and

$$\mathbf{E}[\widetilde{G}_{c_a}(t-X^{(a)})] \ge \widetilde{G}_{c_a}(t), \quad 0 \le t < 2\delta t(a). \tag{2.176}$$

Let us first examine $\widehat{G}'_{1+\varepsilon}$ for $t \geq 2\delta t(a)$. Using that $t - \delta t(a) \geq \nu$ for $t \geq 2\delta t(a)$, the definition of $\widehat{G}'_{1+\varepsilon}$ gives

$$a(1+\varepsilon)\mathbf{E}[\widehat{G}'_{1+\varepsilon}(t - X^{(a)})] = \int_{-\infty}^{t-\delta t(a)} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z)$$
$$\geq \int_{-\infty}^{\nu} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z). \qquad (2.177)$$

Integrating by parts we get

$$\int_0^{\nu} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z)$$
$$= \overline{F}(0)L^{(N)}(t)\overline{F}^I(t) - \overline{F}(\nu)L^{(N)}(t-\nu)\overline{F}^I(t-\nu) + \int_0^{\nu} \overline{F}(z)L^{(N)}(t-z)\overline{F}(t-z)dz$$
$$+ \int_0^{\nu} \overline{F}(z)(\partial_z L^{(N)}(t-z))\overline{F}^I(t-z)dz. \qquad (2.178)$$

Since $g(u)/u$ is decreasing,

$$L^{(N)}(t-z) \geq L^{(N)}(t) \qquad (2.179)$$

for all $z \in [0,\nu]$. Using that $g(u)/u$ is decreasing in u and regarding (2.14), one sees by a straightforward calculation that, for $z \in [0,\nu]$,

$$\partial_z L^{(N)}(t-z) \geq -\partial_t L^{(N)}(t) + o(-\partial_t L^{(N)}(t)). \qquad (2.180)$$

Furthermore, since there exists some $\varepsilon_2 > 0$ such that $g(u) \gg u^{\varepsilon_2}$ as $u \to \infty$,

$$\overline{F}(\nu) \leq e^{-\nu^{\varepsilon_2}} \leq e^{-(t/g(t))^{\varepsilon_2(1-\delta_1)}} = o\Bigg(\left(\frac{g(t)}{t}\right)^2\Bigg)$$

and consequently (2.162) implies that

$$\overline{F}(\nu)L^{(N)}(t-\nu)\overline{F}^I(t-\nu) \leq N\overline{F}(\nu)\overline{F}^I(t-\nu) = o\left(\frac{g(t)}{t}\overline{F}(t)\right). \qquad (2.181)$$

Plugging (2.179), (2.180) and (2.181) into (2.178), we attain

$$\int_0^{\nu} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z)$$
$$= \overline{F}(0)L^{(N)}(t)\overline{F}^I(t) + L^{(N)}(t)\int_0^{\nu} \overline{F}(z)\overline{F}(t-z)dz - \partial_t L^{(N)}(t)\int_0^{\nu} \overline{F}(z)\overline{F}^I(t-z)dz$$
$$+ o\left(\frac{g(t)}{t}\overline{F}(t)\right). \qquad (2.182)$$

On the other side, again by integration by parts,

$$\int_{-\infty}^{0} F(dz)L^{(N)}(t-z)\overline{F}^{I}(t-z)$$
$$= F(0)L^{(N)}(t)\overline{F}^{I}(t) - \int_{-\infty}^{0} F(z)L^{(N)}(t-z)\overline{F}(t-z)dz$$
$$- \int_{-\infty}^{0} F(z)(\partial_z L^{(N)}(t-z))\overline{F}^{I}(t-z)dz. \qquad (2.183)$$

In analogy to (2.161) one can show by using (2.154) instead of (2.127) that $\overline{F}(t+\nu) \sim \overline{F}(t)$ and hence, by virtue of $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$,

$$\int_{-\infty}^{-\nu} F(z)L^{(N)}(t-z)\overline{F}(t-z)dz \leq N\overline{F}(t+\nu)\int_{-\infty}^{-\nu} F(z)dz$$
$$\sim N\overline{F}(t)\int_{-\infty}^{-\nu} F(z)dz = o\left(\frac{g(t)}{t}\overline{F}(t)\right) \qquad (2.184)$$

for δ_1 small enough. One has

$$L^{(N)}(t-z) \leq L^{(N)}(t),$$

and consequently

$$\int_{-\nu}^{0} F(z)L^{(N)}(t-z)\overline{F}(t-z)dz \leq L^{(N)}(t)\int_{-\nu}^{0} F(z)\overline{F}(t-z)dz. \qquad (2.185)$$

An easy calculation shows that there exists a constant $C' > 0$ such that, for $z \in [\nu, \infty)$,

$$\partial_z L^{(N)}(t-z) \leq \frac{C'}{t-\nu} \sim \frac{C'}{t}$$

and, by virtue of (2.160),

$$\overline{F}^{I}(t+\nu) \leq \overline{F}^{I}(t) \sim \frac{t\overline{F}(t)}{\gamma g(t)}.$$

Therefore, $\mathbf{E}[|X^{(a)}|^{2+\varepsilon}] < \infty$ implies

$$\int_{-\infty}^{-\nu} F(z)(\partial_z L^{(N)}(t-z))\overline{F}^{I}(t-z)dz \leq (1+o(1))\frac{C'\overline{F}^{I}(t+\nu)}{t}\int_{-\infty}^{-\nu} F(z)dz$$
$$\leq \frac{C'\overline{F}(t)}{\gamma g(t)}\int_{-\infty}^{-\nu} F(z)dz = o\left(\frac{g(t)}{t}\overline{F}(t)\right). \qquad (2.186)$$

Using that $g(u)/u$ is decreasing in u, a straightforward calculation gives

$$\partial_z L^{(N)}(t-z) \leq -\partial_t L^{(N)}(t)$$

for $z \in [-\nu, 0]$ and thus,

$$\int_{-\nu}^{0} F(z)(\partial_z L^{(N)}(t-z))\overline{F}^I(t-z)dz \le -\partial_t L^{(N)}(t) \int_{-\nu}^{0} F(z)\overline{F}^I(t-z)dz. \tag{2.187}$$

Plugging (2.184), (2.185), (2.186) and (2.187) into (2.183), we attain

$$\begin{aligned}
&\int_{-\infty}^{0} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z) \\
&\ge F(0)L^{(N)}(t)\overline{F}^I(t) - L^{(N)}(t)\int_{-\nu}^{0} F(z)\overline{F}(t-z)dz \\
&\qquad + \partial_t L^{(N)}(t)\int_{-\nu}^{0} F(z)\overline{F}^I(t-z)dz + o\left(\frac{g(t)}{t}\overline{F}(t)\right).
\end{aligned}$$

The latter inequality combined with (2.182) implies

$$\begin{aligned}
&\int_{-\infty}^{\nu} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z) \\
&\ge L^{(N)}(t)\overline{F}^I(t) + L^{(N)}(t)\left(\int_{0}^{\nu} \overline{F}(z)\overline{F}(t-z)dz - \int_{-\nu}^{0} F(z)\overline{F}(t-z)dz\right) \\
&- \partial_t L^{(N)}(t)\left(\int_{0}^{\nu} \overline{F}(z)\overline{F}^I(t-z)dz - \int_{-\nu}^{0} F(z)\overline{F}^I(t-z)dz\right) + o\left(\frac{g(t)}{t}\overline{F}(t)\right). \qquad (2.188)
\end{aligned}$$

Using (2.10) instead of (2.9), one can show similar to (2.127) that, for all $\varepsilon_1 > 0$ such that $\gamma - 2\varepsilon_1 > 0$,

$$g(t) - g(t-w) \ge (\gamma - 2\varepsilon_1)w\frac{g(t)}{t}, \quad w \ll t. \tag{2.189}$$

Hence, in analogy to (2.83) and (2.90), one can show by regarding (2.92), (2.93) and (2.94) that,

$$\int_{0}^{\nu} \overline{F}(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz - \int_{-\nu}^{0} F(z)\frac{\overline{F}(t-z)}{\overline{F}(t)}dz \ge -a\left(1 - \frac{(\gamma-2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right). \tag{2.190}$$

The relation (2.164) combined with (2.146) and (2.94) gives

$$\begin{aligned}
&- \partial_t L^{(N)}(t)\left(\int_{0}^{\nu} \overline{F}(z)\overline{F}^I(t-z)dz - \int_{-\nu}^{0} F(z)\overline{F}^I(t-z)dz\right) \\
&\qquad = O\left(-\frac{a\overline{F}(t)}{\gamma g(t)}\right) = o\left(\frac{g(t)}{t}\overline{F}(t)\right) \qquad (2.191)
\end{aligned}$$

and by plugging (2.190) and (2.191) into (2.188) one obtains

$$\begin{aligned}
&\int_{-\infty}^{\nu} F(dz)L^{(N)}(t-z)\overline{F}^I(t-z) \\
&\ge L^{(N)}(t)\overline{F}^I(t) - aL^{(N)}(t)\overline{F}(t)\left(1 - \frac{(\gamma-2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right), \qquad (2.192)
\end{aligned}$$

where we used $\theta_a = 2a/\sigma^2 + o(a)$ and (2.94) in the last equality. Combining the latter inequality with (2.177), one attains

$$\mathbf{E}[\widehat{G}'_{1+\varepsilon}(t - X^{(a)})] - \widehat{G}'_{1+\varepsilon}(t) \geq -\frac{L^{(N)}(t)\overline{F}(t)}{1+\varepsilon}\left(1 - \frac{(\gamma - 2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right). \quad (2.193)$$

Now, let us examine $\widetilde{G}_{c_a}$. The result (2.107) proves that (2.176) is valid for $1/a \leq t < 2\delta t(a)$ and (2.111) and (2.112) together with (2.107) show that (2.176) is valid for $t \leq 1/a$. It remains to show (2.175) for $t \geq 2\delta t(a)$. Put $\lambda_a = \theta_a + c_a$ and consider $2\delta t(a) \leq t \leq t(a) - C\ln(1/a)/\theta_a$ with a large constant C. We will see later what large means in this context. Then, (2.113) implies

$$\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) \geq ac_a e^{-\lambda_a t} + o(ac_a e^{-\lambda_a t})$$

and by combining the latter result with (2.193) we obtain

$$\begin{aligned}&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}'_{1+\varepsilon}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) - L\widehat{G}'_{1+\varepsilon}(t)\\ &\geq ac_a e^{-\lambda_a t} - \frac{L^{(N)}(t)\overline{F}(t)}{1+\varepsilon}\left(1 - \frac{(\gamma - 2\varepsilon_1)g(t)}{\theta_a t} + o(1)\right) + o(\overline{F}(t)) + o(ac_a e^{-\lambda_a t}).\end{aligned}$$

Due to (2.94), the latter implies (2.175) if one can show that

$$\overline{F}(t) \ll ac_a e^{-\lambda_a t} \quad \text{as } a \to 0 \quad (2.194)$$

for $2\delta t(a) \leq t \leq t(a) - C\ln(1/a)/\theta_a$. Since $\lambda_a y - g(y) = y(\lambda_a - g(y)/y)$ is convex and takes its maximum at one of the edges, it is sufficient to consider $t = 2\delta t(a)$ and $t = t(a) - C\ln(1/a)/\theta_a$. First, we show that the latter holds for $t = 2\delta t(a)$. By (2.9) and $g(t(a)) \sim \theta_a t(a)$,

$$g(\delta t(a)) \geq (\delta t(a))^{\gamma+\varepsilon_1}\frac{g(t(a))}{t(a)^{\gamma+\varepsilon_1}} = \delta^{\gamma+\varepsilon_1}g(t(a)) \sim \delta^{\gamma+\varepsilon_1}\lambda_a t(a)$$

for all $\varepsilon_1 > 0$ and therefore

$$\overline{F}(\delta t(a))e^{\lambda_a \delta t(a)} = e^{\lambda_a \delta t(a) - g(\delta t(a))} \leq e^{\lambda_a t(a)(\delta - \delta^{\gamma+\varepsilon_1})}.$$

Choosing ε_1 so small that $\gamma + \varepsilon_1 < 1$, we conclude by virtue of (2.52) and the definition of c_a that

$$\overline{F}(\delta t(a))e^{\lambda_a \delta t(a)} \leq e^{\lambda_a t(a)(\delta - \delta^{\gamma+\varepsilon_1})} = o(ac_a). \quad (2.195)$$

On the other hand, by virtue of (2.130), there exist constants $\widehat{\gamma}$ and ε_3 with $0 < \widehat{\gamma} < \widehat{\gamma} + \varepsilon_3 < 1$ such that

$$e^{-g(t(a) - C\ln(1/a)/\theta_a)} \leq (1 + o(1))\theta_a a^{1+C(1-\widehat{\gamma}(1+\varepsilon_3))}e^{-\theta_a(t(a) - C\ln(1/a)/\theta_a)}. \quad (2.196)$$

For C large enough, (2.122), the definition of c_a and the relation $\theta_a = 2a/\sigma^2 + o(a)$ give $a^{C(1-\widehat{\gamma}(1+\varepsilon_3))} \ll c_a$ and we attain

$$e^{-g(t(a) - C\ln(1/a)/\theta_a)} \ll ac_a e^{-\theta_a(t(a) - C\ln(1/a)/\theta_a)}. \quad (2.197)$$

This means (2.194) is true for $t = 2\delta t(a)$ and $t = t(a) - C\ln(1/a)/\theta_a$ and therefore (2.175) holds for $2\delta t(a) \leq t \leq t(a) - C\ln(1/a)/\theta_a$.

Now, consider $t > t(a) - C\ln(1/a)/\theta_a$. Then, (2.118) gives

$$\begin{aligned}
&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) \\
&\geq ac_a e^{-\lambda_a t} + \lambda_a e^{-\lambda_a t} \int_{t-C_1\ln(1/a)/\theta_a}^{t} e^{\lambda_a u}\overline{F}(u)du + o(ac_a e^{-\lambda_a t}).
\end{aligned} \tag{2.198}$$

Using (2.189) one can show similar to (2.140) that, for $t > t(a) - C\ln(1/a)/\theta_a$,

$$\lambda_a e^{-\lambda_a t} \int_{t-C\ln(1/a)/\theta_a}^{t} e^{\lambda_a u}\overline{F}(u)du \geq \frac{\overline{F}(t) + o(\overline{F}(t))}{1 - (\gamma - 2\varepsilon_1)g(t)/(\theta_a t)}$$

for all $\varepsilon_1 > 0$ such that $\gamma - 2\varepsilon_1 > 0$ and by plugging this result into (2.198),

$$\begin{aligned}
&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) \\
&\geq ac_a e^{-\lambda_a t} + \frac{\overline{F}(t)}{1 - (\gamma - 2\varepsilon_1)g(t)/(\theta_a t)} + o(\overline{F}(t)) + o(ac_a e^{-\lambda_a t}).
\end{aligned} \tag{2.199}$$

Combining (2.193) and (2.199), we attain

$$\begin{aligned}
&\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}'_{1+\varepsilon}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) - \widehat{G}'_{1+\varepsilon}(t) \\
&\geq ac_a e^{-\lambda_a t} - \frac{L^{(N)}(t)\overline{F}(t)}{1+\varepsilon}\left(1 - \frac{(\gamma - 2\varepsilon_1)g(t)}{\theta_a t}\right) + \frac{\overline{F}(t)}{1 - (\gamma - 2\varepsilon_1)g(t)/(\theta_a t)} \\
&\qquad + o(\overline{F}(t)) + o(ac_a e^{-\lambda_a t}).
\end{aligned}$$

For N large enough and $\varepsilon_1 > 0$ small enough, the definition of $L^{(N)}(t)$ gives

$$\frac{L^{(N)}(t)}{1+\varepsilon} = \frac{1}{(1+\varepsilon)(1 - \gamma g(t)/(\theta_a t))^2} < \frac{1}{(1 - (\gamma + 2\varepsilon_1)g(t)/(\theta_a t))^2},$$

which immediately implies

$$\mathbf{E}[\widetilde{G}_{c_a}(t - X^{(a)})] + \mathbf{E}[\widehat{G}'_{1+\varepsilon}(t - X^{(a)})] - \widetilde{G}_{c_a}(t) - \widehat{G}'_{1+\varepsilon}(t) \geq 0$$

for $t > t(a) - C\ln(1/a)/\theta_a$. Summing up all the above results this means that $(\widetilde{Y}_n^{(2)})$ is a non-negative submartingale.

2.3.5 Proof of Lemma 2.6

By virtue of $x \geq \delta x(a)$, (2.159), (2.14) and $e^{-\theta_a x} = O(\overline{F}^I(x)/a)$,

$$e^{-\theta_a x} = O\left(\frac{x}{ag(x)}e^{-g(x)}\right).$$

Equivalently, one has

$$\limsup_{a\to 0}\left(g(x)-\theta_a x-\ln\left(\frac{x}{ag(x)}\right)\right)<\infty.$$

Since $g(x)\to\infty$ as $x\to\infty$ and $x\geq 1/a$, we can infer from the latter that

$$\limsup_{a\to 0}\left(g(x)-\theta_a x-2\ln x\right)<\infty. \tag{2.200}$$

Because of $\gamma>0$, it is $\ln x=o(g(x))$ as $x\to\infty$ and hence the latter is equivalent to

$$\limsup_{a\to 0}(g(x)-\theta_a x)<\infty$$

and therefore

$$\limsup_{a\to 0}\frac{g(x)}{\theta_a x}\leq 1.$$

2.3.6 Proof of Theorem 2.7

Let us first consider the case $F\in K_\gamma$, $\gamma\in(0,1)$. Fix some $\delta\in(0,1)$ and $N>1/(1-\gamma)$ such that $(\widetilde{Y}_n^{(1)})$ is a non-negative supermartingale for all a small enough. This is possible due to Proposition 2.5. Then,

$$\begin{aligned}
&e^{-(\theta_a-c_a)x}+\frac{L^{(N)}(x)}{a(1-\varepsilon)}\overline{F}^I(x)\mathbf{1}\{x\geq\delta t(a)\}=\widetilde{Y}_0^{(1)}\geq\mathbf{E}[\widetilde{Y}_\infty^{(1)}]\\
&=\mathbf{E}\left[G_{-c_a}(x-S_{\mu_x}^{(a)});\mu_x<\infty\right]+\mathbf{E}\left[\widehat{G}'_{1-\varepsilon}(x-S_{\mu_{x-\delta x(a)}}^{(a)});\mu_{x-\delta x(a)}<\infty\right].
\end{aligned}\tag{2.201}$$

Furthermore, by the definition of G_{-c_a} and $\widehat{G}'_{1-\varepsilon}$,

$$\mathbf{E}[G_{-c_a}(x-S_{\mu_x}^{(a)});\mu_x<\infty]=\mathbf{P}(\mu_x<\infty)=\mathbf{P}(M^{(a)}>x)$$

and

$$\mathbf{E}[\widehat{G}'_{1-\varepsilon}(x-S_{\mu_{x-\delta x(a)}}^{(a)});\mu_{x-\delta x(a)}<\infty]=0.$$

Plugging these results into (2.201), we attain

$$\mathbf{P}(M^{(a)}>x)\leq e^{-(\theta_a-c_a)x}+\frac{L^{(N)}(x)}{a(1-\varepsilon)}\overline{F}^I(x)\mathbf{1}\{x\geq\delta x(a)\}.$$

In analogy to (2.194) one can show that $(1/a)\overline{F}^I(x)\gg e^{-(\theta_a-c_a)x}$ for all $x\geq x(a)+C\ln(1/a)/\theta_a$ for C large enough and, by the definition of c_a, $c_a x\to 0$ as $a\to 0$ for all x such that $x\leq x(a)+C\ln(1/a)/\theta_a$. Furthermore, since $\varepsilon>0$ is arbitrary, one can let $\varepsilon\to 0$ and we conclude that, uniformly in $x\geq 0$,

$$\mathbf{P}(M^{(a)}>x)\leq(1+o(1))\left(e^{-\theta_a x}+\frac{L^{(N)}(x)}{a}\overline{F}^I(x)\mathbf{1}\{x\geq\delta x(a)\}\right). \tag{2.202}$$

By regarding (2.160) and (2.94), one sees similar to (2.174) that, for a large enough constant $C > 0$,

$$\overline{F}^I(x) \ll a e^{-\theta_a x}, \quad \delta x(a) \le x \le x(a) - C\ln(1/a)/\theta_a \tag{2.203}$$

Furthermore, for $x > x(a) - C\ln(1/a)/\theta_a$ the bound from (2.137) gives

$$\sum_{k=0}^{\infty} \left(\frac{\gamma g(x)}{\theta_a x}\right)^k = \frac{1}{1-\gamma g(x)/(\theta_a x)}.$$

and since $N > 1/(1-\gamma)$ one has

$$L^{(N)}(x) = \frac{1}{(1-\gamma g(x)/(\theta_a x))^2}, \quad x > x(a) - C\ln(1/a)/\theta_a. \tag{2.204}$$

Consequently, (2.94), (2.202) and (2.204) give that, uniformly in $x \ge 0$,

$$\mathbf{P}(M^{(a)} > x) \le (1+o(1))\left(e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x > x(a) - C\ln(1/a)/\theta_a\}}{a(1-\gamma g(x)/(\theta_a x))^2}\right). \tag{2.205}$$

On the other hand, by virtue of Proposition 2.5 one can choose $\delta \in (0,1/2)$ arbitrary small and $N > 1/(1-\gamma)$ such that $\widetilde{Y}_n^{(2)}$ is a non-negative submartingale. Hence,

$$\begin{aligned} &e^{-(\theta_a+c_a)x} + \frac{L^{(N)}(x)}{a(1+\varepsilon)}\overline{F}^I(x)\mathbf{1}\{x \ge \delta x(a)\} = \widetilde{Y}_0^{(2)} \le \mathbf{E}[\widetilde{Y}_\infty^{(2)}] \\ &\quad = \mathbf{E}\left[\widetilde{G}_{c_a}(x - S_{\mu_x}^{(a)}); \mu_x < \infty\right] + \mathbf{E}\left[\widehat{G}'_{1+\varepsilon}(x - S^{(a)}_{\mu_{x-2\delta x(a)}}); \mu_{x-2\delta x(a)} < \infty\right]. \end{aligned} \tag{2.206}$$

By definition of $\widetilde{G}_{c_a}$,

$$\mathbf{E}\left[\widetilde{G}_{c_a}(x - S_{\mu_x}^{(a)}); \mu_x < \infty\right] = e^{\alpha}\mathbf{P}(M^{(a)} > x) \tag{2.207}$$

and, since $\widehat{G}'_{1+\varepsilon}(u)$ is decreasing in u, the definition of $L^{(N)}$ implies that

$$\begin{aligned} \mathbf{E}\left[\widehat{G}'_{1+\varepsilon}(x - S^{(a)}_{\mu_{x-2\delta x(a)}}); \mu_{x-2\delta x(a)} < \infty\right] &\le \widehat{G}'_{1+\varepsilon}(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) \\ &\le \frac{N}{a(1+\varepsilon)}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)). \end{aligned} \tag{2.208}$$

Let us bound the term on the right hand side of the latter inequality and to do so we consider different regions of x separately. First, consider $x \le k_\gamma \delta x(a)$, where $k_\gamma := \min\{k \ge 4 : k^{1-\gamma} > 2\}$. In this case, one sees similar to (2.174) that, for δ small enough,

$$\overline{F}(x) \ll a c_a e^{-k_\gamma \theta_a x}, \quad \delta x(a) \le x \le k_\gamma \delta x(a). \tag{2.209}$$

From this, we infer from (2.160) and the definition of c_a,

$$\frac{1}{a}\overline{F}^I(\delta x(a)) \le \frac{x(a)}{a}\overline{F}(\delta x(a)) \ll x(a) c_a e^{-\theta_a k_\gamma \delta x(a)} \ll e^{-\theta_a k_\gamma \delta x(a)} \le e^{-\theta_a x} \tag{2.210}$$

for all $x \leq k_\gamma \delta x(a)$. This gives

$$\frac{1}{a}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) = o(e^{-\theta_a x}), \quad x \leq k_\gamma \delta x(a). \tag{2.211}$$

Next, consider $x > k_\gamma \delta x(a)$. Using the definition of $L^{(N)}$ and the bound from (2.205), one obtains

$$\begin{aligned}
&\frac{1}{a}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) \\
&\leq (1+o(1))\Bigg(\frac{1}{a}\overline{F}^I(\delta x(a))e^{-\theta_a(x-2\delta x(a))} \\
&\qquad + \frac{\overline{F}^I(\delta x(a))\overline{F}^I(x-2\delta x(a))\mathbf{1}\{x-2\delta x(a) > x(a) - C\ln(1/a)/\theta_a\}}{a^2(1-\gamma g(x-2\delta x(a))/(\theta_a(x-2\delta x(a))))^2}\Bigg).
\end{aligned} \tag{2.212}$$

Let us examine the first term of the right hand side of (2.212). The result from (2.210) implies that, for all x,

$$\frac{1}{a}\overline{F}^I(\delta x(a))e^{-\theta_a(x-2\delta x(a))} = o(e^{-\theta_a x}) \tag{2.213}$$

and it remains to bound the second term on the right hand side of (2.212) for $x > k_\gamma \delta x(a)$. Due to (2.94),

$$\frac{1}{(1-\gamma g(x-2\delta x(a))/(\theta_a(x-2\delta x(a))))^2} = O(1) \tag{2.214}$$

for $x - 2\delta x(a) \geq \delta x(a)$ and since $k_\gamma \geq 3$ the latter relation especially holds for $x > k_\gamma \delta x(a)$. Since g is concave, $g'(x-2\delta x(a)) \geq g'(x)$ and therefore, by using (2.159) and (1.17),

$$\begin{aligned}
\overline{F}^I(x-2\delta x(a)) &\sim \frac{\overline{F}(x-2\delta x(a))}{g'(x-2\delta x(a))} \leq \frac{\overline{F}(x-2\delta x(a))}{g'(x)} \\
&\leq \frac{\overline{F}(x)e^{2\delta\alpha x(a)g(x)/x}}{g'(x)} \sim \overline{F}^I(x)e^{2\delta\alpha x(a)g(x)/x}
\end{aligned}$$

for some $0 < \alpha < 1$. The definition of k_γ implies that $k_\gamma^{1-\widehat{\gamma}} \geq 2$ if $\widehat{\gamma}$ with $\gamma < \widehat{\gamma} < 1$ is sufficiently close to γ. Then, using (2.9) and $x > k_\gamma \delta x(a)$, we obtain

$$\delta x(a)\frac{g(x)}{x} \leq \left(\frac{\delta x(a)}{x}\right)^{1-\widehat{\gamma}} g(\delta x(a)) \leq k_\gamma^{\widehat{\gamma}-1} g(\delta x(a)) \leq g(\delta x(a))/2$$

and consequently

$$\overline{F}^I(x-2\delta x(a)) \leq (1+o(1))\overline{F}^I(x)e^{\alpha g(\delta x(a))}.$$

Combining the latter result with (2.160), one gets

$$\frac{1}{a^2}\overline{F}^I(\delta x(a))\overline{F}^I(x-2\delta x(a)) \leq \frac{x(a)}{a^2}\overline{F}^I(x)e^{-(1-\alpha)g(\delta x(a))}.$$

From the latter we infer by using $g(u) \gg u^{\varepsilon_2}$ as $u \to \infty$ for some $\varepsilon_2 > 0$ and $x(a) \geq 1/a$ that

$$\frac{1}{a^2}\overline{F}^I(\delta x(a))\overline{F}^I(x - 2\delta x(a)) = o\left(\frac{1}{a}\overline{F}^I(x)\right), \quad x > k_\gamma \delta x(a). \tag{2.215}$$

Plugging the results from (2.213), (2.214) and (2.215) into (2.212), we get

$$\frac{1}{a}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) = o(e^{-\theta_a x}) + o\left(\frac{1}{a}\overline{F}^I(x)\right), \quad x > k_\gamma \delta x(a),$$

and, by additionally regarding (2.211), we obtain that, uniformly in x,

$$\frac{1}{a}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) = o(e^{-\theta_a x}) + o\left(\frac{1}{a}\overline{F}^I(x)\right).$$

Combining the latter relation with the results from (2.206), (2.207) and (2.208) gives

$$\mathbf{P}(M^{(a)} > x) \geq (1 + o(1))\left(e^{-\alpha}e^{-(\theta_a + c_a)x} + \frac{e^{-\alpha}L^{(N)}(x)}{a(1+\varepsilon)}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right).$$

Consequently, by letting $\varepsilon, \alpha \to 0$ and since $c_a x \to 0$ as $a \to 0$ for x such that $x \leq x(a) + C\ln(1/a)/\theta_a$ and $(1/a)\overline{F}^I(x) \gg e^{-\theta_a x}$ for all $x \geq x(a) + C\ln(1/a)/\theta_a$ with C sufficiently large, we obtain that, uniformly in x,

$$\mathbf{P}(M^{(a)} > x) \geq (1 + o(1))\left(e^{-\theta_a x} + \frac{L^{(N)}(x)}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right).$$

Hence, (2.203) and (2.204) give

$$\mathbf{P}(M^{(a)} > x) \geq (1 + o(1))\left(e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x > x(a) - C\ln(1/a)/\theta_a\}}{a(1 - \gamma g(x)/(\theta_a x))^2}\right)$$

in the case $F \in K_\gamma$ with $\gamma \in (0,1)$.

Now, consider $F \in K_0$. By Proposition 2.4, one can choose $\delta > 0$ arbitrary small such that $(Y_n^{(1)})$ is a non-negative supermartingale and proceeding like in the case $\gamma \in (0,1)$, one can show by using (2.114) instead of (2.194) that

$$\mathbf{P}(M^{(a)} > x) \leq (1 + o(1))\left(e^{-\theta_a x} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right). \tag{2.216}$$

In analogy to (2.114) one can show that $(1/a)\overline{F}^I(x) \gg e^{-\theta_a x}$ for all $x \geq x(a) + C\ln(1/a)/\theta_a$ with $C > 0$ suitable. On the other hand, (2.60) implies $x(a) + C\ln(1/a)/\theta = o(a^{-2})$ and therefore $\theta_a = 2a/\sigma^2 + O(a^2)$ gives

$$\mathbf{P}(M^{(a)} > x) \leq (1 + o(1))\left(e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right). \tag{2.217}$$

To establish the lower bound choose $\delta > 0$ small enough such that $(Y_n^{(2)})$ is a non-negative submartingale. Then, proceeding like in (2.206), (2.207) and (2.208), one obtains

$$e^{\alpha}\mathbf{P}(M^{(a)} > x) + \frac{1}{a}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) \\ \geq e^{-(\theta_a + c_a)x} + \frac{1}{a(1+\varepsilon)}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\} \tag{2.218}$$

and, exactly like in (2.211), one shows

$$\frac{1}{a}\overline{F}^I(\delta x(a))\mathbf{P}(M^{(a)} > x - 2\delta x(a)) = o(e^{-\theta_a x}), \quad x \leq 4\delta x(a). \tag{2.219}$$

On the other hand, in the case $x > 4\delta x(a)$, using (2.102), (2.50), (1.16) and the concavity of g, we get

$$\overline{F}^I(x - 2\delta x(a)) \sim \frac{\overline{F}(x - 2\delta x(a))}{(1-b)g'(x - 2\delta x(a))} \leq \frac{\overline{F}(x - 2\delta x(a))}{(1-b)g'(x)} \\ \leq \frac{\overline{F}(x)}{(1-b)g'(x)}e^{4\varepsilon_1 \delta x(a) g(x)/x} \leq \frac{\overline{F}(x)}{(1-b)g'(x)}e^{4\varepsilon_1 g(\delta x(a))} \sim \overline{F}^I(x)e^{4\varepsilon_1 g(\delta x(a))}$$

for all $\varepsilon_1 > 0$. On the other side, by (2.102), (2.12) and $g(x(a)) \geq (2+\varepsilon)\ln(x(a))$,

$$\overline{F}^I(\delta x(a)) \sim \frac{\overline{F}(\delta x(a))}{(1-b)g'(\delta x(a))} \leq \frac{x(a)\ln x(a)}{(1-b)g(x(a))}\overline{F}(\delta x(a)) \leq x(a)\overline{F}(\delta x(a)).$$

Combining the latter results and using $g(x(a)) \geq (2+\varepsilon)\ln x(a)$ once more, we attain

$$\frac{1}{a^2}\overline{F}^I(\delta x(a))\overline{F}^I(x - 2\delta x(a)) \leq (1+o(1))x(a)e^{-g(\delta x(a))(1-4\varepsilon_1)}\overline{F}^I(x) \\ \leq (1+o(1))\delta^{-(2+\varepsilon)(1-4\varepsilon_1)}\frac{1}{a^2}x(a)^{1-(2+\varepsilon)(1-4\varepsilon_1))}\overline{F}^I(x).$$

Since $x(a) \gg 1/a$, we infer from the latter that for $\varepsilon_1 > 0$ small enough

$$\frac{1}{a^2}\overline{F}^I(\delta x(a))\overline{F}^I(x - 2\delta x(a)) = o\left(\frac{1}{a}\overline{F}^I(x)\right), \quad x > 4\delta x(a), \tag{2.220}$$

and, by plugging the results from (2.219) and (2.220) into (2.218), we conclude

$$e^{\alpha}\mathbf{P}(M^{(a)} > x) \geq (1+o(1))\left(e^{-(\theta_a + c_a)x} + \frac{1}{a(1+\varepsilon)}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right).$$

By the same reasons as in the case $\gamma \in (0,1)$, we obtain from the latter that

$$\mathbf{P}(M^{(a)} > x) \geq (1+o(1))\left(e^{-\theta_a x} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right)$$

and, like in (2.217), one sees that this is the same as

$$\mathbf{P}(M^{(a)} > x) \geq (1+o(1))\left(e^{-2ax/\sigma^2} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\}\right).$$

2.3.7 Proof of Corollary 2.9

Since we use similar techniques as the proof of Propositions 2.4 and 2.5, we will sometimes refer to these proofs and not go into details in every aspect. During the whole proof we assume without loss of generality that a is sufficiently small.

One has

$$\int_{1/a}^{x} e^{-\theta_a(x-y)} dF^I(y) = e^{-\theta_a x} \int_{1/a}^{x} e^{\theta_a y} \overline{F}(y) dy. \tag{2.221}$$

First, let us consider x such that $1/a \le x \le x(a) - C\ln(1/a)/\theta_a$ with $x(a)$ defined in Lemma 2.3 and a constant $C > 0$ small enough in the case $\gamma = 0$ and C large enough in the case $\gamma > 0$. For a further discussion what small enough and large enough means see the proof of Proposition 2.4 and 2.5. The results (2.73) and (2.133) imply

$$\int_{1/a}^{x} e^{\theta_a y} \overline{F}(y) dy = o(a).$$

Furthermore,

$$\int_{1/a}^{x} (x-y) e^{-\theta_a(x-y)} \overline{F}(y) dy \le x(a) e^{-\theta_a x} \int_{1/a}^{x} e^{\theta_a y} \overline{F}(y) dy. \tag{2.222}$$

Since $c_a = o(1/x(a))$, the relations (2.73) and (2.133) give

$$x(a) \int_{1/a}^{x} e^{\theta_a y} \overline{F}(y) dy = o(1)$$

and hence,

$$\int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y) \right) e^{-\theta_a(x-y)} \overline{F}(y) dy = o(e^{-\theta_a x}) \tag{2.223}$$

for $1/a \le x \le x(a) - C\ln(1/a)/\theta_a$ with $C > 0$ as mentioned before.

Now, suppose $x > x(a) - C\ln(1/a)/\theta_a$. In this case we split the integral like in the proof of Propositions 2.4 and 2.5:

$$\int_{1/a}^{x} = \int_{1/a}^{x - C_1 \ln x/\theta_a} + \int_{x - C_1 \ln x/\theta_a}^{x}. \tag{2.224}$$

The results from (2.78), (2.79) and (2.139) give

$$\frac{e^{-\theta_a x}}{a} \int_{1/a}^{x - C_1 \ln x/\theta_a} e^{\theta_a y} \overline{F}(y) dy = o(e^{-\theta_a x}) + o\left(\frac{1}{a^2} \overline{F}(x) \right)$$

for $C_1 > 0$ small enough in the case $\gamma = 0$ and C_1 large enough in the case $\gamma > 0$. On the other side, one has the inequality

$$\int_{1/a}^{x - C_1 \ln x/\theta_a} (x-y) e^{\theta_a y} \overline{F}(y) dy \le x \int_{1/a}^{x - C_1 \ln x/\theta_a} e^{\theta_a y} \overline{F}(y) dy$$

and therefore one can show like in (2.78), (2.79) and (2.139) that

$$e^{-\theta_a x}\int_{1/a}^{x-C_1\ln x/\theta_a}(x-y)e^{\theta_a y}\overline{F}(y)dy = o(e^{-\theta_a x}) + o\left(\frac{1}{a^2}\overline{F}(x)\right) \tag{2.225}$$

for the same choice of C_1 and $x > x(a) - C\ln(1/a)/\theta_a$. The only difference here is that we have the term x in front of the integral. However this makes no difference if $x = O(x(a))$ and one can proceed like in (2.78), (2.79) and (2.139). If $x \gg x(a)$, one can show similar to (2.114) and (2.174) that $xe^{-\theta_a x} \ll \overline{F}^I(x)$ and therefore the proof also goes along the same line in this case.

To calculate the second integral from (2.224) we consider two cases. First let $\gamma > 0$. In this case, one can show like in (2.140) that for arbitrary $\varepsilon_1 > 0$

$$e^{-\theta_a x}\int_{x-C_1\ln x/\theta_a}^{x} e^{\theta_a y}\overline{F}(y)dy \le (1+o(1))\frac{\overline{F}(x)}{\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}. \tag{2.226}$$

Furthermore, by (2.52) and (2.127),

$$\begin{aligned} e^{-\theta_a x}\int_{x-C_1\ln x/\theta_a}^{x}(x-y)e^{\theta_a y}\overline{F}(y)dy &= \int_0^{C_1\ln x/\theta_a} we^{-\theta_a w - g(x-w)}dw \\ \le \overline{F}(x)\int_0^{C_1\ln x/\theta_a} we^{-w\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}dw \end{aligned}$$

for sufficiently large x and by integration by parts,

$$\begin{aligned} &\overline{F}(x)\int_0^{C_1\ln x/\theta_a} we^{-w\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}dw \\ &= -\frac{C_1 x^{-C_1(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}\ln x}{\theta_a^2(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}\overline{F}(x) \\ &\quad + \frac{\overline{F}(x)}{\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}\int_0^{C_1\ln x/\theta_a} e^{-w\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}dw. \end{aligned} \tag{2.227}$$

For all $\varepsilon_1 > 0$ such that $\gamma + \varepsilon_1 < 1$ one can easily see by regarding (2.137) that

$$\int_0^{C_1\ln x/\theta_a} e^{-w\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))}dw = \frac{1}{\theta_a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))} + o(1/\theta_a).$$

Because of (2.159), (2.14) and (2.137),

$$\frac{1}{a^2}\overline{F}(x) \sim \frac{1}{a^2}g'(x)\overline{F}^I(x) \le (\gamma+\varepsilon_1)\frac{g(x)}{a^2 x}\overline{F}^I(x) = O\left(\frac{1}{a}\overline{F}^I(x)\right)$$

for $x > x(a) - C_1\ln(1/a)/\theta_a$. We sum up the latter results and conclude

$$\begin{aligned} &\int_{x-C_1\ln x/\theta_a}^{x}\left(\frac{1}{a}+\frac{2}{\sigma^2}(x-y)\right)e^{-\theta_a(x-y)}\overline{F}(y)dy \\ &\le \left(\frac{(\theta_a/a)(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x)) + 2/\sigma^2}{\theta_a^2(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))^2}\right)(\gamma+\varepsilon_1)\frac{g(x)}{x}\overline{F}^I(x) + o\left(\frac{1}{a}\overline{F}^I(x)\right) \end{aligned} \tag{2.228}$$

and by plugging the results from (2.223), (2.225) and (2.228) into (2.221) and using $\theta_a \sim 2a/\sigma^2$,

$$\int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y)\right) e^{-\theta_a(x-y)} dF^I(y)$$
$$\leq \left(\frac{2-(\gamma+\varepsilon_1)g(x)/(\theta_a x)}{a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))^2}\right)(\gamma+\varepsilon_1)\frac{g(x)}{\theta_a x}\overline{F}^I(x) + o(e^{-\theta_a x}) + o\left(\frac{1}{a}\overline{F}^I(x)\right). \tag{2.229}$$

A straightforward calculation gives

$$\frac{1}{a}\overline{F}^I(x) + \left(\frac{2-(\gamma+\varepsilon_1)g(x)/(\theta_a x)}{a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))^2}\right)(\gamma+\varepsilon_1)\frac{g(x)}{\theta_a x}\overline{F}^I(x)$$
$$= \frac{\overline{F}^I(x)}{a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))^2}$$

and therefore

$$\frac{1}{a}\overline{F}^I(x) + \int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y)\right) e^{-\theta_a(x-y)} dF^I(y)$$
$$\leq \frac{\overline{F}^I(x)}{a(1-(\gamma+\varepsilon_1)g(x)/(\theta_a x))^2} + o(e^{-\theta_a x}) + o\left(\frac{1}{a}\overline{F}^I(x)\right). \tag{2.230}$$

Since this bound is valid for arbitrary $\varepsilon_1 > 0$ one can let $\varepsilon_1 \to 0$ and thus we get by regarding (2.203) that (2.30) implies

$$\mathbf{P}(M^{(a)} > x) \leq (1+o(1))\left(e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x \geq x(a) - C\ln(1/a)/\theta_a\}}{a(1-\gamma g(x)/(\theta_a x))^2}\right)$$

with a sufficiently large constant $C > 0$.

It remains to show that the term on the right hand side of the latter inequality is also a asymptotical lower bound. Recall that for all $\varepsilon_1 > 0$ such that $\gamma + \varepsilon_1 < 1$ inequality (2.137) gives

$$\frac{C_1 x^{-C_1(1-(\gamma+\varepsilon_2)g(x)/(\theta_a x))}\ln x}{\theta_a^2(1-(\gamma+\varepsilon_2)g(x)/(\theta_a x))} = o(1/\theta_a^2)$$

for $x > x(a) - C\ln x/\theta_a$. Then, following the same line, one can show similar to (2.230) by using $g'(x) \geq (\gamma - \varepsilon_1)g(x)/x$ instead of $g'(x) \leq (\gamma+\varepsilon_1)g(x)/x$ that

$$\frac{1}{a}\overline{F}^I(x) + \int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y)\right) e^{-\theta_a(x-y)} dF^I(y)$$
$$\geq \frac{\overline{F}^I(x)}{a(1-(\gamma-\varepsilon_1)g(x)/(\theta_a x))^2} + o(e^{-\theta_a x}) + o\left(\frac{1}{a}\overline{F}^I(x)\right).$$

Hence, by plugging this into (2.30) and regarding (2.203) and $x(a) \geq 1/a$,

$$\mathbf{P}(M^{(a)} > x) \geq (1+o(1)) \left(e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x \geq x(a) - C\ln(1/a)/\theta_a\}}{a(1-(\gamma-\varepsilon_1)g(x)/(\theta_a x))^2} \right)$$

and, since $\varepsilon_1 > 0$ can be chosen arbitrary small,

$$\mathbf{P}(M^{(a)} > x) \geq (1+o(1)) \left(e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x \geq x(a) - C\ln(1/a)/\theta_a\}}{a(1-\gamma g(x)/(\theta_a x))^2} \right).$$

Finally, this means that the result from Blanchet and Lam from (2.30) states that, in the special case of a $M/G/1$ queue,

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} + \frac{\overline{F}^I(x)\mathbf{1}\{x \geq x(a) - C\ln(1/a)/\theta_a\}}{a(1-\gamma g(x)/(\theta_a x))^2} \quad \text{as } a \to 0$$

if $\gamma > 0$.

Let us consider the case $\gamma = 0$. One can show similar to (2.226) by using (2.50) instead of (2.127) that for an arbitrary $\varepsilon_1 > 0$ and C_1 small enough,

$$e^{-\theta_a x} \int_{x-C_1 \ln x/\theta_a}^{x} e^{\theta_a y}\overline{F}(y)dy \leq (1+o(1)) \frac{\overline{F}(x)}{\theta_a(1-\varepsilon_1 g(x)/(\theta_a x))}. \tag{2.231}$$

Using (2.13) and proceeding like in (2.229), we get

$$\begin{aligned} &\int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y)\right) e^{-\theta_a(x-y)} dF^I(y) \\ &\leq \left(\frac{2-\varepsilon_1 g(x)/(\theta_a x)}{a(1-\varepsilon_1 g(x)/(\theta_a x))^2} \right) \varepsilon_1 \frac{g(x)}{\theta_a x} \overline{F}^I(x) + o(e^{-\theta_a x}) + o\left(\frac{1}{a}\overline{F}^I(x)\right). \end{aligned}$$

By regarding $g(x)/(\theta_a x) \leq C_2$ with C_2 large enough, we conclude that

$$\int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y)\right) e^{-\theta_a(x-y)} dF^I(y) \leq \frac{\varepsilon_1 C_3}{a} \overline{F}^I(x) \tag{2.232}$$

with a suitable constant C_3. Since ε_1 was arbitrary, we let $\varepsilon_1 \to 0$ and by plugging the latter inequality into (2.30) one obtains by regarding (2.223) that

$$\mathbf{P}(M^{(a)} > x) \leq (1+o(1)) \left(e^{-\theta_a x} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq x(a) - C\ln(1/a)/\theta_a\} \right).$$

Since C can be chosen arbitrary small in the latter inequality, (2.48) implies $\delta x(a) < x(a) - C\ln(1/a)/\theta_a$ and therefore the latter inequality is equivalent to

$$\mathbf{P}(M^{(a)} > x) \leq (1+o(1)) \left(e^{-\theta_a x} + \frac{1}{a}\overline{F}^I(x)\mathbf{1}\{x \geq \delta x(a)\} \right).$$

On the other side,

$$\int_{1/a}^{x} \left(\frac{1}{a} + \frac{2}{\sigma^2}(x-y) \right) e^{-\theta_a (x-y)} dF^I(y) \geq 0$$

and due to $x(a) \geq 1/a$ we infer from the latter and (2.30) that

$$\mathbf{P}(M^{(a)} > x) \geq (1 + o(1)) \left(e^{-\theta_a x} + \frac{1}{a} \overline{F}^I(x) \mathbf{1}\{x \geq \delta x(a)\} \right).$$

By combining the latter results, we restate our result in the case $\gamma = 0$ in the special case of a $M/G/1$ queue:

$$\mathbf{P}(M^{(a)} > x) \sim e^{-\theta_a x} + \frac{1}{a} \overline{F}^I(x) \mathbf{1}\{x \geq \delta x(a)\}.$$

3 Upper bounds for the maximum of a random walk with negative drift

The results in this chapter mostly arose from joint work with Dr. Vitali Wachtel and are mainly from the paper [36]. However, there are also results which I produced independently - like Theorem 3.11 - and which have not been published yet. Theorem 3.9 has been published in a weaker form, namely there was a supplementary factor e^r on the right hand side of condition (3.19), so Theorem 3.9 stated here holds for a wider range of values x.

3.1 Introduction, motivation and description of the method

We use the notation introduced in Chapter 1.7 and since we will only consider fixed expectation $-a$ we will omit the superscript $^{(a)}$ everywhere.

In this chapter again the tail distribution of M is the object of interest. Since the exact form of the distribution of the maximum of this random walk is known in some special cases only, good estimates are required. In the literature, the tail of M is usually approximated by it's asymptotic form (see (1.28), (1.30) and (1.31) in the light-tailed case and (1.29), (1.30) and (1.31) in the heavy tailed case). However this is not necessarily a good approximation, for a discussion and numerical results on the accuracy see for example Kalashnikov [31]. Therefore it is of great interest to have non-asymptotic properties of the distribution of M.

In the light-tailed case, the first result goes back, apparently, to Cramér and Lundberg (see Chapter 1.8): If

$$\mathbf{E}[e^{h_0 X}] = 1 \quad \text{for some } h_0 > 0, \tag{3.1}$$

one has for all $x > 0$ the so-called Lundberg inequality

$$\mathbf{P}(M > x) \le e^{-h_0 x}. \tag{3.2}$$

It is known that if (3.1) holds and, in addition, $\mathbf{E}[Xe^{h_0 X}] < \infty$, then there exists a constant $c_0 \in (0, 1)$ such that

$$\mathbf{P}(M > x) \sim c_0 e^{-h_0 x} \quad \text{as } x \to \infty. \tag{3.3}$$

This means that (3.2) has optimal order and the error of the Lundberg inequality is only a constant.

In the case when (3.1) is not fulfilled, upper bounds for $\mathbf{P}(M > x)$ have been derived by Kalashnikov [31] and by Richards [46]. The approach in these papers is based on the representation of M as a geometric sum of independent random variables:

$$\mathbf{P}(M > x) = \sum_{k=0}^{\infty} q(1-q)^k \mathbf{P}(\chi_1^+ + \chi_2^+ \ldots + \chi_k^+ > x), \tag{3.4}$$

where $\{\chi_k^+\}$ are independent random variables and $q = \mathbf{P}(M = 0)$. The main difficulty in this approach is the fact that one has to know the distribution of χ_k^+ and the parameter

q. In some special cases this information can be obtained from the initial data. But in general one has to obtain appropriate estimates for q and $\mathbf{P}(\chi_1^+ > x)$. Bounds for $\mathbf{P}(\chi_1^+ > x)$ are given, for example, in Chapter 4 of [8].

The main purpose of the present chapter is to derive good upper bounds for $\mathbf{P}(M > x)$ if the Cramér-Lundberg condition does not hold. Therefore, we will assume the existence of power moments of X only, avoid the representation via geometric sum and use a supermartingale-construction instead. The most important advantage of this method is that it will give specific upper bounds without any unknown factors like q and $\mathbf{P}(\chi_1^+ > x)$ in the bounds of Kalashnikov [31] and Richards [46]. As it is usual for deriving upper bounds, we are going to truncate summands and to use inequalities, which are based on truncated exponential moments. But the problem is that we have infinitely many random variables X_i's, so we can not truncate all of them at the same level. Thus, we have to split the time axis into intervals of finite length and then choose a level of truncation on each of these intervals. One can take, for example, a deterministic strictly increasing sequence (k_n) with $k_0 = 0$ and consider the intervals $I_n := (k_n, k_{n+1}]$:

$$\begin{aligned}\mathbf{P}(M \geq x) = \mathbf{P}\Big(\bigcup_{k\geq 0}\{S_k \geq x\}\Big) &\leq \sum_{n=0}^{\infty}\mathbf{P}\left(\bigcup_{k\in I_n}\{S_k \geq x\}\right)\\ &\leq \sum_{n=0}^{\infty}\mathbf{P}\left(\max_{k\leq k_{n+1}}(S_k - ka) \geq x - k_n a\}\right). \qquad (3.5)\end{aligned}$$

Now, one can apply the Fuk-Nagaev inequalities, see [42], to every probability in the last line. It is clear that replacing $\sup_{k\in I_n}(S_k - ka)$ by $\sup_{k\leq k_{n+1}}(S_k - ka)$ is not too rough if and only if k_{n+1} and $k_{n+1} - k_n$ are comparable. Thus, one has to take k_n exponentially growing. Using this approach with $k_n = x2^n$, Borovkov [10] obtained a version of the Markov inequality for M.

Our strategy is quite different and consists in splitting $[0,\infty)$ into random intervals defined by a sequence of stopping times. More precisely, we introduce the stopping time

$$\tau_z := \min\{k \geq 0 : S_k \leq -z\}, \quad z \geq 0.$$

Let $M_{\tau_z} = \max_{1\leq k\leq \tau_z} S_k$. We split the tail probability

$$\mathbf{P}(M > x) \leq \mathbf{P}(M_{\tau_z} > x) + \mathbf{P}\left(\max_{k\geq\tau_z} S_k > x\right) \qquad (3.6)$$

and consider the continuation of the process (S_k) beyond τ_z as a probabilistic replica of the entire process. By $S_{\tau_z} \leq -z$ follows

$$\mathbf{P}\left(\max_{k\geq\tau_z} S_k > x\right) \leq \mathbf{P}(M > x + z).$$

As a result, we have

$$\mathbf{P}(M > x) \leq \mathbf{P}(M_{\tau_z} > x) + \mathbf{P}(M > x + z),$$

and inductively we conclude

$$\mathbf{P}(M > x) \leq \sum_{j=0}^{\infty} \mathbf{P}(M_{\tau_z} > x + jz). \tag{3.7}$$

It is worth mentioning that the difference between (3.5) and (3.7) is the same as between Riemann and Lebesgue integrals: We do not fit the random walk S_n into a fixed splitting of the time, but choose the splitting depending on the paths of the random walk.

A decomposition similar to (3.6) has been used by Denisov [18] for deriving the asymptotics of $\mathbf{P}(M_{\tau_0} > x)$ from that of $\mathbf{P}(M \in [x, x - S_{\tau_0}))$. In the present chapter we use the opposite approach: We obtain estimates for $\mathbf{P}(M > x)$ from the ones for $\mathbf{P}(M_{\tau_z} > x)$ and the estimates for $\mathbf{P}(M_{\tau_z} > x)$ are derived by using a martingale construction similar to the one used in (1.34)

3.2 Upper bounds for $\mathbf{P}(M_{\tau_z} > x)$ and $\mathbf{P}(M > x)$

We first state our results on M_{τ_z}.

Theorem 3.1. *Assume that $A_t := \mathbf{E}[|X|^t] < \infty$ for some $t \in (1, 2]$. For all y satisfying $y^{t-1} \geq (e-1)A_t a^{-1}$ we have the following inequality:*

$$\mathbf{P}(M_{\tau_z} > x) \leq \frac{A_t^{x/y}}{a^{x/y-1}} \mathbf{E}[\tau_z] y^{-1-(t-1)x/y} \ln\left(1 + a y^{t-1}/A_t\right) + \left(1 + \frac{A_t^{x/y}}{a^{x/y}} y^{-(t-1)x/y}\right) \mathbf{E}[\tau_z] \mathbf{P}(X > y). \tag{3.8}$$

Remark 3.2. We show in the proof that (3.8) remains true, if one replaces a and A_t by $-\mathbf{E}[X; |X| \leq y]$ and $A_t(y) = \mathbf{E}[|X|^t; |X| \leq y]$ respectively. In this case the restriction $y^{t-1} > (e-1)a^{-1}A_t$ should be replaced by $\mathbf{E}[X; |X| \leq y] < 0$. The use of truncated moments is more convenient in theoretical applications, but for deriving concrete estimates for M it is easier to use full moments. ◇

Let us turn to the case $t > 2$. Fix $\varepsilon > 0$, $\alpha \in (0, 1)$ and put $\beta = 1 - \alpha$. Assume $\mathbf{P}(X > 0) > 0$. We use the notation

$$A_{t,+} := \mathbf{E}[X^t; X > 0], \quad \psi_1(x) := \exp\left(\frac{2\alpha a x}{e^{\varepsilon}\mathbf{E}[X^2]}\right) - 1, \quad \psi_2 := \frac{\beta a}{A_{t,+}}.$$

Theorem 3.3. *Assume $\sigma^2 = \mathbf{Var}(X) < \infty$ and $A_{t,+} < \infty$ for some $t > 2$.*

(i) If y satisfies the condition

$$\frac{2\alpha a}{e^{\varepsilon}\mathbf{E}[X^2]} \leq \frac{1}{y} \ln\left(1 + \frac{\beta a}{A_{t,+}} y^{t-1}\right), \tag{3.9}$$

then

$$\mathbf{P}(M_{\tau_z} > x) \leq \left(1 + \frac{1}{\psi_1(x)}\right) \mathbf{E}[\tau_z] \mathbf{P}(X > y) + \frac{2\alpha a^2 \mathbf{E}[\tau_z]}{e^{\varepsilon}\mathbf{E}[X^2]\psi_1(x)}. \tag{3.10}$$

(ii) If y satisfies the condition

$$\frac{2\alpha a}{e^{\varepsilon}\mathbf{E}[X^2]} \geq \frac{1}{y}\ln\left(1+\frac{\beta a}{A_{t,+}}y^{t-1}\right), \tag{3.11}$$

then

$$\mathbf{P}\left(M_{\tau_z} > x\right) \leq \psi_2^{-x/y}\mathbf{E}[\tau_z]ay^{-1-(t-1)x/y}\ln\left(1+\psi_2 y^{t-1}\right) + \left(1+\psi_2^{-x/y}y^{-(t-1)x/y}\right)\mathbf{E}[\tau_z]\mathbf{P}(X > y). \tag{3.12}$$

Remark 3.4. Analogously to Theorem 3.1 one can replace $\mathbf{E}[X^2]$ and $A_{t,+}$ by the corresponding truncated expectations $B^2(-\infty, y) = \mathbf{E}[X^2; X \leq y]$ and $A_{t,+}(y) = \mathbf{E}[X^t; X \in (0, y]]$ respectively. ◇

Corollary 3.5. *Assume that $\mathbf{P}(|X| > x) = L(x)x^{-r}$ for some $r > 1$ and*

$$\mathbf{P}(X > x))/\mathbf{P}(|X| > x) \to p \in (0,1) \quad \text{as } x \to \infty.$$

Then, it follows from (3.8) and (3.12) that

$$\limsup_{x\to\infty} \frac{\mathbf{P}\left(M_{\tau_z} > x\right)}{\mathbf{P}(X > x)} \leq \mathbf{E}[\tau_z]$$

for every $z > 0$.

But it follows from the results of Asmussen [4] (see also Denisov [18] and Foss, Zachary [29]), that

$$\lim_{x\to\infty} \frac{\mathbf{P}\left(M_{\tau_z} > x\right)}{\mathbf{P}(X > x)} = \mathbf{E}[\tau_z]$$

under the condition that the tail of F is regular varying. This means that the inequalities (3.8) and (3.12) are asymptotically precise in the case of regular varying tails.

In all these inequalities we have $\mathbf{E}[\tau_z]$ on the right hand side. It is really hard to get an exact expression for this value via initial data, but there are good upper bounds in the literature: Since $\mathbf{E}[\tau_z] < \infty$ (see, for example, Feller [25]) Wald's identity gives

$$\mathbf{E}[\tau_z] = \frac{z + \mathbf{E}[R_z]}{a}, \tag{3.13}$$

where $R_z = -z - S_{\tau_z}$ denotes the overshoot in τ_z. Hence, we get upper bounds for $\mathbf{E}[\tau_z]$ by the inequality of Lorden (see [38]): If $\mathbf{E}[X] < 0$ and $\mathbf{E}[(X^-)^2] < \infty$,

$$\mathbf{E}[R_z] \leq \frac{\mathbf{E}[(X^-)^2]}{a} \tag{3.14}$$

and the one from Mogul'skii [40]: If $\mathbf{E}[X] \leq 0$ and $\mathbf{E}[|X|^3] < \infty$,

$$\mathbf{E}[R_z] \leq A\frac{3}{2}\frac{\mathbf{E}[|X|^3]}{\mathbf{E}[X^2]}, \tag{3.15}$$

where A is a certain constant, $A \leq 2$. The disadvantage of these bounds is, that we have to assume the existence of the second or even the third moment. We give another bound, which only requires the finiteness of the moment of order t, $t \in (1, 2]$.

Proposition 3.6. *Assume that* $A_{t,-} := \mathbf{E}[(X^-)^t] < \infty$ *for some* $t \in (1,2]$*, then, for every* $z > 0$*,*

$$\mathbf{E}[R_z] \le \frac{t^{t/(t-1)} A_{t,-}^{1/(t-1)}}{(t-1)a^{t/(t-1)}} \left(\mathbf{E}[-X; X < 0] + \frac{z^{2-t}}{t} A_{t,-} \right). \tag{3.16}$$

Combining (3.13) with (3.14), (3.15) or (3.16) we obtain upper bounds for $\mathbf{E}[\tau_z]$. Plugging these bounds into the inequalities in Theorems 3.1 and 3.3 we get bounds for $\mathbf{P}(M_{\tau_z} > x)$, which contain information on X only. So, they can be used for concrete calculations.

We come back to the global maximum. The results in this section follow from the results on M_{τ_z} via the formula (3.7) we attained through the random time splitting.

Theorem 3.7. *Fix some* $\theta \in (0,1)$ *and define*

$$c_1 := \frac{3A_t^{1/\theta}\theta^{-(t-1)/\theta}}{(t-1)a^{1/\theta-1}}, \quad c_2 := \frac{3A_{t,+}^{1/\theta}\theta^{-(t-1)/\theta}}{(t-1)\beta^{1/\theta}a^{1/\theta-1}},$$

$$\psi_3(x) := \frac{a\theta^{t-1}x^{t-1}}{A_t}, \quad \psi_4(x) := \frac{\beta a\theta^{t-1}x^{t-1}}{A_{t,+}}.$$

(i) Assume that $A_t < \infty$ *for some* $t \in (1,2]$*. Then, for every* x *satisfying* $x^{t-1} \ge \theta^{1-t}(e^\theta - 1)A_t a^{-1}$ *and* $x \ge z(t-1)\theta^{-1}$*, we have*

$$\mathbf{P}(M > x) \le c_1 \frac{\mathbf{E}[\tau_z]}{z} \ln(1 + \psi_3(x))\, x^{-(t-1)/\theta} + \left(1 + \psi_3(x)^{-1/\theta}\right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z}\overline{F}^I(\theta x) + \mathbf{P}(X > \theta x) \right). \tag{3.17}$$

(ii) Assume that $\mathbf{Var}(X) < \infty$ *and* $A_{t,+} < \infty$ *for some* $t > 2$*. Then, for every* x *satisfying (3.11) for* $y = \theta x$ *and the conditions* $x^{t-1} \ge \theta^{1-t}(e^\theta - 1)A_{t,+}\beta^{-1}a^{-1}$ *and* $x \ge z(t-1)\theta^{-1}$*, we have*

$$\mathbf{P}(M > x) \le c_2 \frac{\mathbf{E}[\tau_z]}{z} \ln(1 + \psi_4(x))\, x^{-(t-1)/\theta} + \left(1 + \psi_4(x)^{-1/\theta}\right) \mathbf{E}[\tau_z] \left(\frac{1}{\theta z}\overline{F}^I(\theta x) + \mathbf{P}(X > \theta x) \right). \tag{3.18}$$

Corollary 3.8. *If the assumptions of Corollary 3.5 hold, then it follows from Theorem 3.7 that*

$$\limsup_{x\to\infty} \frac{\mathbf{P}(M > x)}{\overline{F}^I(x)} \le \frac{\mathbf{E}[\tau_z]}{z}\theta^{-r}.$$

Since the left-hand side does not depend on θ and z, we can let $\theta \to 1$ and $z \to \infty$. Noting that each of (3.14) and (3.15) combined with (3.13) yields

$$\frac{\mathbf{E}[\tau_z]}{z} \to \frac{1}{a} \quad \text{as } z \to \infty,$$

we conclude

$$\limsup_{x\to\infty} \frac{\mathbf{P}(M > x)}{\overline{F}^I(x)} \le \frac{1}{a}.$$

Comparing this with (0.2), we see that the inequalities in Theorem 3.7 are asymptotically precise. This even remains valid, if we bound $\mathbf{E}[\tau_z]$ in the inequalities of Theorem 3.7 by combining (3.14) or (3.15) with (3.13).

The reason why we are able to obtain asymptotically precise bounds is, that we may choose z arbitrary large. That possibility seems to be a quite important advantage of our method compared to geometric sums. If the distribution of χ_1^+ is subexponential, then it follows easily from (3.4) that

$$\mathbf{P}(M > x) \sim \left(\frac{1}{q} - 1\right) \mathbf{P}(\chi_1^+ > x) \quad \text{as } x \to \infty.$$

Therefore, in order to obtain an upper bound for the maximum we need to control the quantity $1/q$. It is well known that $1/q = \mathbf{E}[-S_{\tau_0}] = \mathbf{E}[R_0]$. Thus, we may apply (3.14), (3.15) or (3.16) with $z = 0$. But corresponding inequalities for M will not be asymptotically precise. Summarizing, the approach via geometric sums can only lead to asymptotically precise bounds if q is known.

3.3 Asymptotic implications of the bounds from chapter 3.2

In this section we test our inequalities in the heavy traffic and heavy tail regimes. Consider a family of random walks $\{S^{(a)}, a \ge 0\}$ with $\mathbf{E}[X^{(a)}] = -a$ and consider x depending on a. We shall assume that $X^{(a)} = X^{(0)} - a$ for all $a > 0$. Let $M^{(a)}$ denote the corresponding maximum. It is known that if $X^{(0)}$ belongs to the domain of attraction of a stable law, then there exists a regular varying function $g(a)$ such that $g(a)M^{(a)}$ converges weakly as $a \to 0$. It turns out that our inequalities may be applied to large deviation problems in the heavy traffic convergence mentioned above. More precisely, they give asymptotically precise bounds for the probabilities $\mathbf{P}(M^{(a)} > x)$ if $x \gg 1/g(a)$. In the case of $\sigma^2 := \mathbf{Var}(X^{(0)})$ being finite, one has $g(a) = a$ and the weak limit of $aM^{(a)}$ is the exponential distribution with parameter $2/\sigma^2$.

Theorem 3.9. *Assume that $\sigma^2 < \infty$ and the right tail of the distribution function of $X^{(0)}$ is regular varying with index $r > 2$, that is, $\mathbf{P}(X^{(0)} > u) = u^{-r}L(u)$, where L is slowly varying. If*

$$\liminf_{a\to 0} \frac{x}{a^{-1}\ln a^{-1}} > \frac{(r-2)}{2}\sigma^2, \tag{3.19}$$

then

$$\mathbf{P}(M^{(a)} > x) \sim \frac{x^{-r+1}L(x)}{(r-1)a} \quad \text{as } a \to 0. \tag{3.20}$$

Theorem 3.10. *Assume that $\mathbf{E}[(\min\{0, X^{(0)}\})^2] < \infty$ and $\mathbf{P}(X^{(0)} > u) = u^{-r}L(u)$ with $r \in (1, 2)$. If*

$$\liminf_{a\to 0} g(a)x = \infty, \tag{3.21}$$

then

$$\mathbf{P}(M^{(a)} > x) \sim \frac{x^{-r+1}L(x)}{(r-1)a} \quad \text{as } a \to 0. \tag{3.22}$$

We have imposed the condition $\mathbf{E}[(\min\{0, X^{(0)}\})^2] < \infty$ just to use the Lorden inequality for the overshoot. If one replaces that condition by $\mathbf{E}[|\min\{0, X^{(0)}\}|^t] < \infty$ with $t \in (1, 2)$, then, using Proposition 3.6, one can show that (3.22) holds for $x \gg a^{-t/(t-1)^2}$ only. The reason is the roughness of Proposition 3.6 for small values of a. Indeed, if we use (3.16) even with $t = 2$, we get the bound $\mathbf{E}[R_z] \leq Ca^{-2}$, which is much worse then the Lorden inequality.

Theorem 3.11. *Assume that* $\mathbf{E}[|X^{(0)}|^r] < \infty$ *with* $r > 2$. *If*

$$\limsup_{a \to 0} \frac{x}{a^{-1} \ln a^{-1}} < \frac{(r-2)}{2} \sigma^2, \tag{3.23}$$

then

$$\mathbf{P}(M^{(a)} > x) \sim e^{-2ax/\sigma^2} \quad \text{as } a \to 0. \tag{3.24}$$

This result generalizes a result from Olvera-Cravioto, Blanchet and Glynn [43] (see also Blanchet and Lam [7]). They have shown in the setting of a $M/G/1$ queue with regular varying processing time distribution, that there exists some critical value

$$x_{RV}(a) \approx \frac{(r-2)\sigma^2}{2} \frac{1}{a} \ln \frac{1}{a},$$

under which the heavy traffic approximation (3.24) and above which the heavy tail asymptotic (3.20) holds. However, our results correspond to the more general case of a $G/G/1$ queue with regular varying processing times in which the arrivals don't have to be exponential, but only independent of each other. Furthermore, in contrast to Theorem 3.10, Olvera-Cravioto, Blanchet and Glynn [43] only consider the case $r > 2$ with finite variance.

Remark 3.12. Theorems 3.9 and 3.11 are special cases of Theorem 2.7, see Example 1 in chapter 2.2. In the present chapter, the asymptotical results were derived using the computable inequalities. As shown in Theorems 3.9 and 3.11 these inequalities are asymptotically precise and serve well to attain asymptotical results in the case of regular varying distributions. However, it turned out that our inequalities do not seem to give precise asymptotics in the Weibull-case. It seems the overshoot plays a more important role in this case and the bound $\mathbf{E}[e^{hS_{\tau_z}}] \geq 1 + h\mathbf{E}[S_{\tau_z}]$ used in the proof of Theorem 3.1 seems to be too rough.

3.4 Proofs

3.4.1 Proofs of Theorems 3.1 and 3.3

We set for brevity $\tau = \tau_z$.

Lemma 3.13. *For all h satisfying*

$$\mathbf{E}[e^{hX}; X \le y] \le 1 \tag{3.25}$$

we have the inequality

$$\mathbf{P}(M_\tau > x) \le \Big(1 + \frac{1}{e^{hx}-1}\Big)\mathbf{E}[\tau]\mathbf{P}(X > y) + \mathbf{E}[\tau]\frac{ah}{e^{hx}-1}. \tag{3.26}$$

Beweis. Our strategy is to truncate the random variables X_i in the level y:

$$\begin{aligned}\mathbf{P}\left(M_\tau > x\right) &\le \mathbf{P}\left(M_\tau > x, \max_{1\le k\le\tau} X_k \le y\right) + \mathbf{P}\left(\max_{1\le k\le\tau} X_k > y\right) \\ &= \mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le\tau} X_k \le y\right\} > x\right) + \mathbf{P}\left(\max_{1\le k\le\tau} X_k > y\right).\end{aligned} \tag{3.27}$$

From the Wald identity follows

$$\mathbf{P}\left(\max_{1\le k\le\tau} X_k > y\right) \le \mathbf{E}\left[\sum_{k=1}^{\tau} \mathbf{1}\{X_k > y\}\right] = \mathbf{E}\left[\tau\right]\mathbf{P}\left(X > y\right). \tag{3.28}$$

To examine the first term on the right-hand-side of (3.27) we introduce the process $\{W_k\}$ defined by

$$W_0 := 1, \quad W_k := \prod_{i=1}^{k} e^{hX_i}\mathbf{1}\{X_i \le y\},\ k \ge 1.$$

It is clear that if h satisfies (3.25), $\{W_k\}$ is a positive supermartingale. Define

$$\sigma_y := \min\{k \ge 1 : X_k > y\},\ t_x := \min\{k \ge 1 : S_k > x\} \text{ and } T := \min\{\sigma_y, t_x, \tau\}.$$

Applying the Optional Stopping Theorem to the supermartingale $\{W_{k\wedge T}\}$ gives us

$$1 = W_0 \ge \mathbf{E}[W_T] = \mathbf{E}\left[W_T; t_x < \tau, t_x < \sigma_y\right] + \mathbf{E}\left[W_T; \tau < t_x, \tau < \sigma_y\right].$$

We analyze the two terms on the right-hand-side separately:

$$\mathbf{E}\left[W_T; t_x < \tau, t_x < \sigma_y\right] \ge e^{hx}\mathbf{P}(t_x < \tau < \sigma_y) = e^{hx}\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le\tau} X_k \le y\right\} > x\right)$$

and

$$\begin{aligned}\mathbf{E}\left[W_T; \tau < t_x, \tau < \sigma_y\right] &= \mathbf{E}\left[e^{hS_\tau}\right] - \mathbf{E}\left[e^{hS_\tau}; \{M_\tau > x\} \cup \left\{\max_{1\le k\le\tau} X_k > y\right\}\right] \\ &\ge \mathbf{E}\left[e^{hS_\tau}\right] - e^{-hz}\left(\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le\tau} X_k \le y\right\} > x\right) + \mathbf{P}\left(\max_{1\le k\le\tau} X_k > y\right)\right).\end{aligned}$$

Consequently,

$$\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le\tau} X_k \le y\right\} > x\right) \le \frac{1 - \mathbf{E}\left[e^{hS_\tau}\right] + \mathbf{P}\left(\max_{1\le k\le\tau} X_k > y\right)}{e^{hx} - 1}$$

and hence by applying (3.28),

$$\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le \tau} X_k \le y\right\} > x\right) \le \frac{1-\mathbf{E}\left[e^{hS_\tau}\right]+\mathbf{E}[\tau]\mathbf{P}\left(X>y\right)}{e^{hx}-1}.$$

It is easy to see that

$$\mathbf{E}\left[e^{hS_\tau}\right] \ge \mathbf{E}\left[1+hS_\tau\right] = 1+h\mathbf{E}[S_\tau]$$

and as a result we have

$$\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le \tau} X_k \le y\right\} > x\right) \le \mathbf{E}[\tau]\frac{ah+\mathbf{P}\left(X>y\right)}{e^{hx}-1}. \tag{3.29}$$

Applying (3.28) and (3.29) to the summands in (3.27) finishes the proof. □

To prove Theorems 3.1 and 3.3 we need to choose a specific h for which (3.25) holds. The optimal choice would be the positive solution of the equation $\mathbf{E}[e^{hX}; X \le y] = 1$, which is in the spirit of the Cramér-Lundberg condition. But it is not clear how to solve this equation. For this reason we replace $\mathbf{E}[e^{hX}; X \le y] = 1$ by the equation $\phi(h, y) = 1$, where $\phi(h, y)$ is an appropriate upper bound for $\mathbf{E}[e^{hX}; X \le y]$.

If $A_t < \infty$, we may use a bound from the proof of Theorem 2 from [30], which says

$$\mathbf{E}[e^{hX}; X \le y] \le 1 + h\mathbf{E}[X; |X| \le y] + \frac{e^{hy}-1-hy}{y^t}A_t. \tag{3.30}$$

Using the Markov inequality we also obtain

$$\mathbf{E}[X; |X| \le y] \le -a - \mathbf{E}[X; X \le -y] \le -a + \frac{A_t}{y^{t-1}},$$

and therefore

$$\mathbf{E}[e^{hX}; X \le y] \le 1 - ha + \frac{e^{hy}-1}{y^t}A_t.$$

Put $h_0 := \frac{1}{y}\ln\left(1+ay^{t-1}/A_t\right)$. It is easy to see that

$$-h_0 a + \frac{e^{h_0 y}-1}{y^t}A_t \le 0$$

for all y such that $y^{t-1} \ge (e-1)A_t a^{-1}$ and this implies that h_0 satisfies (3.25). Using (3.26) with $h = h_0$ and applying the inequality

$$(1+u)^{x/y} \ge 1 + u^{x/y},\ x \ge y,$$

we obtain

$$\begin{aligned}\mathbf{P}(M_\tau > x) \le {}& \frac{A_t^{x/y}}{a^{x/y-1}}\mathbf{E}[\tau]y^{-1-(t-1)x/y}\ln\left(1+ay^{t-1}/A_t\right)\\ &+\left(1+\frac{A_t^{x/y}}{a^{x/y}}y^{-(t-1)x/y}\right)\mathbf{E}[\tau]\mathbf{P}(X>y).\end{aligned}$$

Thus, the proof of Theorem 3.1 is complete.

In order to show that one can replace $\mathbf{E}[X]$ and A_t by the corresponding truncated moments, see Remark 3.2, we first note that analogously to (3.30) and by additionally using $e^x - 1 \le xe^x$,

$$\mathbf{E}[e^{hX}; X \le y] \le 1 + h\mathbf{E}[X; |X| \le y] + h\frac{e^{hy}-1}{y^{t-1}}\mathbf{E}[|X|^t; |X| \le y].$$

If $\mathbf{E}[X; |X| \le y] < 0$, then

$$h_0 := \frac{1}{y}\ln\left(1 + \frac{|\mathbf{E}[X; |X| \le y]|y^{t-1}}{\mathbf{E}[|X|^t; |X| \le y]}\right)$$

is strictly positive and solves

$$h\mathbf{E}[X; |X| \le y] + h\frac{e^{hy}-1}{y^{t-1}}\mathbf{E}[|X|^t; |X| \le y] = 0.$$

Therefore, we may use Lemma 3.13 with $h = h_0$ and get an inequality with truncated moments.

To proof Theorem 3.3, we want to apply Lemma 3.13 again and therefore need to bound $\mathbf{E}[e^{hX}; X \le y]$ under the conditions of Theorem 3.3. We proceed similar to the proof of Theorem 3 from [42] and get

$$\mathbf{E}[e^{hX}; X \le y] \le 1 - ha + \mathbf{E}[e^{hX} - 1 - hX; X \le \epsilon/h] + \mathbf{E}[e^{hX} - 1 - hX; X \in (\epsilon/h, y]]. \quad (3.31)$$

We consider the last two terms of this inequality separately. As you can easily see,

$$\mathbf{E}[e^{hX} - 1 - hX; X \le \epsilon/h] \le \frac{e^\epsilon h^2\sigma^2}{2}, \quad (3.32)$$

and to bound the second term we distinguish two cases.

At first, let $y \le t/h$. Then,

$$\mathbf{E}[e^{hX} - 1 - hX; X \in (\epsilon/h, y]] \le \frac{e^t h^2}{2}\mathbf{E}[X^2; X \in (\epsilon/h, t/h]],$$

and if $y > t/h$, we obtain

$$\begin{aligned}\mathbf{E}[e^{hX} - 1 - hX; X \in (\epsilon/h, y]] \\ \le \frac{e^t h^2}{2}\mathbf{E}[X^2; X \in (\epsilon/h, t/h]] + \mathbf{E}[e^{hX} - 1 - hX; X \in (t/h, y]].\end{aligned}$$

The function $(e^{hu} - 1 - hu)/u^t$ is increasing for $u > t/h$, hence,

$$\mathbf{E}[e^{hX} - 1 - hX; X \in (t/h, y]] \le \frac{e^{hy} - 1 - hy}{y^t}A_{t,+}$$

and thereby,

$$\mathbf{E}[e^{hX}-1-hX; X\in(\epsilon/h,y]] \le \frac{e^t h^2}{2}\mathbf{E}[X^2; X\in(\epsilon/h,t/h]] + \frac{e^{hy}-1-hy}{y^t}A_{t,+}. \quad (3.33)$$

As a consequence the second bound holds for all values of y and combining (3.31), (3.32) and (3.33) gives us the following bound:

$$\mathbf{E}[e^{hX}; X\le y] \le 1-ha+\frac{e^\epsilon h^2\sigma^2}{2}+\frac{e^t h^2}{2}\mathbf{E}[X^2; X\in(\epsilon/h,t/h]]+\frac{e^{hy}-1-hy}{y^t}A_{t,+}. \quad (3.34)$$

Following further the method of the proof of Theorem 3 from [42] we split the right hand side of (3.34) into three parts:

$$f_1(h) := -\alpha ha + e^\epsilon\sigma^2\frac{h^2}{2},$$

$$f_2(h) := -\beta ha + \frac{e^{hy}-1-hy}{y^t}A_{t,+},$$

$$f_3(h) := -\gamma ha + \frac{e^t h^2}{2}\mathbf{E}[X^2; X\in(\epsilon/h,t/h]],$$

where $\gamma\in(0,1)$ with $\gamma = 1-\alpha-\beta$.

We consider f_1, f_2 and f_3 separately. It is clear that

$$h_1 := \frac{2\alpha a}{e^\varepsilon \mathbf{E}[X^2]}$$

is the positive solution of the equation $f_1(h)=0$. Moreover, $f_1(h)<0$ for all $h\in(0,h_1)$.

Furthermore, it is easy to see that f_2 takes it's unique minimum in

$$h_2 := \frac{1}{y}\ln\left(1+\frac{\beta a}{A_{t,+}}y^{t-1}\right).$$

Since f_2 is convex, one has

$$f_2(h)<0 \quad \text{for all } h\in(0,h_2]. \quad (3.35)$$

Obviously, $\mathbf{E}[X^2; X\in(\epsilon/h,t/h]]\to 0$ as $a\to 0$ for all $h=o(1)$ as $a\to 0$. Hence, $f_3(h)<0$ for all h such that $h/a=O(1)$ as $a\to 0$.

The assumption in Theorem 3.3(i) means that $h_1\le h_2$. In this case, taking into account (3.35) and $h_1/a=O(1)$, we obtain

$$f_1(h_1)+f_2(h_1)+f_3(h_1)<0.$$

From the latter inequality we conclude that h_1 satisfies (3.25) and by applying (3.26) with $h=h_1$ we obtain (3.10).

Under the conditions of Theorem 3.3 (ii) we have $h_2\le h_1$. By the same arguments we get

$$f_1(h_2)+f_2(h_2)+f_3(h_2)<0.$$

Then, applying (3.26) with $h=h_2$ and using the inequality $(1+u)^{x/y}-1\ge u^{x/y}$, $u\ge 0$, we obtain (3.12).

3.4.2 Proof of Proposition 3.6

We want to use Theorem 2.1 from [12]. If we put $F := F_{-X}$ the conditions (G1)-(G3) of this theorem are fulfilled in our setting. Hence we get

$$\mathbf{E}[R_z] \leq c \int_0^\infty \mathbf{P}(-X > u)du + c \int_0^\infty \int_u^{u+z} \mathbf{P}(-X > v)dvdu, \tag{3.36}$$

where

$$c = \frac{b^*(\epsilon a)}{a(1-\epsilon)} \tag{3.37}$$

with $b^*(u) = \min\{v : -\mathbf{E}[X; X < -v] \leq u\}$ and $\epsilon \in (0,1)$ arbitrary. Clearly,

$$\int_0^\infty \mathbf{P}(-X > u)du = \mathbf{E}[-X; X < 0]. \tag{3.38}$$

Changing the order of integration gives us

$$\begin{aligned}\int_0^\infty \int_u^{u+z} \mathbf{P}(-X > v)dvdu &= \int_0^z v\mathbf{P}(-X > v)dv + z\int_z^\infty \mathbf{P}(-X > v)dv \\ &\leq z^{2-t} \int_0^\infty v^{t-1}\mathbf{P}(-X > v)dv = \frac{z^{2-t}}{t} A_{t,-}.\end{aligned} \tag{3.39}$$

As you can easily see,

$$b^*(u) \leq \left(\frac{A_{t,-}}{u}\right)^{1/(t-1)},$$

therefore by (3.37)

$$c \leq \frac{A_{t,-}^{1/(t-1)}}{a^{t/(t-1)}\epsilon^{1/(1-t)}(1-\epsilon)},$$

and by minimization over $\epsilon \in (0,1)$

$$c \leq \frac{t^{t/(t-1)} A_{t,-}^{1/(t-1)}}{(t-1)a^{t/(t-1)}}. \tag{3.40}$$

Finally, combining (3.36), (3.38), (3.39) and (3.40) gives us the desired result.

3.4.3 Proof of Theorem 3.7

We prove (3.17) only. The proof of the second bound goes along the same line.

Using Theorem 3.1 with $y = \theta(x + jz)$, we obtain

$$\begin{aligned}\mathbf{P}(M_\tau > x + jz) &\leq \frac{A_t^{1/\theta}\theta^{-1-(t-1)/\theta}\mathbf{E}[\tau_z]}{a^{1/\theta-1}(x+jz)^{1+(t-1)/\theta}} \ln\left(1 + \frac{a\theta^{t-1}(x+jz)^{t-1}}{A_t}\right) \\ &+ \left(1 + \frac{A_t^{1/\theta}\theta^{-(t-1)/\theta}}{a^{1/\theta}}(x+jz)^{-(t-1)/\theta}\right)\mathbf{E}[\tau_z]\mathbf{P}(X > \theta(x+jz)),\end{aligned}$$

and in view of (3.7),

$$\begin{aligned}\mathbf{P}(M > x) \le & \frac{A_t^{1/\theta}\theta^{-1-(t-1)/\theta}}{a^{1/\theta-1}}\mathbf{E}[\tau_z]\Sigma_1(x,z) \\ & + \left(1+\frac{A_t^{1/\theta}\theta^{-(t-1)/\theta}}{a^{1/\theta}}x^{-(t-1)/\theta}\right)\mathbf{E}[\tau_z]\left(\mathbf{P}(X>\theta x)+\Sigma_2(x,z)\right), \quad (3.41)\end{aligned}$$

where

$$\Sigma_1(x,z) := \sum_{j=0}^{\infty} \ln\left(1+\frac{a\theta^{t-1}(x+jz)^{t-1}}{A_t}\right)(x+jz)^{-1-(t-1)/\theta}$$

and

$$\Sigma_2(x,z) := \sum_{j=1}^{\infty} \mathbf{P}(X > \theta(x+jz)).$$

Define

$$\widetilde{\Sigma}_1(x,z) := \sum_{j=1}^{\infty} \ln\left(1+\frac{a\theta^{t-1}(x+jz)^{t-1}}{A_t}\right)(x+jz)^{-1-(t-1)/\theta}.$$

The summands in this sum are strictly decreasing, so we conclude by the integral criteria for sums:

$$\begin{aligned}\widetilde{\Sigma}_1(x,z) \le & \sum_{j=1}^{\infty}\int_{j-1}^{j} \ln\left(1+\frac{a\theta^{t-1}(x+uz)^{t-1}}{A_t}\right)(x+uz)^{-1-(t-1)/\theta}du \\ = & \frac{1}{z}\int_x^{\infty} \ln\left(1+\frac{a\theta^{t-1}w^{t-1}}{A_t}\right)w^{-1-(t-1)/\theta}dw\end{aligned}$$

and further by integration by parts,

$$\begin{aligned}&\frac{1}{z}\int_x^{\infty} \ln\left(1+\frac{a\theta^{t-1}w^{t-1}}{A_t}\right)w^{-1-(t-1)/\theta}dw \\ &\quad \le \frac{\theta}{z(t-1)}\ln\left(1+\frac{a\theta^{t-1}x^{t-1}}{A_t}\right)x^{-(t-1)/\theta} + \frac{\theta^2}{z(t-1)}x^{-(t-1)/\theta}.\end{aligned}$$

Therefore, for all x sufficing $x^{t-1} \ge \theta^{1-t}(e^\theta - 1)A_t a^{-1}$ and $x \ge z(t-1)\theta^{-1}$,

$$\Sigma_1(x,z) \le \frac{3\theta}{z(t-1)}\ln\left(1+\frac{a\theta^{t-1}x^{t-1}}{A_t}\right)x^{-(t-1)/\theta}.$$

Furthermore, it is easy to see that

$$\Sigma_2(x,z) \le \sum_{j=1}^{\infty}\int_{j-1}^{j} \mathbf{P}(X > \theta(x+uz))du = \frac{1}{\theta z}\overline{F}^I(\theta x). \quad (3.42)$$

and Theorem 3.7(i) is proved.

The proof of Theorem 3.7(ii) goes along the same line with using Theorem 3.3(ii) instead of Theorem 3.1.

3.4.4 Proof of Theorem 3.9

Foss, Korshunov and Zachary have shown, see Theorem 5.1 in [27], that for any random walk with the drift $-a$ and x with $x \to \infty$ as $a \to 0$ one has the following lower bound:

$$\liminf_{a\to 0} \frac{\mathbf{P}(M^{(a)} > x)}{a^{-1}\overline{F}^I(x)} \geq 1. \tag{3.43}$$

It follows from the regular variation of $\mathbf{P}(X^{(0)} > u)$, that

$$\overline{F}^I(x) \sim \frac{1}{r-1} x^{-r+1} L(x) \quad \text{as } a \to \infty, \tag{3.44}$$

therefore

$$\mathbf{P}(M^{(a)} \geq x) \geq (1+o(1)) \frac{x^{-r+1}L(x)}{(r-1)a} \quad \text{as } a \to 0.$$

Thus, we only have to derive an upper bound.

During the rest of this proof we assume a to be sufficiently small in every inequality. We want to apply Theorem 3.7(ii) with $2 < t < r$ and arbitrary $\varepsilon > 0$. It is clear that

$$A_{t,+}^{(a)} := \mathbf{E}[(X^{(a)})^t; X^{(a)} > 0] \leq \mathbf{E}[(X^{(0)})^t; X^{(0)} > 0] = A_{t,+}^{(0)},$$

therefore $A_{t,+}^{(a)}$ is finite for $t < r$ and

$$\lim_{a\to 0} A_{t,+}^{(a)} = A_{t,+}^{(0)} > 0.$$

Furthermore, we have to show that (3.11) is fulfilled for $y = \theta x$ and $\varepsilon > 0$ sufficiently small under our assumptions. Since the function $y^{-1}\ln(1 + \beta a y^{t-1}/A_{t,+}^{(a)})$ is decreasing for $y \gg a^{1/(t-1)}$, we have the following bound for $x \geq c a^{-1} \ln a^{-1}$:

$$\begin{aligned} \frac{1}{\theta x} \ln\left(1 + \frac{\beta\theta^{t-1} a x^{t-1}}{A_{t,+}^{(a)}}\right) &\leq \frac{a}{\theta c \ln a^{-1}} \ln\left(1 + \frac{\beta\theta^{t-1}c^{t-1}}{A_{t,+}^{(a)}} a^{2-t} \ln^{t-1} a^{-1}\right) \\ &= \frac{(t-2)}{\theta c} a(1+o(1)). \end{aligned} \tag{3.45}$$

This implies that if we choose $c > (r-2)\sigma^2/2$ and $\theta = (1-\delta)(t-2)/(r-2)$, we can choose $\alpha < 1$ so close to 1 and $\delta, \varepsilon, a > 0$ so close to 0 that

$$\frac{t-2}{\theta c} = \frac{r-2}{(1+\delta)c} < \frac{2\alpha}{e^\varepsilon \mathbf{E}[X^2]}$$

and consequently x satisfies (3.11) for a small enough.

We take $z = z(a)$ satisfying $a^{-1} \ll z \ll x$. Then, combining (3.13) and (3.14), we get

$$\frac{\mathbf{E}[\tau_z]}{z} \sim \frac{1}{a} \quad \text{as } a \to 0. \tag{3.46}$$

Since $a^{-1} \ll x$ and $(t-1)/\theta - (r-1) > 1/\theta - 1$ for $\theta < (t-2)/(r-2)$,

$$a^{-1/\theta+1}\frac{\mathbf{E}[\tau_z]}{z}\ln\left(1+\frac{\beta\theta^{t-1}ax^{t-1}}{A_{t,+}^{(a)}}\right)x^{-(t-1)/\theta} = o\left(a^{-1}x^{-r+1}L(x)\right). \tag{3.47}$$

Furthermore, it follows from the condition $z = o(x)$ and the regular variation of $\mathbf{P}(X^{(0)} > x)$ that

$$z\mathbf{P}(X^{(a)} > x) = o\left(x^{-r+1}L(x)\right). \tag{3.48}$$

By combining (3.46) with (3.48) and (3.44) and regarding $L(\theta x) \sim L(x)$, we conclude

$$\left(1+\left(\frac{A_{t,+}^{(a)}}{\beta\theta^{t-1}ax^{t-1}}\right)^{1/\theta}\right)\mathbf{E}[\tau_z]\left(\frac{1}{\theta z}\overline{F}^I(\theta x)+\mathbf{P}(X^{(a)} > \theta x)\right)$$
$$\sim \theta^{-r}(r-1)^{-1}a^{-1}x^{-r+1}L(x). \tag{3.49}$$

Plugging (3.47) and (3.49) into (3.18) gives us

$$\limsup_{a\to 0}\frac{\mathbf{P}(M^{(a)} > x)}{a^{-1}x^{-r+1}L(x)} \le \frac{\theta^{-r}}{r-1}.$$

To complete the proof it suffices to note, that we can choose θ arbitrary close to 1 by choosing t close to r. This implies that the previous inequality is valid even with $\theta = 1$.

3.4.5 Proof of Theorem 3.10

Since (3.43) is valid for all distributions with negative expectation, we again need an upper bound only. Let a be sufficiently small during this proof.

It follows from the assumptions in the theorem that $S_n^{(0)}/c_n$ converges weakly to a stable law of index r. The sequence c_n can be taken from the equation $c_n^{-r}L(c_n) = 1/n$. It is known that the function $g(a)$ in the heavy traffic approximation can be defined by the relations

$$g(a) = 1/c_{n_a} \text{ and } an_a \sim c_{n_a}.$$

The latter can be rewritten as

$$c_{n_a} \sim a\frac{(c_{n_a})^r}{L(c_{n_a})}.$$

From this we infer that (3.21) is equivalent to

$$\frac{ax^{r-1}}{L(x)} \to \infty \quad \text{as } a \to 0. \tag{3.50}$$

We want to apply Theorem 3.1 for $t = 2$ and $y = \theta x$ with $-\mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x]$ and $A_2(\theta x)$ instead of a and A_2 respectively. According to Remark 3.2 we have to show that $\mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x]$ is negative. Using the Markov inequality, we have

$$\mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x] \le -a + (\theta x)^{-1}\mathbf{E}[(\min\{0, X^{(0)}\})^2].$$

In view of (3.50), $ax \to \infty$. Therefore,

$$\mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x] \le -a(1 + o(1)).$$

Furthermore,

$$A_2(\theta x) \sim \frac{r}{2-r}(\theta x)^{2-r} L(\theta x) \tag{3.51}$$

and consequently by $-\mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x] \sim a$,

$$\begin{aligned} A_2^{1/\theta}(\theta x)\mathbf{E}[\tau_z] & \frac{(-\mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x])^{1-1/\theta}}{\theta^{1+1/\theta} x^{1+1/\theta}} \ln\left(1 - \frac{\theta x \mathbf{E}[X^{(a)}; |X^{(a)}| \le \theta x]}{A_2(\theta x)}\right) \\ & \le (1 + o(1)) k_1 \mathbf{E}[\tau_z] \mathbf{P}(X^{(a)} > x) \ln\left(1 + k_2 \frac{ax^{r-1}}{L(x)}\right)\left(\frac{ax^{r-1}}{L(x)}\right)^{-(1/\theta - 1)} \end{aligned} \tag{3.52}$$

with appropriate constants k_1 and k_2. Then, (3.50) implies that

$$\ln\left(1 + c_2 \frac{ax^{r-1}}{L(x)}\right)\left(\frac{ax^{r-1}}{L(x)}\right)^{-(1/\theta - 1)} = o(1). \tag{3.53}$$

Furthermore,

$$\frac{A_2^{1/\theta}(\theta x)}{a^{1/\theta}} \theta^{-1/\theta} x^{-1/\theta} \sim k_3 \left(\frac{ax^{r-1}}{L(x)}\right)^{-1/\theta}$$

with k_3 suitable and hence by (3.50),

$$\left(1 + \frac{A_2^{1/\theta}(\theta x)}{a^{1/\theta}} \theta^{-1/\theta} x^{-1/\theta}\right) = 1 + o(1). \tag{3.54}$$

Then, combining (3.52), (3.53) and (3.54), Theorem 3.1 with $t = 2$ and $y = \theta x$ gives us

$$\mathbf{P}(M_\tau^{(a)} > x) \le (1 + o(1))\theta^{-r} \mathbf{E}[\tau_z] \mathbf{P}(X^{(a)} > x),$$

where $\theta \in (0, 1)$ is arbitrary. By the summation formula (3.7) we get a bound for the total maximum:

$$\mathbf{P}(M^{(a)} > x) \le (1 + o(1))\theta^{-r} \mathbf{E}[\tau_z] \sum_{j=0}^{\infty} \mathbf{P}(X^{(a)} > x + jz). \tag{3.55}$$

Combining (3.42) and (3.44) with $a^{-1} \ll z \ll x$ gives us

$$\begin{aligned} \sum_{j=0}^{\infty} \mathbf{P}(X^{(a)} > x + jz) & \le (1 + o(1))\left(x^{-r} L(x) + \frac{x^{-r+1} L(x)}{z(r-1)}\right) \\ & \sim (1 + o(1)) \frac{x^{-r+1} L(x)}{z(r-1)}. \end{aligned}$$

Regarding (3.46) and letting $\theta \to 1$ completes the proof.

3.4.6 Proof of Theorem 3.11

We write τ instead of τ_z for brevity. To derive the lower bound we use the following

Lemma 3.14. *If* $\mathbf{E}[|X^{(0)}|^r] < \infty$ *for some* $r > 2$, *then*

$$\mathbf{P}\left(M_\tau^{(a)} > x\right) \geq (1 - e^{-h_a z})e^{-h_a(x+k/a)}, \tag{3.56}$$

where $k > 0$ *is an arbitrary constant and*

$$h := h_a = \frac{2a}{\sigma^2} + O(a^{1+\delta}). \tag{3.57}$$

Beweis. We omit the superscript $^{(a)}$ during this proof. To derive the bound (3.56) we introduce the process $\{W_k\}_{k\geq 0}$ defined by

$$W_0 := 1, \quad W_k := \prod_{i=1}^{k} e^{hX_i}\mathbf{1}\{X_i \leq y\}, \ k \geq 1.$$

It is clear, that for every h sufficing

$$\mathbf{E}[e^{hX}; X \leq y] = 1 \tag{3.58}$$

$\{W_k\}$ is a martingale. To find a suitable choice of h we put $y := k/a$ with a constant $k > 0$ and split the expected value on the left hand side of (3.58) as follows:

$$\begin{aligned}\mathbf{E}[e^{hX}, X \leq k/a] = 1 - ha + \frac{h^2}{2}(\sigma^2 + a^2) - \mathbf{P}(X > k/a) - h\mathbf{E}[X; X > k/a] \\ - \frac{h^2}{2}\mathbf{E}[X^2; X > k/a] + \mathbf{E}\left[e^{hX} - 1 - hX - \frac{h^2}{2}X^2; X \leq k/a\right].\end{aligned} \tag{3.59}$$

By restricting ourselves to such h that satisfy $h \leq ca$, $c > 0$, for a small enough we conclude by the Markov inequality:

$$\mathbf{P}(X > k/a) + h\mathbf{E}[X; X > k/a] + \frac{h^2}{2}\mathbf{E}[X^2; X > k/a] = O(a^r). \tag{3.60}$$

If $r \geq 3$, one can easily infer from the Taylor formula that

$$\mathbf{E}\left[e^{hX} - 1 - hX - \frac{h^2}{2}X^2; X \leq k/a\right] = O(a^3) \tag{3.61}$$

and if $r \in (2, 3)$, one can show with the Taylor formula that

$$\mathbf{E}\left[e^{hX} - 1 - hX - \frac{h^2}{2}X^2; |X| \leq k/a\right] \leq \tilde{c}_1 a^3 \mathbf{E}[|X|^3; |X| \leq -k/a]$$

with a suitable constant $\widetilde{c}_1 > 0$. Furthermore,

$$\mathbf{E}[|X|^3; |X| \le k/a] \le k^{3-r} a^{r-3} \mathbf{E}[|X|^r].$$

Thus,

$$\mathbf{E}\left[e^{hX} - 1 - hX - \frac{h^2}{2} X^2; |X| \le k/a\right] = O(a^r).$$

On the other side, $|e^{hX} - 1| < 1$ for $X < -k/a$ and thus

$$\begin{aligned}&\left|\mathbf{E}\left[e^{hX} - 1 - hX - \frac{h^2}{2} X^2; X < -k/a\right]\right| \\ &\le \mathbf{P}(X < -k/a) + h\mathbf{E}[|X|; X < -k/a] + h^2 \mathbf{E}[|X|^2; X < -k/a]\end{aligned}$$

By the Markov inequality and the existence of a power moment of order r we infer from the latter inequality that

$$\mathbf{E}\left[e^{hX} - 1 - hX - \frac{h^2}{2} X^2; X < -k/a\right] = o(a^r).$$

Hence, we have for all $r \in (2, 3)$

$$\mathbf{E}\left[e^{hX} - 1 - hX + \frac{h^2}{2} X^2; X \le k/a\right] = O(a^r). \tag{3.62}$$

Combining (3.61) and (3.62) we obtain for all $r \in (2, 3)$,

$$\mathbf{E}[e^{hX} - 1 - hX + \frac{h^2}{2} X^2; X \le k/a] = O(a^{\min\{3,r\}})$$

and, using (3.60) once again, one can rewrite (3.59) as

$$\mathbf{E}[e^{hX}; X \le k/a] = 1 - ha + \frac{h^2}{2}(\sigma^2 + a^2) + O(a^{2+\delta})$$

with $\delta = \min\{1, r - 2\} > 0$. Thereby,

$$h = h_a = \frac{2a}{\sigma^2} + O(a^{1+\delta}) \tag{3.63}$$

solves the equation (3.58) and $\{W_k\}$ with $h = h_a$ is a martingale.

We want to use this result to prove Lemma 3.14 and therefore introduce the stopping times

$$\sigma := \min\{k \ge 1 : X_k > k/a\},\ t_x := \min\{k \ge 1 : S_k > x\} \text{ and } T := \min\{\sigma, t_x, \tau\}.$$

Applying the Optional Stopping Theorem to the martingale $\{W_k\}$ and the stopping time T leads

$$1 = W_0 = \mathbf{E}[W_T] = \mathbf{E}\left[W_T; t_x < \tau, t_x < \sigma_{k/a}\right] + \mathbf{E}\left[W_T; \tau < t_x, \tau < \sigma_{k/a}\right]. \tag{3.64}$$

We consider the two terms on the right-hand-side separately:

$$\begin{aligned}\mathbf{E}\left[W_T; t_x < \tau, t_x < \sigma_{k/a}\right] &\le e^{h_a(x+k/a)}\mathbf{P}(t_x < \tau, t_x < \sigma_{k/a}) \\ &\le e^{h_a(x+k/a)}\mathbf{P}\left(M_\tau > x\right)\end{aligned}$$

and

$$\mathbf{E}\left[W_T; \tau < t_x, \tau < \sigma_{k/a}\right] \le \mathbf{E}[e^{h_a S_\tau}] \le e^{-h_a z}.$$

If we plug the last two results into (3.64), we obtain the stated result:

$$\mathbf{P}(M_\tau > x) \ge (1 - e^{-h_a z})e^{-h_a(x+k/a)}.$$

□

To establish the lower bound in (3.30) we apply Lemma 3.14 with $z \gg a^{-1}$ and get

$$\mathbf{P}(M_\tau^{(a)} > x) \ge (1 + o(1))e^{-2(ax+k)/\sigma^2}.$$

This inequality holds for all $k > 0$, hence by $k \to 0$,

$$\mathbf{P}(M_\tau^{(a)} > x) \ge (1 + o(1))e^{-2ax/\sigma^2}$$

and we get the desired lower bound for the total maximum by stopping the total maximum in τ:

$$\mathbf{P}(M^{(a)} > x) \ge \mathbf{P}(M_\tau^{(a)} > x).$$

To get an upper bound we need another

Lemma 3.15. *For all $h > 0$ satisfying* (3.25) *we have the inequality*

$$\mathbf{P}(M_\tau^{(a)} > x) \le (1 - e^{-h(z+y)})e^{-hx} + \mathbf{E}[\tau]\mathbf{P}(|X^{(a)}| > y). \tag{3.65}$$

Beweis. Let us again omit the superscript $^{(a)}$ for reasons of clarity. By truncating the random variables $|X_i|$ in the level y follows

$$\mathbf{P}\left(M_\tau > x\right) \le \mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\le k\le \tau}|X_k| \le y\right\} > x\right) + \mathbf{P}\left(\max_{1\le k\le\tau}|X_k| > y\right) \tag{3.66}$$

and due to Wald's Identity,

$$\mathbf{P}\left(\max_{1\le k\le\tau}|X_k| > y\right) \le \mathbf{E}\left[\sum_{k=1}^{\tau}\mathbf{1}\{|X_k| > y\}\right] = \mathbf{E}\left[\tau\right]\mathbf{P}\left(|X| > y\right). \tag{3.67}$$

To bound the first term on the right hand side of (3.66) we define

$$\widetilde{\sigma}_y := \min\{k \ge 1 : |X_k| > y\} \text{ and } \widetilde{T} := \min\{\widetilde{\sigma}_y, t_x, \tau\}.$$

Applying the Optional Stopping Theorem to the martingale $\{W_k\}$ from the proof of Lemma 3.14 and the stopping time $\widetilde{T}$ yields

$$1 = W_0 = \mathbf{E}[W_{\widetilde{T}}] \ge \mathbf{E}\left[W_{\widetilde{T}}\mathbf{1}\{t_x < \tau, t_x < \widetilde{\sigma}_y\}\right] + \mathbf{E}\left[W_{\widetilde{T}}\mathbf{1}\{\tau < t_x, \tau < \widetilde{\sigma}_y\}\right].$$

We bound these two terms as follows:

$$\begin{aligned}\mathbf{E}\left[W_{\widetilde{T}}; t_x < \tau, t_x < \widetilde{\sigma}_y\right] &\geq e^{hx}\mathbf{P}(t_x < \tau < \widetilde{\sigma}_y) \\ &= e^{hx}\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\leq k\leq \tau} |X_k| \leq y\right\} > x\right)\end{aligned}$$

and

$$\mathbf{E}\left[W_{\widetilde{T}}; \tau < t_x, \tau < \widetilde{\sigma}_y\right] \geq e^{-h(z+y)}$$

Therefore,

$$\mathbf{P}\left(M_\tau \mathbf{1}\left\{\max_{1\leq k\leq \tau} |X_k| \leq y\right\} > x\right) \leq (1 - e^{-h(z+y)})e^{-hx}, \tag{3.68}$$

and by combining (3.66), (3.67) and (3.68) we attain (3.65). □

To get a bound for the total maximum from this lemma we split the tail probability

$$\mathbf{P}(M^{(a)} > x) \leq \mathbf{P}(M_\tau^{(a)} > x) + \mathbf{P}\left(\sup_{k\geq\tau} S_k^{(a)} > x\right).$$

Since we can consider the continuation of the process (S_k) beyond τ as a probabilistic replica of the entire process with starting point S_τ it follows by $S_\tau \leq -z$, that

$$\mathbf{P}\left(\sup_{k\geq\tau} S_k^{(a)} > x\right) \leq \mathbf{P}(M^{(a)} > x + z)$$

and hence,

$$\mathbf{P}(M^{(a)} > x) \leq \mathbf{P}(M_\tau^{(a)} > x) + \mathbf{P}(M^{(a)} > x + z). \tag{3.69}$$

Let

$$j_0 = \min\left\{j \in \mathbb{N} : x + jz > (r-2)\sigma^2 \ln(1/a)/a\right\},$$

then we conclude inductively from (3.52), that

$$\mathbf{P}(M^{(a)} > x) \leq \sum_{j=0}^{j_0-1} \mathbf{P}(M_\tau^{(a)} > x + jz) + \mathbf{P}(M^{(a)} > x + j_0 z). \tag{3.70}$$

We consider the two terms on the right hand side of (3.70) separately. To bound the first term we use Lemma 3.15 with $h = h_a$ and $z = y = k/a$.

$$\sum_{j=0}^{j_0-1} \mathbf{P}(M_\tau^{(a)} > x + jz) \leq \sum_{j=0}^{j_0-1} (1 - e^{-h_a(z+y)})e^{-h_a(x+jz)} + j_0 \mathbf{E}[\tau]\mathbf{P}(|X^{(a)}| > y). \tag{3.71}$$

As you can easily see, $j_0 \leq \widetilde{c} \ln a^{-1}$ with a positive constant $\widetilde{c}$. Therefore our choice of y, z and h_a includes

$$\sum_{j=0}^{j_0-1} e^{-h_a z j} = \frac{1 - e^{-h_a z j_0}}{1 - e^{-h_a z}} \sim \frac{1}{1 - e^{-h_a z}}. \tag{3.72}$$

We recall $z = y$ and conclude

$$\frac{1 - e^{-h_a(z+y)}}{1 - e^{-h_a z}} = 1 + e^{-h_a y} \sim 1 + e^{-2k/\sigma^2} \tag{3.73}$$

and, by combining (3.72) and (3.73),

$$\sum_{j=0}^{j_0-1} (1 - e^{-h_a(z+y)}) e^{-h_a(x+jz)} \sim e^{-h_a x}(1 + e^{-2k/\sigma^2}). \tag{3.74}$$

On the other hand, Wald's identity and the inequality of Lorden (see (3.14)) give

$$\mathbf{E}[\tau] \le \frac{k + \mathbf{E}[((X^{(a)})^-)^2]}{a^2}. \tag{3.75}$$

Hence, we can follow by using $j_0 \le \widetilde{c} \ln a^{-1}$ and the Markov inequality that

$$j_0 \mathbf{E}[\tau] \mathbf{P}(|X| > y) \le \widetilde{c} A_r (k + \mathbf{E}[((X^{(a)})^-)^2]) k^{-r} a^{r-2} \ln a^{-1}$$

and by plugging the latter results into (3.71) we attain

$$\sum_{j=0}^{j_0-1} \mathbf{P}(M_\tau^{(a)} > x + jz) \le (1 + e^{-2k/\sigma^2} + o(1)) e^{-h_a x} + \widetilde{c}_2 a^{r-2} \ln a^{-1}. \tag{3.76}$$

with a suitable constant $\widetilde{c}_2$. To bound the second term in (3.70) denote by c a constant with $(r-2)\sigma^2/2 < c \le (r-2)\sigma^2$. The definition of j_0 gives

$$\mathbf{P}(M^{(a)} > x + j_0 z) \le \mathbf{P}\left(M^{(a)} > ca^{-1} \ln a^{-1}\right). \tag{3.77}$$

Let us apply Theorem 3.7(ii) to the probability on the right hand side of the latter inequality. Put $x = ca^{-1} \ln a^{-1}$ and $y = \theta x$ with $\theta = (t-2)/(r-2)$, where $2 < t < r$. Then, one can see similar to (3.45) that (3.11) is fulfilled for $\alpha < 1$ close enough to 1 and a, ε close enough to 0. As one can easily see, the other two conditions of Theorem 3.7(ii) are also fulfilled for a small enough if one chooses $z = 1/a$. Therefore we may apply Theorem 3.7(ii) with these values of t, θ, ε and z for a small enough. Consequently,

$$\mathbf{P}(M^{(a)} > ca^{-1} \ln a^{-1}) \le P_1(a) + P_2(a) \tag{3.78}$$

with

$$P_1(a) = \widetilde{c}_3 \mathbf{E}[\tau] \ln\left(1 + \widetilde{c}_3 a^{2-t} \ln^{t-1} a^{-1}\right) a^{(t-1)/\theta + 1} \ln^{-(t-1)/\theta} a^{-1} \tag{3.79}$$

and

$$P_2(a) = \left(1 + \widetilde{c}_4^{-1/\theta} a^{(t-2)/\theta} \ln^{-(t-1)/\theta} a^{-1}\right) \mathbf{E}[\tau] \cdot \\ \cdot \left(a\theta^{-1} \overline{F}^I(\theta ca^{-1} \ln a^{-1}) + \mathbf{P}(X^{(a)} > \theta ca^{-1} \ln a^{-1})\right) \tag{3.80}$$

where $\widetilde{c}_3$ and $\widetilde{c}_4$ are positive constants that are independent of a. Since $\theta = (t-2)/(r-2)$ with $2 < t < r$ we obtain by regarding (3.75),

$$P_1(a) \leq \widetilde{c}_5 a^{(t-1)/\theta - 1} \ln^{1-(t-1)/\theta} a^{-1}. \tag{3.81}$$

with and an appropriate constant $\widetilde{c}_5$. On the other hand, the Markov inequality gives

$$a\theta^{-1}\overline{F}^I(\theta c a^{-1} \ln a^{-1}) \leq \widetilde{c}_6 a^r \ln^{1-r} a^{-1}$$

and

$$\mathbf{P}(X^{(a)} > \theta c a^{-1} \ln a^{-1}) \leq \widetilde{c}_7 a^r \ln^{-r} a^{-1}$$

with $\widetilde{c}_6$ and $\widetilde{c}_7$ appropriate and obviously, for every $2 < t < r$,

$$1 + \widetilde{c}_4^{-1/\theta} a^{(t-2)/\theta} \ln^{-(t-1)/\theta} a^{-1} \sim 1.$$

Consequently, by additionally considering (3.75),

$$P_2(a) \leq \widetilde{c}_8 a^{r-2} \ln^{1-r} a^{-1} \tag{3.82}$$

with $\widetilde{c}_8$ suitable. By $2 < t < r$,

$$(r-2)\theta = t - 2 < t - 1 - \theta.$$

Therefore the combination of (3.78), (3.81) and (3.82) gives $P_1(a) = o(P_2(a))$ and consequently

$$\mathbf{P}(M^{(a)} > c a^{-1} \ln a^{-1}) \leq (1 + o(1)) \widetilde{c}_8 a^{r-2} \ln^{1-r} a^{-1}. \tag{3.83}$$

By plugging the latter result and (3.76) into (3.70), we get the following asymptotic bound:

$$\begin{aligned} \mathbf{P}(M^{(a)} > x) \leq &(1 + o(1)) \left(e^{-h_a x} + \widetilde{c}_2 a^{r-2} \ln a^{-1} + \widetilde{c}_4 a^{r-2} \ln^{1-(t-1)/\theta} a^{-1}\right) \\ = &(1 + o(1)) \left(e^{-h_a x} + \widetilde{c}_2 a^{r-2} \ln a^{-1}\right) \end{aligned}$$

By virtue of the relation (3.23),

$$x \leq c a^{-1} \ln a^{-1}$$

with some $c < (r-2)\sigma^2/2$ for a small enough, therefore

$$e^{-h_a x} \geq (1 + o(1)) a^{2c/\sigma^2},$$

and consequently

$$a^{r-2} \ln a^{-1} = o(e^{-h_a x}).$$

Finally,

$$\mathbf{P}(M^{(a)} > x) \leq (1 + o(1)) e^{-h_a x},$$

and the proof is complete.

4 A local limit theorem for the maximum of a random walk in the heavy traffic regime

This chapter contains results from the publication [37]. To the knowledge of the author there are no known local limit theorems concerning the heavy traffic regime and the results attained in this chapter are totally new.

4.1 Introduction and statement of results

We again use the notation from Chapter 1.7 and again consider a random walk with negative drift $-a$. In the case $a = 0$ write S, X_i and X instead of $S^{(0)}$, $X_i^{(0)}$ and $X^{(0)}$ respectively. As stated in Chapter 1.8, the random walk drifts to $-\infty$ for all $a > 0$ and as $a \to 0$ the so-called heavy traffic asymptotics (see (1.30)) holds for every fixed value of x:

$$\mathbf{P}(M^{(a)} > x) \sim e^{-2ax/\sigma^2} \quad \text{as } a \to 0. \tag{4.1}$$

In interesting mathematical question is whether there also a local version of this result and this is what this chapter is about.

We assume that $X^{(a)}$ possesses a $\Delta^{(a)}$-lattice distribution with zero shift, that means there exists some $\Delta^{(a)} > 0$ such that $\mathbf{P}(X^{(a)} \in \Delta^{(a)}\mathbb{Z}) = 1$ and $\Delta^{(a)}$ is the maximal positive number with this property. Assume that $\Delta^{(a)} \to \Delta^{(0)} > 0$ as $a \to 0$ and in the case $a = 0$ write Δ instead of $\Delta^{(0)}$. Due to rescaling we can assume without loss of generality that $\Delta^{(a)} = \Delta \equiv 1$. Suppose that

$$X^{(a)} \xrightarrow{w} X \quad \text{as } a \to 0 \tag{4.2}$$

and

$$\sup_{a\in[0,a_0]} \mathbf{E}[X^{(a)}]^2 < \infty \quad \text{and} \quad \sup_{a\in[0,a_0]} \mathbf{E}[(\max\{0, X^{(a)}\})^{2+\varepsilon}] < \infty \tag{4.3}$$

for some $a_0, \varepsilon > 0$.

Our main result is a local limit theorem for the probability $\mathbf{P}(M^{(a)} = y)$ as $a \to 0$ for y such that $y \to \infty$ and $ay = O(1)$. The main idea for our proof is to find a a representation of the probability $\mathbf{P}(M^{(a)} = y)$ as a geometric sum and to derive and apply a uniform renewal theorem to find the asymptotic behaviour of this sum. This uniform renewal theorem will be a generalization of a result attained by Nagaev [41].

It is worth mentioning that the approach used in this chapter is similar to the method used in [7], where the authors use the well-known representation of $\mathbf{P}(M^{(a)} > y)$ as a geometric sum of independent random variables (see for example [3]) and a uniform renewal theorem from [6] to establish the asymptotic behaviour of $\mathbf{P}(M^{(a)} > y)$ as $a \to 0$ and $y \to \infty$ for subexponential distributions. In [6] there is also a uniform renewal theorem used to develop asymptotic expansions of the distribution of a geometric sum. Let us also mention that the local behaviour of the probability $\mathbf{P}(M^{(a)} = x)$ as $x \to \infty$ is known in the sense that there exists a local limit theorem for $x \to \infty$. Namely, Theorem 5.13 in [27] states that for a random walk with negative drift whose increments are long tailed and strong subexponential,

$$\mathbf{P}(M^{(a)} \in (x, x+T]) \sim \frac{T}{a}\overline{F}(x) \quad \text{as } x \to \infty.$$

In the lattice case x and T should be restricted to the lattice and for a 1-lattice with zero shift we attain

$$\mathbf{P}(M^{(a)} = x) \sim \frac{1}{a}\overline{F}(x) \quad \text{as } x \to \infty.$$

We now state our main result.

Theorem 4.1. *Assume that* (4.2) *and* (4.3) *hold and suppose that* $X^{(a)}$ *possesses an aperiodic* 1*-lattice distribution with zero shift for a small enough. Then, as* $a \to 0$*,*

$$\mathbf{P}(M^{(a)} = y) \sim \frac{2a}{\sigma^2} \exp\left\{-\frac{2ay}{\sigma^2}\right\} \tag{4.4}$$

uniformly for all y *such that* $y \to \infty$ *and* $ay = O(1)$ *as* $a \to 0$.

Remark that our model excludes the case $S_1^{(a)} = S_1^{(0)} - a$. Examining this case would be desireable, however this would be a different problem. In the non-local case it is known (see for example Wachtel and Shneer [49]) that one only needs to assume $\lim_{a\to 0} \mathbf{Var} X^{(a)} = \sigma^2 \in (0, \infty)$ and a Lindeberg-type condition

$$\lim_{a\to 0} \mathbf{E}[(X^{(a)})^2; |X_1^{(a)}| > K/a] = 0 \quad \text{for all } K > 0$$

to establish (4.1). This means that we must make stronger assumptions to establish our local result than it is needed in the non-local case.

Remark that it seems likely that one can get a non-lattice version of Theorem 4.1 with similar methods used here. The reason is the following: The proof of the uniform renewal Theorem which we need for the proof of Theorem 4.1 is based on results from Nagaev [41]. The results in [41] are derived under the assumption that the increments posess a absolutely continuous component and the proof should work similar in our case. The rest of the proof of a non-lattice analogy of Theorem 4.1 should work similar to the approach used in chapter 4.4.

It is also worth mentioning that Theorem 4.1 restates the heavy traffic asymptotics (4.1): As $a \to \infty$,

$$\begin{aligned}\mathbf{P}(M^{(a)} \geq y) &= \sum_{x=y}^{\infty} \mathbf{P}(M^{(a)} = x) \sim \frac{2a}{\sigma^2} \sum_{x=y}^{\infty} e^{-2ax/\sigma^2} \\ &= \frac{2a}{\sigma^2} \frac{e^{-2ay/\sigma^2}}{1 - e^{-2ay/\sigma^2}} \sim e^{-2ay/\sigma^2}\end{aligned}$$

for all y such that $y \to \infty$ and $ay = o(1)$ as $a \to 0$.

4.2 Uniform renewal theorem

In this section we prove a modification of Theorem 1 in Nagaev [41] which is, unlike the uniform renewal theorem from Nagaev, even uniform in the expected value. This renewal theorem is the key to the proof of our main result.

Consider a family of non-negative 1-latticed and aperiodic random variables $\{Z^{(b)}, b \in I\}$ with $\mathbf{E}[Z^{(b)}] = \mu^{(b)}$, a non-empty set $I \subseteq \mathbb{R}$ that contains at least one accumulation point and $\inf_{b\in I} \mu^{(b)} > 0$. Denote by $F^{(b)}$ the distribution function of $Z^{(b)}$ and by $F_k^{(b)}$ the k-fold convolution of $F^{(b)}$ with itself. Let

$$H(x, b, A) = \sum_{k=0}^{\infty} A^k F_k^{(b)}(x), \quad A > 0.$$

In renewal theory one usually studies the asymptotic behavior of $H(x+h, b, 1) - H(x, b, 1)$, $h > 0$. However, the case $A \neq 1$ is also of great interest. In [41], Nagaev's motivation for studying the case $A \neq 1$ comes from branching processes, since there arises a need for an asymptotic representation for $H(x + h, b, A) - H(x, b, A)$ as $x \to \infty$ with an estimate for the remainder term which is uniform in A. For our purposes we seek a representation for $H(x + h, b, A) - H(x, b, A)$ as $x \to \infty$ and the estimate for the remainder shall be uniform in A and b. Assume that there exists some $s > 1$ such that

$$\sup_{b\in I} \mathbf{E}[(Z^{(b)})^s] < \infty \tag{4.5}$$

and that

$$Z^{(b)} \xrightarrow{w} Z^{(0)}. \tag{4.6}$$

Put

$$f_k^{(b)} = F^{(b)}(k) - F^{(b)}((k-1)), \quad f_y^{(b)}(z) = \sum_{k=0}^{y} f_k^{(b)} z^k,$$

$$\mu_y^{(b)}(z) = f_y^{(b)\prime}(z) = \sum_{k=1}^{y} k f_k^{(b)} z^{k-1}.$$

Proposition 4.2. *Let $\lambda_y^{(b)}(A)$ be the real non-negative root of the equation $Af_y^{(b)}(z)$ $= 1$. Assume that* (4.5) *holds for some $s > 1$. For every accumulation point b_0 of I, there exists a positive constant α such that*

$$\sum_{k=1}^{\infty} A^k \left(F_k^{(b)}(y) - F_k^{(b)}((y-1))\right) = \frac{\left(\lambda_y^{(b)}(A)\right)^{-y-1}}{A\mu_y^{(b)}(\lambda_y^{(b)}(A))} + O(y^{-\min\{1,s-1\}} \ln y) \tag{4.7}$$

uniformly in $b \in I \cap \{b \in I : |b - b_0| \leq \alpha\}$ and $A_y \leq A \leq 1$, where

$$A_y = 1 - C/y \tag{4.8}$$

with a fixed positive number C.

4.3 Proof of the uniform renewal theorem

Although the uniform renewal theorem is a generalization of Theorem 1 in Nagaev [41], the main idea of the proof is the same. However, for reasons of completeness, we give the whole proof.

Let us assume without loss of generality $I = [0, b_1]$ with $b_1 > 0$. Suppose that y is sufficiently large in this section, even if it is not explicitly mentioned and throughout the following $\int_a^b g(x)dF^{(b)}(x)$ is to be interpreted as $\int_{a+}^{b+} g(x)dF^{(b)}(x)$.

Lemma 4.3. *Assume that* (4.5) *holds for some* $s > 1$. *Put* $U_y(\delta) = \{z : 1 \le |z| \le e^{h_y}, |\arg z| \le \delta\}$ *for some* $h_y = O(1/y)$. *Then,*

$$\lim_{\delta\to 0}\lim_{y\to\infty}\sup_{b\in I, z\in U_y(\delta)} |\mu_y^{(b)}(z) - \mu^{(b)}| = 0. \tag{4.9}$$

Beweis. First of all,

$$\begin{aligned}|\mu_y^{(b)}(z) - \mu^{(b)}| &= \left|\int_0^y xz^{x-1}dF^{(b)}(x) - \int_0^\infty xdF^{(b)}(x)\right| \\ &\le \int_0^y x|z^{x-1} - 1|dF^{(b)}(x) + \int_y^\infty xdF^{(b)}(x).\end{aligned} \tag{4.10}$$

When $x, |z| \ge 1$, one can easily see by Taylor's approximation that

$$|z^{x-1} - 1| \le x|z-1||z|^x.$$

Using this estimate we obtain for all $z \in U_y(\delta)$ and $N \le y$,

$$\begin{aligned}\int_0^N x|z^{x-1} - 1|dF^{(b)}(x) &\le |z-1|\int_0^N x^2|z|^x dF^{(b)}(x) \\ &\le |z-1|e^{h_y y}\int_0^N x^2 dF^{(b)}(x) \le N^2|z-1|e^{h_y y}.\end{aligned}$$

Furthermore, a straightforward trigonometric calculation shows that for δ sufficiently small,

$$|z-1| \le |z - e^{i\arg z}| + |1 - e^{i\arg z}| = |z| - 1 + \sqrt{2(1-\cos(\arg z))} \le e^{h_y} - 1 + 2\delta$$

for all $z \in U_y(\delta)$ and hence, as $y \to \infty$,

$$\int_0^N x|z^{x-1} - 1|dF^{(b)}(x) \le (e^{h_y} - 1 + 2\delta)N^2 e^{h_y y} = (2\delta + h_y + o(h_y))N^2 e^{h_y y}$$

uniformly in $b \in I$. At the same time, for $z \in U_y(\delta)$, assumption (4.5) and $h_y y = O(1)$ imply that there exists an absolute number $K > 0$ such that for all $N \le y$,

$$\begin{aligned}\int_N^y x|z^{x-1} - 1|dF^{(b)}(x) &\le (1 + e^{h_y y})\int_N^y xdF^{(b)}(x) \\ &\le \frac{1 + e^{h_y y}}{N^{s-1}}\int_N^\infty x^s dF^{(b)}(x) \le KN^{1-s}\end{aligned}$$

and by setting $N = (2\delta + h_y)^{-1/3}$ and choosing K_1 such that $e^{h_y y} \le K_1$, we attain

$$\begin{aligned}\int_0^y x|z^{x-1} - 1| dF^{(b)}(x) &\le (2\delta + h_y)^{1/3} e^{h_y y} + K(2\delta + h_y)^{(s-1)/3} + o(h_y) \\ &\le 2^{1/3}\delta^{1/3} K_1 + K 2^{(s-1)/3}\delta^{(s-1)/3} + o(1) \end{aligned} \tag{4.11}$$

uniformly in $b \in I$ as $y \to \infty$. Plugging (4.11) into (4.10) and using (4.5) once more, we conclude

$$|\mu_y^{(b)}(z) - \mu^{(b)}| \le 2^{1/3} K_1 \delta^{1/3} + K 2^{(s-1)/3}\delta^{(s-1)/3} + o(1)$$

uniformly in $b \in I$ as $y \to \infty$. □

Lemma 4.4. *Assume that* (4.5) *holds for some* $s > 1$. *Then, for large enough* y, $\lambda_y^{(b)}(A) < e^{h_y}$ *for all* $A_y \le A \le 1$ *and* $b \in I$, *where* $A_y = 1 - C/y$ *with some constant* $C > 0$ *and* $h_y = C_1/(\mu^{(0)} y)$ *with* $C_1 > C\mu^{(0)}/\inf_{b\in I}\mu^{(b)}$.

Beweis. We want to estimate the difference $\lambda_y^{(b)}(A) - 1$. First of all, by using the definition of $\lambda_y^{(b)}(A)$,

$$\begin{aligned}\int_{0-}^y \left((\lambda_y^{(b)}(A))^x - 1\right) dF^{(b)}(x) &= f_y^{(b)}(\lambda_y^{(b)}(A)) - \int_{0-}^y dF^{(b)}(x) \\ &= \frac{1}{A} - 1 + \int_y^\infty dF^{(b)}(x) = \frac{1-A}{A} + \int_y^\infty dF^{(b)}(x).\end{aligned}$$

Furthermore, $\lambda_y^{(b)}(A) \ge 1$ for $A \le 1$ and therefore by the binomial formula,

$$(\lambda_y^{(b)}(A))^x - 1 \ge x(\lambda_y^{(b)}(A) - 1), \quad x \ge 0.$$

Thus, (4.5) gives that uniformly in $b \in I$,

$$\begin{aligned}(\lambda_y^{(b)}(A) - 1)\int_{0-}^y x dF^{(b)}(x) &\le \int_{0-}^y \left((\lambda_y^{(b)}(A))^x - 1\right) dF^{(b)}(x) \\ &= \frac{1-A}{A} + \int_y^\infty dF^{(b)}(x) = \frac{1-A}{A} + O(y^{-s}).\end{aligned} \tag{4.12}$$

The condition $A_y \le A \le 1$ implies that $1 - A \le C/y$, hence

$$\frac{1}{A} = 1 + \frac{1-A}{A} = 1 + O\left(\frac{1}{y}\right)$$

and consequently

$$\frac{1-A}{A} \le \frac{C}{Ay} = \frac{C}{y} + O\left(\frac{1}{y^2}\right). \tag{4.13}$$

From the inequalities (4.12), (4.13) and (4.5) we conclude that

$$\begin{aligned}\lambda_y^{(b)}(A) - 1 &\le \frac{C/y + O(y^{-2}) + O(y^{-s})}{\mu^{(b)} - \int_y^\infty x dF^{(b)}(x)} = \frac{C/(\mu^{(b)} y)}{1 - O(y^{1-s})} + O(y^{-2}) + O(y^{-s}) \\ &= \frac{C}{\mu^{(b)} y} + O(y^{-2}) + O(y^{-s}) < \frac{C_1}{\mu^{(0)} y}\end{aligned}$$

uniformly in $b \in I$ for all y large enough. Therefore, since $e^x - 1 \geq x$ for all $x > 0$, $\lambda_y^{(b)}(A) < e^{h_y}$ uniformly in $A_y \leq A \leq 1$ and $b \in I$, if y is sufficiently large. □

Lemma 4.5. *Assume that (4.5) and (4.6) hold for some $s > 1$. Put $h_y = C_1/(\mu^{(0)}y)$ with a constant $C_1 > C\mu^{(0)}/\inf_{b\in I}\mu^{(b)}$. Then, there exists some $b_2 > 0$ such that for y large enough, $Af_y^{(b)}(z) - 1$ has no other zeros in the disc $|z| < e^{h_y}$ apart from $\lambda_y^{(b)}(A)$ and this holds uniform in $A_y \leq A \leq 1$ and $0 \leq b \leq b_2$.*

Beweis. First of all, for all $|z| \leq e^{h_y}$,

$$|\mu_y^{(b)}(z)| \leq \int_0^y x|z|^{x-1} dF^{(b)}(x) \leq e^{h_y y}\mu^{(b)}.$$

Using in addition $h_y y = O(1)$ and (4.5), we conclude

$$\sup_{y, b\leq b_1, |z|\leq e^{h_y}} |\mu_y^{(b)}(z)| < \infty. \tag{4.14}$$

Since the convergence radius of the derivative of a power series is the same as the convergence radius of the power series, we attain

$$\begin{aligned} &\lim_{y\to\infty} \sup_{b\leq b_1} \sup_{\substack{1\leq r\leq e^{h_y}\\ 0\leq\varphi\leq 2\pi}} \left| f_y^{(b)}(re^{i\varphi}) - f_y^{(b)}(e^{i\varphi})\right| \\ &= \lim_{y\to\infty} \sup_{b\leq b_1} \sup_{\substack{1\leq r\leq e^{h_y}\\ 0\leq\varphi\leq 2\pi}} \left|\mu_y^{(b)}(e^{i\varphi})\right| |r-1| = 0. \end{aligned} \tag{4.15}$$

On the other hand,

$$\begin{aligned} \lim_{y\to\infty} \sup_{b\leq b_1} \sup_{0\leq\varphi\leq 2\pi} \left|f_y^{(b)}(e^{i\varphi}) - f_\infty^{(b)}(e^{i\varphi})\right| &\leq \lim_{y\to\infty} \sup_{b\leq b_1} \sup_{0\leq\varphi\leq 2\pi} \int_y^\infty \left|e^{i\varphi x}\right| dF^{(b)}(x) \\ &= \lim_{y\to\infty} \sup_{b\leq b_1} \overline{F}^{(b)}(y) = 0. \end{aligned} \tag{4.16}$$

As $b \to 0$, $F^{(b)} \xrightarrow{w} F^{(0)}$ due to (4.6) and $F^{(0)}$ is not defective because of (4.5). Obviously, $u_\varphi(\cdot) = e^{i\varphi\cdot}$ is equicontinuous with $|u_\varphi| = 1 < \infty$. Hence, by a corollary in Chapter VIII.1 in Feller [25],

$$\int_0^\infty e^{i\varphi x} dF^{(b)}(x) \to \int_0^\infty e^{i\varphi x} dF^{(0)}(x) \tag{4.17}$$

uniformly in $0 \leq \varphi \leq \pi$ as $b \to 0$.

Now, let us first consider values of z in the circle $|z| < e^{h_y}$ that are not in the vicinity of $\lambda_y^{(b)}(A)$. Due to Lemma 4.4, these values can be characterized as those values that satisfy $|z| < e^{h_y}$ and $\delta \leq |\arg z| \leq \pi$, $\delta > 0$. We have

$$\sup_{\delta\leq\varphi\leq\pi} \left|f_\infty^{(0)}(e^{i\varphi})\right| = \sup_{\delta\leq\varphi\leq\pi} \left|\int_0^\infty e^{i\varphi x} dF^{(0)}(x)\right| < \sup_{\delta\leq\varphi\leq\pi} \int_0^\infty \left|e^{i\varphi x}\right| dF^{(0)}(x) = 1.$$

Combining the latter inequality with (4.17), we conclude that for $b_2 > 0$ sufficiently small

$$\sup_{b\leq b_2}\sup_{\delta\leq\varphi\leq\pi}|f_\infty^{(b)}(e^{i\varphi})| < 1$$

and since this inequality is strict,

$$m(\delta) := \inf_{b\leq b_2}\inf_{A\leq 1}\inf_{\delta\leq\varphi\leq\pi}\left|Af_\infty^{(b)}(e^{i\varphi}) - 1\right| > 0. \tag{4.18}$$

By combining (4.15), (4.16) and (4.18), we conclude that for y large enough and $A \in \mathfrak{A}_y = \{A : 1 - C/y = A_y \leq A \leq 1\}$,

$$\inf_{b\leq b_2}\inf_{\substack{1\leq r\leq e^{h_y}\\ \delta\leq\varphi\leq 2\pi}}\left|Af_y^{(b)}(re^{i\varphi}) - 1\right| > \frac{m(\delta)}{2} > 0. \tag{4.19}$$

On the basis of (4.19) we can assert that if $Af_y^{(b)}(z) - 1$ has a zero $\widetilde{\lambda}_y^{(b)}(A)$ in the disc $|z| \leq e^{h_y}$ differing from $\lambda_y^{(b)}(A)$, then $\widetilde{\lambda}_y^{(b)}(A)$ will lie outside the region $\{z : 1 \leq |z| \leq e^{h_y}, |\arg z| \geq \delta\}$ Note that the region $\{z : 1 \leq |z| \leq e^{h_y}, |\arg z| \geq \delta\}$ does not depend on b and A, so the latter holds for all values of $b \in I$ and $A \in \mathfrak{A}_y$.

Next, consider the region $U_y(\delta) = \{z : 1 \leq |z| \leq e^{h_y}, |\arg z| < \delta\}$. The equicontinuity of the family $\{f_y^{(b)}(\cdot) : b \in I, A \in \mathfrak{A}_y\}$ implies the existence of a $\delta_1(b, A) > 0$ such that $|Af_y^{(b)}(z) - 1|$ has no other zeros in the disc $|z - \lambda_y^{(b)}(A)| \leq \delta_1(b, A)$ apart from $\lambda_y^{(b)}(A)$. Therefore,

$$\widetilde{m}(\delta_2) := \inf_{b\leq b_2}\inf_{A\in\mathfrak{A}_y}\inf_{\substack{z:|z-\lambda_y^{(b)}(A)|\leq\delta_2\\ z\neq\lambda_y^{(b)}(A)}}|Af_y^{(b)}(z) - 1| > 0.$$

where $\delta_2 = \inf_{b\leq b_2}\inf_{A\in\mathfrak{A}_y}\delta_1(b, A) > 0$. This implies that for δ small enough, say $\delta \leq \delta_3$, $\widetilde{\lambda}_y^{(b)}(A)$ cannot lie in the region $\{z : 1 \leq |z| \leq e^{h_y}, |\arg z| < \delta_3\}$ and this holds uniform in $b \leq b_2$ and $A \in \mathfrak{A}_y$ with y large enough. Setting $\delta = \delta_3$ in (4.19) we conclude that $\widetilde{\lambda}_y^{(b)}(A)$ cannot lie in the annulus $1 \leq |z| \leq e^{h_y}$. Since $|\widetilde{\lambda}_y^{(b)}(A)| \geq 1$ for all $A \leq 1$, we finally obtain that $\widetilde{\lambda}_y^{(b)}(A)$ does not lie in the disc $|z| \leq e^{h_y}$, so $\lambda_y^{(b)}(A)$ is the only root of the equation $Af_y^{(b)}(A) = 1$ in the disc $|z| \leq e^{h_y}$ and this holds uniformly in $b \leq b_2$ and $A_y \leq A \leq 1$. □

Proof of Proposition 4.2. Let γ_y be a circle of radius $r_y = e^{h_y}$ with $h_y = C_1/(\mu^{(0)}y)$, $C_1 > \mu^{(0)} + C\mu^{(0)}/\inf_{b\leq b_1}\mu^{(b)}$ and C from (4.8). Then, according to Lemma 4.4 and Lemma 4.5, there exists some $b_2 > 0$ such that for all $0 \leq b \leq b_2$ and $A \in \mathfrak{A}_y$, the function $1 - Af_y^{(b)}(z)$ is zero in the disc $|z| \leq e^{h_y}$, if and only if $z = \lambda_y^{(b)}(A)$. Hence, the Residue theorem states that

$$\frac{1}{2\pi i}\int_{\gamma_y}\frac{z^{-y-1}}{1 - Af_y^{(b)}(z)}dz = \operatorname{Res}\left(\frac{z^{-y-1}}{1 - Af_y^{(b)}(z)}, \lambda_y^{(b)}(A)\right) + \operatorname{Res}\left(\frac{z^{-y-1}}{1 - Af_y^{(b)}(z)}, 0\right). \tag{4.20}$$

for $0 \le b \le b_2$ and $A \in \mathfrak{A}_y$.

In the following denote by $C_n(f(z))$, $n \ge 1$, the coefficient of z^n in the Taylor series of the function $f(z)$. An easy calculation shows that

$$A^n (f_\infty^{(b)}(z))^n = A^n \sum_{j=1}^{\infty} \left(F_n^{(b)}(j) - F_n^{(b)}(j-1) \right) z^j$$

and consequently, by changing the order of summation, it is not hard to see that

$$\sum_{k=1}^{\infty} A^k \left(F_k^{(b)}(n) - F_k^{(b)}(n-1) \right) = C_n \left(\frac{1}{1 - A f_\infty^{(b)}(z)} \right).$$

On the other hand, when $n \le y$,

$$C_n \left(\frac{1}{1 - A f_\infty^{(b)}(z)} \right) = C_n \left(\frac{1}{1 - A f_y^{(b)}(z)} \right)$$

and thus, for $n \le y$,

$$\sum_{k=1}^{\infty} A^k \left(F_k^{(b)}(n) - F_k^{(b)}(n-1) \right) = C_n \left(\frac{1}{1 - A f_y^{(b)}(z)} \right). \tag{4.21}$$

Using (4.21) with $n = y$, one can easily verify

$$\mathrm{Res} \left(\frac{z^{-y-1}}{1 - A f_y^{(b)}(z)}, 0 \right) = \sum_{k=1}^{\infty} A^k \left(F_k^{(b)}(y) - F_k^{(b)}(y-1) \right).$$

The pole of the function $z^{-y-1}/(1 - A f_y^{(b)}(z))$ in $z = \lambda_y^{(b)}(A)$ is of order 1. Therefore, it is not hard to see that

$$\mathrm{Res} \left(\frac{z^{-y-1}}{1 - A f_y^{(b)}(z)}, \lambda_y^{(b)}(A) \right) = - \frac{\lambda_y^{(b)}(A)^{-y-1}}{A \mu_y^{(b)}(\lambda_y^{(b)}(A))}$$

and by combining the latter results we obtain

$$\sum_{k=1}^{\infty} A^k \left(F_k^{(b)}(y) - F_k^{(b)}(y-1) \right) = \frac{\left(\lambda_y^{(b)}(A)\right)^{-y-1}}{A \mu_y^{(b)}(\lambda_y^{(b)}(A))} + \frac{1}{2\pi i} \int_{\gamma_y} \frac{z^{-y-1}}{1 - A f_y^{(b)}(z)} dz.$$

It remains to show that under the conditions of Proposition 4.2,

$$\frac{1}{2\pi i} \int_{\gamma_y} \frac{z^{-y-1}}{1 - A f_y^{(b)}(z)} dz = o\left(y^{-\min\{1, s-1\}} \ln y \right) \tag{4.22}$$

uniformly in $b \le b_2$ and $A_y \le A \le 1$. Let

$$\varphi_y^{(b)}(z) = A(f_y^{(b)}(z) - f_y^{(b)}(r_y)) - A\mu_y^{(b)}(r_y)(z - r_y),$$
$$\psi_y^{(b)}(z) = 1 - A f_y^{(b)}(r_y) - A\mu_y^{(b)}(r_y)(z - r_y).$$

Then, the following identity holds:

$$\frac{1}{1 - Af_y^{(b)}(z)} - \frac{1}{\psi_y^{(b)}(z)} = \frac{\psi_y^{(b)}(z) - 1 + Af_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)} = \frac{\varphi_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)}. \quad (4.23)$$

Let $\varepsilon > 0$, $\gamma_y(\varepsilon) = \gamma_y \cap U_y(\varepsilon)$ and let $\overline{\gamma}_y(\varepsilon)$ be the complement of $\gamma_y(\varepsilon)$ with respect to γ_y. By (4.23),

$$\begin{aligned}\int_{\gamma_y} \frac{z^{-y-1}}{1 - Af_y^{(b)}(z)} dz = \int_{\gamma_y} \frac{z^{-y-1}}{\psi_y^{(b)}(z)} dz &+ \int_{\gamma_y(\varepsilon)} \frac{z^{-y-1}\varphi_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)} dz \\ &+ \int_{\overline{\gamma}_y(\varepsilon)} \frac{z^{-y-1}\varphi_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)} dz.\end{aligned}$$

Using (4.23) once again, the last integral of the latter identity can be rewritten as

$$\int_{\overline{\gamma}_y(\varepsilon)} \frac{z^{-y-1}\varphi_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)} dz = -\int_{\overline{\gamma}_y(\varepsilon)} \frac{z^{-y-1}}{\psi_y^{(b)}(z)} dz + \int_{\overline{\gamma}_y(\varepsilon)} \frac{z^{-y-1}}{1 - Af_y^{(b)}(z)} dz.$$

Hence,

$$\int_{\gamma_y} \frac{z^{-y-1}}{1 - Af_y^{(b)}(z)} dz = I_1^{(b)}(y) + \sum_{j=2}^{4} I_j^{(b)}(y, \varepsilon), \quad (4.24)$$

where

$$I_1^{(b)}(y) = \int_{\gamma_y} \frac{z^{-y-1}}{\psi_y^{(b)}(z)} dz, \quad I_2^{(b)}(y, \varepsilon) = \int_{\gamma_y(\varepsilon)} \frac{z^{-y-1}\varphi_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)} dz,$$
$$I_3^{(b)}(y, \varepsilon) = -\int_{\overline{\gamma}_y(\varepsilon)} \frac{z^{-y-1}}{\psi_y^{(b)}(z)} dz, \quad I_4^{(b)}(y, \varepsilon) = \int_{\overline{\gamma}_y(\varepsilon)} \frac{z^{-y-1}}{1 - Af_y^{(b)}(z)} dz.$$

To calculate $I_1^{(b)}$ let us examine integrals of the form

$$\int_{|z|=c^2} \frac{z^{-n}}{kz + h} dz, \quad (4.25)$$

where $n > 0$, $k, h \in \mathbb{C}$ and $|h| < c^2|k|$. For $|h| < c^2|k|$, the function $z^{-n}/(kz + h)$ has exactly two singularities in the disc $|z| \leq c^2$, one in 0 and the other in $-h/k$. Consequently the Residue theorem states that

$$\int_{|z|=c^2} \frac{z^{-n}}{kz + h} dz = \text{Res}\left(\frac{z^{-n}}{kz + h}, 0\right) + Res\left(\frac{z^{-n}}{kz + h}, -\frac{h}{k}\right).$$

The pole in $z = 0$ has order n, hence

$$\text{Res}\left(\frac{z^{-n}}{kz + h}, 0\right) = (-1)^{n-1}k^{n-1}h^{-n}$$

and the pole in $z = -h/k$ is of order 1, thus

$$\operatorname{Res}\left(\frac{z^{-n}}{kz+h}, -\frac{h}{k}\right) = (-1)^n k^{n-1} h^{-n}.$$

Therefore,

$$\int_{|z|=c^2} \frac{z^{-n}}{kz+h} dz = [(-1)^{n-1} + (-1)^n] k^{n-1} h^{-n} = 0. \tag{4.26}$$

By the equicontinuity of $\mu_y^{(b)}(\cdot)$, the result from (4.14), Lemma 4.3 and Lemma 4.4, as $y \to \infty$,

$$\begin{aligned} f_y^{(b)}(r_y) - f_y^{(b)}(\lambda_y^{(b)}(A)) &= (r_y - \lambda_y^{(b)}(A))\mu^{(b)}(\lambda_y^{(b)}(A)) + o(r_y - \lambda_y^{(b)}(A)) \\ &= (r_y - \lambda_y^{(b)}(A))\mu^{(b)} + o(r_y - \lambda_y^{(b)}(A)) \end{aligned} \tag{4.27}$$

uniformly in $b \le b_2$ and $A \in \mathfrak{A}_y$. By virtue of Lemma 4.4 and the definition of C_1, $|\lambda_y^{(b)}(A)| \le e^{h_y - 1/y}$ and consequently

$$\begin{aligned} r_y - \lambda_y^{(b)}(A) &\ge e^{h_y}(1 - e^{-1/y}) \\ &= (1 + h_y + o(y^{-1}))(y^{-1} + o(y^{-1})) = y^{-1} + o(y^{-1}) \end{aligned}$$

uniformly in $b \le b_2$ and $A \in \mathfrak{A}_y$. By plugging these results into (4.27),

$$1 - Af_y^{(b)}(r_y) \le -\frac{A\mu^{(b)}}{y} + o\left(\frac{1}{y}\right) < 0 \tag{4.28}$$

for y large enough. Now put $h = 1 - Af_y^{(b)}(r_y) + A\mu_y^{(b)}(r_y)r_y$ and $k = -A\mu_y^{(b)}(r_y)$. Then, since $A\mu_y^{(b)}(r_y)r_y \ge A\mu_y^{(b)}(1) \ne o(1)$, we obtain by virtue of (4.28),

$$|h| \le A\mu_y^{(b)}(r_y)r_y = |k|r_y$$

and consequently by (4.26),

$$I_1^{(b)}(y) = \int_{\gamma_y} \frac{z^{-y-1}}{\psi_y^{(b)}(z)} dz = 0. \tag{4.29}$$

Let us consider $I_2^{(b)}$. Clearly,

$$I_2^{(b)}(y, \varepsilon) = i r_y^{-y} \int_{-\varepsilon}^{\varepsilon} \frac{\varphi_y^{(b)}(r_y e^{it})}{(1 - Af_y^{(b)}(r_y e^{it}))\psi_y^{(b)}(r_y e^{it})} e^{-ity} dt.$$

To bound this integral we use a method similar to the method Taibleson [47] used to bound Fourier coefficients. Denote by g a continuous function with bounded variation on an interval $[\theta_1, \theta_2]$ with $\theta_1, \theta_2 \in \mathbb{R}$ and $\theta_1 < \theta_2$. Let $n\theta_1/(2\pi), n\theta_2/(2\pi) \in \mathbb{Z}$ and

$a_k := 2\pi k/n$ for $k = n\theta_1/(2\pi), n\theta_1/(2\pi)+1, \ldots, n\theta_2/(2\pi)$. Then, $a_k - a_{k-1} = 2\pi/n$ for all k and consequently

$$\int_{a_{k-1}}^{a_k} e^{-itn} dt = -\frac{1}{in} e^{-ina_k} \left(1 - e^{2\pi i}\right) = 0$$

for all k. Hence,

$$\begin{aligned}
\int_{\theta_1}^{\theta_2} g(t)e^{-int} dt &= \sum_{k=n\theta_1/(2\pi)+1}^{n\theta_2/(2\pi)} \int_{a_{k-1}}^{a_k} g(t)e^{-int} dt \\
&= \sum_{k=n\theta_1/(2\pi)+1}^{n\theta_2/(2\pi)} \left(\int_{a_{k-1}}^{a_k} (g(t) - g(a_k))e^{-int} dt + g(a_k) \int_{a_{k-1}}^{a_k} e^{-itn} dt \right) \\
&= \sum_{k=n\theta_1/(2\pi)+1}^{n\theta_2/(2\pi)} \int_{a_{k-1}}^{a_k} (g(t) - g(a_k))e^{-int} dt.
\end{aligned}$$

For $a < b$, $a, b \in \mathbb{R}$ denote by $\mathbf{V}_D(g(u))$ the total variation of g on D. Then, for $t \in [a_{k-1}, a_k]$,

$$|g(t) - g(a_{k-1})| \le \mathbf{V}_{[a_{k-1},a_k]}(g(u)).$$

and by recalling $a_{n\theta_1/(2\pi)} = \theta_1$ and $a_{n\theta_2/(2\pi)} = \theta_2$,

$$\begin{aligned}
\left| \int_{\theta_1}^{\theta_2} g(t)e^{-int} dt \right| &\le \sum_{k=n\theta_1/(2\pi)+1}^{n\theta_2/(2\pi)} \int_{a_{k-1}}^{a_k} |(g(t) - g(a_k))| dt \\
&\le \frac{2\pi}{n} \sum_{k=n\theta_1/(2\pi)+1}^{n\theta_2/(2\pi)} \mathbf{V}_{[a_{k-1},a_k]}(g(u)) \le \frac{2\pi}{n} \mathbf{V}_{[\theta_1,\theta_2]}(g(u)). \qquad (4.30)
\end{aligned}$$

To use the latter inequality to bound $I_2^{(b)}$, remark that since we want to consider the case $y \to \infty$ we can always assume $y\varepsilon/(2\pi) \in \mathbb{N}$ without loss of generality. Consequently, by using (4.30) with $n = y$ and $-\theta_1 = \theta_2 = \varepsilon$ one attains for every fixed ε and y,

$$|I_2^{(b)}(y, \varepsilon)| \le \frac{2\pi}{y} \mathbf{V}_{\gamma_y(\varepsilon)} \left(\frac{\varphi_y^{(b)}(z)}{(1 - Af_y^{(b)}(z))\psi_y^{(b)}(z)} \right). \qquad (4.31)$$

The variation of $\omega_y^{(b)}(z) := \varphi_y^{(b)}(z)/((1 - Af_y^{(b)}(z))\psi_y(z))$ on $\gamma_y(\varepsilon)$ can be rewritten as follows:

$$\begin{aligned}
\mathbf{V}_{\gamma_y(\varepsilon)} \left(\omega_y^{(b)}(z) \right) &= \mathbf{V}_{\gamma_y(\varepsilon)} \left(\mathrm{Re}(\omega_y^{(b)}(z)) \right) + \mathbf{V}_{\gamma_y(\varepsilon)} \left(\mathrm{Im}(\omega_y^{(b)}(z)) \right) \\
&= \int_{\gamma_y(\varepsilon)} \left(\left| \frac{d}{dl} \mathrm{Re}(\omega_y^{(b)}(z)) \right| + \left| \frac{d}{dl} \mathrm{Im}(\omega_y^{(b)}(z)) \right| \right) dl,
\end{aligned}$$

where dl is the differential of the arc along $\gamma_y(\varepsilon)$. Due to the binomial formula,

$$\left(\left|\frac{d}{dl}\mathrm{Re}(\omega_y^{(b)}(z))\right| + \left|\frac{d}{dl}\mathrm{Im}(\omega_y^{(b)}(z))\right|\right)^2 \le 2\left(\left|\frac{d}{dl}\mathrm{Re}(\omega_y^{(b)}(z))\right|^2 + \left|\frac{d}{dl}\mathrm{Im}(\omega_y^{(b)}(z))\right|^2\right)$$
$$= 2\left|\frac{d}{dz}\omega_y^{(b)}(z)\right|^2$$

and thus,

$$\mathbf{V}_{\gamma_y(\varepsilon)}\left(\frac{\varphi_y^{(b)}(z)}{(1-Af_y^{(b)}(z))\psi_y^{(b)}(z)}\right) \le \sqrt{2}\int_{\gamma_y(\varepsilon)}\left|\frac{d}{dz}\frac{\varphi_y^{(b)}(z)}{(1-Af_y^{(b)}(z))\psi_y^{(b)}(z)}\right| dz$$
$$\le \sqrt{2}\left(\int_{\gamma_y(\varepsilon)}\left|\frac{\psi_y^{(b)\prime}(z)\varphi_y^{(b)}(z)}{(1-Af_y^{(b)}(z))(\psi_y^{(b)}(z))^2}\right| dz + \int_{\gamma_y(\varepsilon)}\left|\frac{A\mu_y^{(b)}(z)\varphi_y^{(b)}(z)}{(1-Af_y^{(b)}(z))^2\psi_y^{(b)}(z)}\right| dz\right.$$
$$\left. + \int_{\gamma_y(\varepsilon)}\left|\frac{\varphi_y^{(b)\prime}(z)}{(1-Af_y^{(b)}(z))\psi_y^{(b)}(z)}\right| dz\right)$$
$$= \sqrt{2}(I_{21}^{(b)} + I_{22}^{(b)} + I_{23}^{(b)}). \tag{4.32}$$

Let us bound the terms appearing in the integrands of the integrals from the latter inequality. Using the definition of the complex absolute value, an easy calculation shows that

$$|Af_y^{(b)}(z)-1|^2 = A^2|f_y^{(b)}(z) - f_y^{(b)}(r_y)|^2 + (Af_y^{(b)}(r_y)-1)^2$$
$$- 2A(Af_y^{(b)}(r_y)-1)\mathrm{Re}(f_y^{(b)}(r_y) - f_y^{(b)}(z)).$$

By the equicontinuity of the family $\{f_y^{(b)} : b \in I, A \in \mathfrak{A}_y\}$ and Lemma 4.3 with $\delta = \varepsilon$ small enough, as $y \to \infty$,

$$|f_y^{(b)}(r_y) - f_y^{(b)}(z)| = |r_y - z|\mu^{(b)}(z) + o(r_y - z) \ge (1-\varepsilon)\mu^{(b)}|z - r_y| \tag{4.33}$$

and

$$|f_y^{(b)}(r_y) - f_y^{(b)}(z)| = |r_y - z|\mu^{(b)}(z) + o(r_y - z) \le (1+\varepsilon)\mu^{(b)}|z - r_y| \tag{4.34}$$

uniformly in $b \le b_2$ and $z \in U_y^{(b)}(\varepsilon)$. Furthermore, for all $z \in U_y(\varepsilon)$ with ε sufficiently small,

$$\mathrm{Re}(r_y - z) = \sin(\arg z)|z - r_y| \le \varepsilon|z - r_y|.$$

By virtue of (4.28), (4.33) and (4.34),

$$|Af_y^{(b)}(z)-1|^2 \ge (Af_y^{(b)}(r_y)-1)^2 + (1-\varepsilon)(\mu^{(b)})^2A^2|z-r_y|^2$$
$$- 2\varepsilon(1+\varepsilon)\mu^{(b)}A(Af_y^{(b)}(r_y)-1)|z-r_y|$$

and by the binomial formula,

$$2\mu^{(b)}A(Af_y^{(b)}(r_y)-1)|z-r_y| \le (\mu^{(b)})^2A^2|z-r_y|^2 + (Af_y^{(b)}(r_y)-1)^2.$$

Using the binomial formula once again, we obtain

$$\begin{aligned}|Af_y^{(b)}(z)-1|^2 &\ge (1-\varepsilon-\varepsilon(1+\varepsilon))\left[|1-Af_y^{(b)}(r_y)|^2+(\mu^{(b)})^2A^2|z-r_y|^2\right]\\ &\ge \frac{1-\varepsilon-\varepsilon(1+\varepsilon)}{2}\left[|1-Af_y^{(b)}(r_y)|+A\mu^{(b)}|z-r_y|\right]^2.\end{aligned}$$

Choose ε so small that $1-\varepsilon-\varepsilon(1+\varepsilon)\ge 1/2$. Then,

$$|Af_y^{(b)}(z)-1| \ge \frac{|Af_y^{(b)}(r_y)-1|}{2}+\frac{A\mu^{(b)}|z-r_y|}{2} \tag{4.35}$$

uniformly in $b\le b_2$ and $z\in U_y^{(b)}(\varepsilon)$. We proceed analogously to bound $|\psi_y^{(b)}(z)|$ for $z\in U_y(\varepsilon)$ from below. One has $\mathrm{Re}(r_y-z)\le|z-r_y|$ and by virtue of Lemma 4.3, $\mu_y^{(b)}(r_y)\in[(1-\widehat{\delta}_1)\mu^{(b)},(1+\widehat{\delta}_1)\mu^{(b)}]$ for arbitrary $\widehat{\delta}_1$ if y is large enough and ε is small enough. Consequently, one can easily see that for $\widehat{\delta}_1$ small enough,

$$\begin{aligned}|\psi_y^{(b)}(z)|^2 &= |1-f_y^{(b)}(r_y)|^2+A^2\left(\mu_y^{(b)}(r_y)\right)^2|z-r_y|^2\\ &\qquad -2A(f_y^{(b)}(r_y)-1)\mu_y^{(b)}(r_y)\mathrm{Re}(r_y-z)\\ &\ge \frac{1-\widehat{\delta}_2}{2}\left[|1-f_y^{(b)}(r_y)|+A\mu^{(b)}|z-r_y|\right]^2 \end{aligned}\tag{4.36}$$

for all $\widehat{\delta}_2\le 1/2$. Hence,

$$|\psi_y^{(b)}(z)| \ge \frac{|1-Af_y^{(b)}(r_y)|}{2}+\frac{A\mu^{(b)}|z-r_y|}{2}. \tag{4.37}$$

On the other hand, one can easily see that for every z on $\gamma_y(\varepsilon)$ with ε sufficiently small,

$$\begin{aligned}|z-r_y| \ge |e^{i\arg z}-1| &= \sqrt{\sin^2(\arg z)+(1-\cos(\arg z))^2}\\ &= \sqrt{2-2\cos(\arg z)} \ge \frac{|\arg z|}{2},\end{aligned}\tag{4.38}$$

where we used $\cos\varphi\le 1-\varphi^2/8$ in the last inequality. Combining inequalities (4.28) and (4.38) with (4.35), we obtain

$$|1-Af_y^{(b)}(z)| \ge \frac{A\mu^{(b)}}{4}\left(\frac{1}{y}+|\arg z|\right). \tag{4.39}$$

for $b\le b_2$ and $z\in U_y^{(b)}(\varepsilon)$. The inequalities (4.28), (4.38) and (4.37) provide

$$|\psi_y^{(b)}(z)| \ge \frac{A\mu^{(b)}}{4}\left(\frac{1}{y}+|\arg z|\right) \tag{4.40}$$

and, moreover, an easy calculation shows

$$|\psi_y^{(b)\prime}(z)| = A\mu_y^{(b)}(r_y) \leq e^{h_y y} A\mu^{(b)}. \tag{4.41}$$

For $z \in U_y(\varepsilon)$,

$$|f_y^{(b)\prime\prime}(z)| \leq \begin{cases} e^{h_y y}\mathbf{E}(Z^{(b)})^2 & : s \geq 2 \\ e^{h_y y} y^{2-s}\mathbf{E}(Z^{(b)})^s & : 1 < s < 2 \end{cases}$$

and, consequently,

$$\frac{\varphi_y^{(b)}(z)}{|z - r_y|^2 y^{\max\{0,2-s\}}} \sim \frac{\varphi_y^{(b)\prime}(z)}{2|z - r_y| y^{\max\{0,2-s\}}} \sim \frac{A f_y^{(b)\prime\prime}(z)}{2y^{\max\{0,2-s\}}} = O(1)$$

as $y \to \infty$. By virtue of (4.38),

$$|z - r_y| = |z||e^{i \arg z} - 1| = \sqrt{2 - 2\cos(\arg z)} \leq \arg z,$$

for all $z \in \gamma_y(\varepsilon)$ if ε is sufficiently small. Hence, if ε is is sufficiently small,

$$\varphi_y^{(b)}(z) = O(y^{\max\{0,2-s\}}|z - r_y|^2) = O(y^{\max\{0,2-s\}} \arg^2(z)) \tag{4.42}$$

and

$$\varphi_y^{(b)\prime}(z) = O(y^{\max\{0,2-s\}}|z - r_y|) = O(y^{\max\{0,2-s\}}|\arg(z)|) \tag{4.43}$$

uniformly in $b \leq b_2$ and $A \in \mathfrak{A}_y$. Considering (4.39), (4.40), (4.41), (4.42) and $h_y y = O(1)$ provides

$$|I_{21}^{(b)}| \leq r_y \int_{-\varepsilon}^{\varepsilon} \frac{|\psi_y^{(b)\prime}(r_y e^{it})||\varphi_y^{(b)}(r_y e^{it})|}{|f_y^{(b)}(r_y e^{it}) - 1||\psi_y^{(b)}(r_y e^{it})|^2} dt = O\left(y^{\max\{0,2-s\}} \int_0^{\varepsilon} \frac{t^2}{(y^{-1} + t)^3} dt\right)$$

uniformly in $b \leq b_2$ and $A \in \mathfrak{A}_y$. Moreover,

$$\begin{aligned} \int_0^{\varepsilon} \frac{t^2}{(y^{-1} + t)^3} dt &= \int_{1/y}^{\varepsilon + 1/y} \frac{(w - y^{-1})^2}{w^3} dw \\ &\sim \ln(\varepsilon + y^{-1}) - \ln(y^{-1}) = \ln(1 + \varepsilon y) \sim \ln(y) \end{aligned} \tag{4.44}$$

as $y \to \infty$ and therefore, uniformly in $b \leq b_2$ and $A \in \mathfrak{A}_y$,

$$|I_{21}^{(b)}| = O(y^{\max\{0,2-s\}} \ln y). \tag{4.45}$$

In analogy, by additionally taking into account that $\mu_y^{(b)}(z) \leq 2\mu^{(b)}$ due to Lemma 4.3 for y large enough, one can easily see that

$$I_{22}^{(b)} = O(y^{\max\{0,2-s\}} \ln y) \tag{4.46}$$

and furthermore, by regarding (4.43),

$$I_{23}^{(b)} = O\left(y^{\max\{0,2-s\}} \int_0^{\varepsilon} \frac{t}{(y^{-1} + t)^2} dt\right) = O(y^{\max\{0,2-s\}} \ln y). \tag{4.47}$$

Finally, plugging (4.45), (4.46) and (4.47) into (4.32) we attain

$$\mathbf{V}_{\gamma_y(\varepsilon)}\left(\frac{\varphi_y^{(b)}(z)}{(1-Af_y^{(b)}(z))\psi_y^{(b)}(z)}\right) = O(y^{\max\{0,2-s\}}\ln y)$$

and hence by (4.31),

$$|I_2^{(b)}(y,\varepsilon)| = O\big(y^{\max\{-1,-(s-1)\}}\ln y\big) \tag{4.48}$$

uniformly in $b \le b_2$ and the admissible values of A. Next, we draw our attention to the integral $I_3^{(b)}$.

$$I_3^{(b)}(y,\varepsilon) = -ir_y^{-y}\int_{\varepsilon\le|t|\le\pi}\frac{e^{-iyt}}{\psi_y^{(b)}(r_ye^{it})}dt. \tag{4.49}$$

To bound this integral we use the bound from (4.30) again:

$$\int_{\varepsilon\le|t|\le\pi}\frac{e^{-iyt}}{\psi_y^{(b)}(r_ye^{it})}dt \le \frac{2\pi}{y}\mathbf{V}_{\overline{\gamma}_y(\varepsilon)}\left(\frac{1}{\psi_y^{(b)}(z)}\right). \tag{4.50}$$

In analogy to (4.32), one can show that

$$\mathbf{V}_{\overline{\gamma}_y(\varepsilon)}\left(\frac{1}{\psi_y^{(b)}(z)}\right) \le \sqrt{2}\int_{\overline{\gamma}_y(\varepsilon)}\left|\frac{d}{dz}\frac{1}{\psi_y^{(b)}(z)}\right|dz = \sqrt{2}\int_{\overline{\gamma}_y(\varepsilon)}\frac{|\psi_y^{(b)\prime}(z)|}{|\psi_y^{(b)}(z)|^2}dz.$$

For all $z \in \overline{\gamma}_y(\varepsilon)$ the inequality (4.37) gives

$$|\psi_y^{(b)}(z)|^2 \ge \frac{A^2(\mu^{(b)})^2}{4}|z-r_y|^2 \ge \frac{A^2(\mu^{(b)})^2}{4}\varepsilon^2,$$

where we used that $|z-r_y| \ge \varepsilon$ for all $z \in \overline{\gamma}_y(\varepsilon)$. Therefore, by virtue of (4.41) and $h_yy = O(1)$,

$$\mathbf{V}_{\overline{\gamma}_y(\varepsilon)}\left(\frac{1}{\psi_y^{(b)}(z)}\right) = O(1)$$

and consequently by combining this result with (4.49), (4.50) and $h_yy = O(1)$,

$$|I_3^{(b)}(y,\varepsilon)| = O\left(\frac{1}{y}\right) \tag{4.51}$$

uniform in $b \le b_2$ and $A \in \mathfrak{A}_y$. It remains to consider $I_4^{(b)}$. Due to (4.30),

$$|I_4^{(b)}(y,\varepsilon)| = \left|ir_y^{-y}\int_{\varepsilon\le|t|\le\pi}\frac{e^{-iyt}}{1-Af_y^{(b)}(r_ye^{it})}dt\right| \le \frac{2\pi}{y}\mathbf{V}_{\overline{\gamma}_y(\varepsilon)}\left(\frac{1}{1-Af_y^{(b)}(z)}\right). \tag{4.52}$$

Furthermore, by (4.14) and (4.19),

$$\begin{aligned}\mathbf{V}_{\overline{\gamma}_y(\varepsilon)}\left(\frac{1}{1-Af_y^{(b)}(z)}\right) &\le \sqrt{2}\int_{\overline{\gamma}_y(\varepsilon)}\left|\frac{d}{dz}\frac{1}{1-Af_y^{(b)}(z)}\right|dz\\ &= \sqrt{2}\int_{\overline{\gamma}_y(\varepsilon)}\frac{A|\mu_y^{(b)}(z)|}{|1-Af_y^{(b)}(z)|^2}dz = O(1)\end{aligned}$$

and consequently

$$I_4^{(b)}(y,\varepsilon) = O\left(\frac{1}{y}\right). \tag{4.53}$$

Finally, by plugging the results attained in (4.29), (4.48), (4.51) and (4.53) into (4.24), we get

$$\int_{\gamma_y} \frac{z^{-y-1}}{1 - Af_y^{(b)}(z)} dz = O(y^{\max\{-1,-(s-1)\}} \ln y) + O(y^{-1}) = O(y^{-\min\{1,s-1\}} \ln y)$$

uniformly in $0 \le b \le b_2$ and $A_y \le A \le 1$.

4.4 Proof of the local limit theorem

Put $\tau_{+,0}^{(a)} = 0$ and define recursively for $i \ge 1$ the i-th strict ascending ladder epoch of the random walk $S^{(a)}$ and its corresponding ladder height by

$$\tau_{+,i}^{(a)} := \min\{k \ge \tau_{+,i-1}^{(a)} : S_k^{(a)} > S_{\tau_{+,i-1}}^{(a)}\} \quad \text{and} \quad \chi_i^{(a)} = S^{(a)}_{\tau_{+,i}^{(a)}} - S^{(a)}_{\tau_{+,i-1}^{(a)}}.$$

In the case $i = 1$ we write $\tau_+^{(a)}$ and $\chi^{(a)}$ instead of $\tau_{+,1}^{(a)}$ and $\chi_1^{(a)}$ respectively and, if additionally $a = 0$, we write τ_+ and χ instead of $\tau_+^{(0)}$ and $\chi^{(0)}$ respectively. Define random variables $Z_i^{(a)}$ as *iid* copies of a random variable $Z^{(a)}$ with

$$\mathbf{P}(Z^{(a)} \in \cdot) = \mathbf{P}(\chi_1^{(a)} \in \cdot | \tau_+^{(a)} < \infty).$$

Denote by $\theta := \min\{k \ge 0 : S_k^{(a)} = M^{(a)}\}$ the first time the random walk reaches its maximum. Then,

$$\mathbf{P}(M^{(a)} = y) = \sum_{n=1}^{\infty} \mathbf{P}(M^{(a)} = y, \theta = n).$$

We further define $M_n^{(a)} := \max_{k \le n} S_k^{(a)}$ and $\theta_n := \min\{k \le n : S_k^{(a)} = M_n^{(a)}\}$. By the Markov property,

$$\mathbf{P}(M^{(a)} = y, \theta = n) = \mathbf{P}(S_n^{(a)} = y, \theta_n = n)\mathbf{P}(\tau_+^{(a)} = \infty).$$

Hence the following representation holds for the maximum:

$$\mathbf{P}(M^{(a)} = y) = \mathbf{P}(\tau_+^{(a)} = \infty) \sum_{n=1}^{\infty} \mathbf{P}(S_n^{(a)} = y, \theta_n = n). \tag{4.54}$$

Clearly,

$$\begin{aligned} \mathbf{P}(S_n^{(a)} = y, \theta_n = n) &= \mathbf{P}(S_n^{(a)} = y, n \text{ is a strict ascending ladder epoch}) \\ &= \sum_{k=1}^{\infty} \mathbf{P}(\chi_1^{(a)} + \chi_2^{(a)} + \dots + \chi_k^{(a)} = y, \tau_{+,1}^{(a)} + \tau_{+,2}^{(a)} + \dots + \tau_{+,k}^{(a)} = n). \end{aligned} \tag{4.55}$$

Denote the distribution function of $Z^{(a)}$ by G and denote the expectation by $\mu^{(a)} := \mathbf{E}[Z^{(a)}]$. Let G^{*k} be the k-fold convolution of G with itself. Then, by using (4.55), changing the order of summation and using the Markov property,

$$\begin{aligned}
&\sum_{n=1}^{\infty}\mathbf{P}(S_n^{(a)} = y, \theta_n = n)\\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty}\mathbf{P}(\chi_1^{(a)} + \chi_2^{(a)} + \cdots + \chi_k^{(a)} = y, \tau_{+,1}^{(a)} + \tau_{+,2}^{(a)} + \cdots + \tau_{+,k}^{(a)} = n)\\
&= \sum_{k=1}^{\infty}\mathbf{P}(\chi_1^{(a)} + \chi_2^{(a)} + \cdots + \chi_k^{(a)} = y | \tau_{+,k}^{(a)} < \infty)\mathbf{P}(\tau_{+,k}^{(a)} < \infty)\\
&= \sum_{k=1}^{\infty} A^k \left(G^{*k}(y) - G^{*k}(y-1)\right) \qquad (4.56)
\end{aligned}$$

with $A = \mathbf{P}(\tau_+^{(a)} < \infty)$. Combining results (4.54) and (4.56) we attain

$$\mathbf{P}(M^{(a)} = y) = \mathbf{P}(\tau_+^{(a)} = \infty)\sum_{k=1}^{\infty} A^k \left(G^{*k}(y) - G^{*k}(y-1)\right). \qquad (4.57)$$

Next, we want to use Proposition 4.2 to determine the asymptotic behaviour of the sum on the right hand side of the latter equality. Therefore, let us first show that under the assumptions of Theorem 4.1,

$$Z^{(a)} \xrightarrow{w} Z^{(0)} \quad \text{as } a \to 0. \qquad (4.58)$$

It is known that

$$\mathbf{P}(\tau_+^{(a)} < \infty) \sim \mathbf{P}(\tau_+ < \infty) = 1. \qquad (4.59)$$

Thus, as $a \to 0$,

$$\mathbf{P}(Z^{(a)} > x) = \frac{\mathbf{P}(\chi^{(a)} > x, \tau_+^{(a)} < \infty)}{\mathbf{P}(\tau_+^{(a)} < \infty)} \sim \mathbf{P}(\chi^{(a)} > x, \tau_+^{(a)} < \infty)$$

and, on the other hand, (4.2) and (4.59) imply that for every $R > 0$, as $a \to 0$,

$$\begin{aligned}
\mathbf{P}(\chi^{(a)} > x, R < \tau_+^{(a)} < \infty) &\le \mathbf{P}(R < \tau_+^{(a)} < \infty)\\
&= \mathbf{P}(\tau_+^{(a)} < \infty) - \mathbf{P}(\tau_+^{(a)} \le R) \sim \mathbf{P}(\tau_+ > R).
\end{aligned}$$

Furthermore, by using (4.2) and the continuous mapping theorem,

$$\begin{aligned}
\mathbf{P}(\chi^{(a)} > x, \tau_+^{(a)} \le R) &= \sum_{k=0}^{R-1}\mathbf{P}\big(S_{k+1}^{(a)} > x, \max_{1\le l\le k} S_l^{(a)} \le 0\big)\\
&\sim \sum_{k=0}^{R-1}\mathbf{P}\big(S_{k+1} > x, \max_{1\le l\le k} S_l \le 0\big) = \mathbf{P}(\chi > x, \tau_+ \le R)
\end{aligned}$$

as $a \to 0$. Thus,

$$\limsup_{a\to 0} \mathbf{P}(Z^{(a)} > x) \le \mathbf{P}(\chi > x, \tau_+ \le R) + \mathbf{P}(\tau_+ > R)$$

and by letting $R \to \infty$ we conclude

$$\limsup_{a\to 0} \mathbf{P}(Z^{(a)} > x) \le \mathbf{P}(\chi > x, \tau_+ < \infty) = \mathbf{P}(Z^{(0)} > x).$$

On the other side, the above calculations give

$$\liminf_{a\to 0} \mathbf{P}(Z^{(a)} > x) \ge \liminf_{a\to 0} \mathbf{P}(\chi^{(a)} > x, \tau_+^{(a)} \le R) = \mathbf{P}(\chi > x, \tau_+ \le R)$$

and by letting $R \to \infty$,

$$\liminf_{a\to 0} \mathbf{P}(Z^{(a)} > x) \ge \mathbf{P}(\chi > x, \tau_+ < \infty) = \mathbf{P}(Z^{(0)} > x).$$

This means that (4.58) holds under our assumptions.

Due to relation (16) of Chow [16] there exists a constant $C_1 > 0$ such that

$$\mathbf{E}[\big(S^{(a)}_{\tau_+^{(a)}}\big)^{1+\varepsilon}; \tau_+^{(a)} < \infty] \le C_1 \int_0^\infty \frac{u^{2+\varepsilon}}{\mathbf{E}\big[|S^{(a)}_{\tau_-^{(a)}}| \wedge u\big]} d\mathbf{P}(\max\{0, X^{(a)}\} < u),$$

where $\tau_-^{(a)} = \min\{k \ge 1 : S_k^{(a)} \le 0\}$ is the first weak descending ladder epoch. Obviously,

$$\mathbf{E}\big[|S^{(a)}_{\tau_-^{(a)}}| \wedge u\big] \ge \mathbf{E}\big[|S^{(a)}_{\tau_-^{(a)}}| \wedge 1\big] \ge \mathbf{P}(S_1^{(a)} < 0) > 0$$

for all $u \ge 1$ and therefore

$$\mathbf{E}[\big(S^{(a)}_{\tau_+^{(a)}}\big)^{1+\varepsilon}; \tau_+^{(a)} < \infty] \le \frac{C_1}{\mathbf{P}(S_1 < 0)} \int_0^\infty u^{2+\varepsilon} d\mathbf{P}(\max\{0, X^{(a)}\} < u).$$

Hence, by virtue of (4.3),

$$\sup_{a \le a_0} \mathbf{E}[(Z^{(a)})^{1+\varepsilon}] < \infty. \tag{4.60}$$

The convergence from (4.58) combined with (4.60) implies

$$\mu^{(a)} \to \mu^{(0)} \quad \text{as } a \to 0 \tag{4.61}$$

by dominated convergence. It is known that for all $a > 0$ the stopping time $\tau_+^{(a)}$ is infinite with positive probability and that

$$\mathbf{P}(\tau_+^{(a)} = \infty) = 1/\mathbf{E}[\tau_-^{(a)}]. \tag{4.62}$$

Totally analogously to (4.60), one can use (15) from Chow [16] to show that the existence of the second moment in assumption (4.3) implies $\sup_{a\le a_0} \mathbf{E}[S^{(a)}_{\tau_-^{(a)}}] < \infty$. Hence, one can use dominated convergence to show that

$$\mathbf{E}[S^{(a)}_{\tau_-^{(a)}}] \to \mathbf{E}[S_{\tau_-^{(0)}}] \quad \text{as } a \to 0.$$

Thus, using (4.62), the known identity

$$\frac{\sigma^2}{2} = -\mu^{(0)}\mathbf{E}[S_{\tau_-^{(0)}}] \tag{4.63}$$

and Wald's identity imply that

$$\mathbf{P}(\tau_+^{(a)} = \infty) = \frac{1}{\mathbf{E}[\tau_-^{(a)}]} \sim \frac{a}{-\mathbf{E}[S_{\tau_-^{(0)}}]} \sim \frac{2a\mu^{(0)}}{\sigma^2}. \tag{4.64}$$

The assumption $ay = O(1)$ implies the existence of a constant C such that $a \le C/y$. Therefore, by (4.64),

$$\mathbf{P}(\tau_+^{(a)} < \infty) \ge 1 - \frac{3C\mu^{(0)}}{\sigma^2 y} \tag{4.65}$$

for a small enough. Summing up the results from (4.60) and (4.65), this means that we can apply Proposition 4.2 for $I = \{\mu^{(a)} : 0 \le a \le a_0\}$ with $a_0 > 0$ small enough, $A_y = 1 - 3C\mu^{(0)}/(\sigma^2 y)$, $A = \mathbf{P}(\tau_+^{(a)} < \infty)$ and $s = 1 + \varepsilon$. Hence,

$$\sum_{k=1}^{\infty} A^k \left(G^{*k}(y) - G^{*k}(y-1)\right) = \frac{\left(\lambda_y^{(a)}(A)\right)^{-y-1}}{A\mu_y^{(a)}(\lambda_y^{(a)}(A))} + O(y^{-\min\{1,\varepsilon\}} \ln y) \tag{4.66}$$

and consequently, by combining equations (4.57), (4.66) and the fact that $1 - A = O(a)$, we attain

$$\mathbf{P}(M^{(a)} = y) = (1-A)\frac{\left(\lambda_y^{(a)}(A)\right)^{-y-1}}{A\mu_y^{(a)}(\lambda_y^{(a)}(A))} + o(ay^{-\min\{1,\varepsilon\}} \ln y). \tag{4.67}$$

Let us determine $\lambda_y^{(a)}(A)$ and $\mu_y^{(a)}(\lambda_y^{(a)}(A))$. Write λ_y and $\mu_y(\lambda_y)$ instead of $\lambda_y^{(a)}(A)$ and $\mu_y^{(a)}(\lambda_y^{(a)}(A))$ respectively for abbreviation and put $\lambda_y = e^{\theta_y}$. According to the definition of λ_y, we want to find θ_y such that

$$\mathbf{E}[\exp\{\theta_y Z^{(a)}\}; Z^{(a)} \le y] = \frac{1}{A}. \tag{4.68}$$

It turns out we don't need an exact solution for this equation and it is sufficient to determine θ_y such that

$$\mathbf{E}[\exp\{\theta_y Z^{(a)}\}; Z^{(a)} \le y] = \frac{1}{A} + O(y^{-1-\varepsilon}). \tag{4.69}$$

By Taylor's formula,

$$\begin{aligned}
&\mathbf{E}[\exp\{\theta_y Z^{(a)}/\}; Z^{(a)} \le y] \\
&= 1 + \theta_y\mu^{(a)} - \mathbf{P}(Z^{(a)} > y) - \theta_y\mathbf{E}[Z^{(a)}; Z^{(a)} > y] \\
&\qquad\qquad + \frac{\theta_y^2}{2}\mathbf{E}[(Z^{(a)})^2 \exp\{\gamma\theta_y Z^{(a)}\}; Z^{(a)} \le y]
\end{aligned}$$

with some random $\gamma \in (-\infty, 1]$. We restrict ourselves to θ_y such that $\theta_y = O(1/y)$. Then, (4.60) implies

$$\mathbf{P}(Z^{(a)} > y) + \theta_y \mathbf{E}[Z^{(a)}; Z^{(a)} > y] = O(y^{-1-\varepsilon})$$

and

$$\frac{\theta_y^2}{2}\mathbf{E}[(Z^{(a)})^2 \exp\{\gamma\theta_y Z^{(a)}\}; Z^{(a)} \le y] = O\left(\theta_y^2 E[(Z^{(a)})^2; Z^{(a)} \le y]\right) = O(y^{-1-\varepsilon}).$$

This means that to find θ_y that satisfies (4.69), it is sufficient to choose θ_y such that

$$1 + \theta_y \mu^{(a)} = \frac{1}{A} + O(y^{-1-\varepsilon})$$

or

$$\theta_y = \frac{1-A}{A\mu^{(a)}} + O(y^{-1-\varepsilon}).$$

Consequently,

$$\lambda_y = \exp\left\{\frac{1-A}{A\mu^{(a)}} + O(y^{-1-\varepsilon})\right\}. \tag{4.70}$$

Furthermore,

$$\begin{aligned}\mu_y^{(a)}(\lambda_y) &= \sum_{k=1}^{y} k f_k^{(a)} \lambda_y^{k-1} = \frac{1}{\lambda_y}\mathbf{E}[Z^{(a)} \exp\{\theta_y Z^{(a)}\}; Z^{(a)} \le y] \\ &= \lambda_y^{-1}\left\{\mathbf{E}[Z^{(a)}; Z^{(a)} \le y] + \theta_y \mathbf{E}[(Z^{(a)})^2 \exp\{\widetilde{\gamma}\theta_y Z^{(a)}\}; Z^{(a)} \le y]\right\}\end{aligned}$$

for some random $\widetilde{\gamma} \in (-\infty, 1]$. For all $\theta_y = O(1/y)$ the result (4.60) gives

$$\mathbf{E}[(Z^{(a)})^2 \exp\{\widetilde{\gamma}\theta_y Z^{(a)}\}; Z^{(a)} \le y] = O(y^{1-\varepsilon})$$

and

$$\mathbf{E}[Z^{(a)}; Z^{(a)} \le y] = \mu^{(a)} + O(y^{-\varepsilon}).$$

Consequently,

$$\mu_y^{(a)}(\lambda_y) = \frac{\mu^{(a)}}{\lambda_y} + O(y^{-\varepsilon}). \tag{4.71}$$

Plugging the results from (4.70) and (4.71) into the right hand side of (4.67), we obtain by regarding $1 - A = O(a)$,

$$\begin{aligned}&\mathbf{P}(M^{(a)} = y) \\ &= \frac{1-A}{A\mu^{(a)} + O(y^{-\varepsilon})}\exp\left\{-\frac{(1-A)y}{A\mu^{(a)}} + O(y^{-\varepsilon})\right\} + o(ay^{-\min\{1,\varepsilon\}}\ln y) \\ &= \frac{1-A}{A\mu^{(a)} + O(y^{-\varepsilon})}\exp\left\{-\frac{(1-A)y}{A\mu^{(a)}}\right\} + o(ay^{-\min\{1,\varepsilon\}}\ln y) \\ &= \frac{1-A}{A\mu^{(a)}}\exp\left\{-\frac{(1-A)y}{A\mu^{(a)}}\right\} + o(ay^{-\min\{1,\varepsilon\}}\ln y) + O(ay^{-\varepsilon})\end{aligned} \tag{4.72}$$

uniformly for all y such that $ay = O(1)$ as $a \to 0$. Here, we applied Taylor's formula in the last line. As a consequence of (4.59), (4.61) and (4.64),

$$\frac{1-A}{A\mu^{(a)}} = \frac{2a}{\sigma^2} + o(a)$$

and hence, by plugging this result into (4.72), we finally obtain

$$\mathbf{P}(M^{(a)} = y) \sim \frac{2a\Delta}{\sigma^2} \exp\left\{-\frac{2ay\Delta}{\sigma^2}\right\}$$

uniformly for all y such that $y \to \infty$ and $ya = O(1)$ as $a \to 0$.

Literatur

[1] Adler, J., Feldman, R. and Taqqu, M. A Practical Guide to Heavy Tails. *Birkhäuser,* 1998.

[2] Asmussen, S. Applied Probability and Queues. *2nd ed., Springer,* 2003.

[3] Asmussen, S. Ruin Probability. *World Scientific Publishing Co. Pte. Ltd.,* 2000.

[4] Asmussen, S. Subexponential asymptotics for stochastic processes: extremal behaviour, stationary distributions and first passage probabilities. *Ann. Appl. Probab.*, 8:354–374, 1998.

[5] Bingham, N.H., Goldie, C.M. and Teugels, J.L. Regular Variation. *Cambridge University Press,* 1987

[6] Blanchet, J. and Glynn, P. Uniform renewal theory with applications to expansions of random geometric sums. *Adv. in Appl. Probab.*, 4: 1070–1097, 2007.

[7] Blanchet, J. and Lam, D. Uniform large deviations for heavy-tailed queues under heavy traffic. *Bull. of the Mex. Math. Soc. (3)*, 19, 2013.

[8] Borovkov, A.A. Stochastic Processes in Queueing Theory. *Springer,* 1976.

[9] Borovkov, A.A. Wahrscheinlichkeitstheorie - Eine Einführung. *Akademie-Verlag,* 1976.

[10] Borovkov, A.A. On subexponential distributions and asymptotics of the distribution of the maximum of sequential sums. *Siberian Math. J.,* 43:995–1022, 2002.

[11] Borovkov, A.A. and Borovkov, K.A. Asymptotic Analysis of Random Walks. *Cambridge University Press,* 2008.

[12] Borovkov, A.A. and Foss, S.G. Estimates for the excess of a random walk over an arbitrary boundary and their applications. *Theory Probab. Appl.,* 44:231–253, 1998.

[13] Callaert, H. and Cohen, J.W. A lemma on regular variation of a transient renewal function. *Z. Wahrscheinlichkeitstheorie verw. Gebiete,* 24:275–278, 1972.

[14] Chistyakov, A. A theorem on sums of independent positive variables and its applications to branching processes. *Theory Probab. Appl.,* 9:640–648, 1964.

[15] Chover, J., Ney, P. and Wainger, S. Functions of probability measures. *J. Analyse Math.,* 26:255–302, 1973.

[16] Chow, Y.S. On moments of ladder height variables. *Adv. in Appl. Math.*, 7:46–54, 1986

[17] Cohen, J.W. Some results on regular variation for distributions in queueing and fluctuation theory. *J. Appl. Probab.,* 10:343–353, 1973.

[18] Denisov, D. A note on the asymptotics for the maximum on a random time interval of a random walk. *Markov Process. Related Fields*, 11:165–169 (2005).

[19] Denisov, D. and Kugler, J. Heavy traffic and heavy tails for subexponential distributions. *arXiv:* 1403.7325.

[20] Denisov, D., Wachtel V. Martingale approach to subexponential asymptotics for random walks. *Electr. Comm. in Probab.*, 17, Article 6, 1–9. 2012.

[21] Embrechts, P., Goldie, C.M. and Veraverbeke, N. Subexponentiality and infinite divisibility. *Z. Wahscheinlichkeitstheorie verw. Gebiete*, 49:335-347, 1979.

[22] Embrechts, P. and Hawkes, J. A limit theorem for the tails of discrete infinitely divisible laws with applications to fluctuation theory. *J. Austral. Math. Soc. Ser. A*, 32:412-422, 1982.

[23] Embrechts, P., Klüppelberg, C. and Mikosch, T. Modelling Extremal Events for Finance and Insurance. *Springer,* 1997.

[24] Embrechts, P. and Veraverbeke, N. Estimates for the probability of ruin with special emphasis on the possibility of large claims. *Insur. Math. Econ.*, 1:55-72, 1982.

[25] Feller, W. An Introduction to Probability Theory and Its Applications, volume II, 2nd ed. *John Wiley & Sons, Inc.,* 1971.

[26] Foss, S. and Korshunov, D. Lower limits and equivalences for convolution tails. *Ann. Probab.*, 35:366-383, 2007

[27] Foss, S., Korshunov, D. and Zachary, S. An Introduction to Heavy-Tailed and Subexponential Distributions. *Springer Science+Business Media*, 2011.

[28] Foss, S., Puhalskii, A.A. On the limit law of a random walk conditioned to reach a high level. *Stochastic Process. Appl.*, 121:288-313, 2011.

[29] Foss, S., Zachary, S. The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Ann. Appl. Probab.*, 13:37-53, 2003.

[30] Fuk, D.Kh. and Nagaev, S.V. Probability inequalities for sums of independent random variables. *Theory Probab. Appl.,* 4:643–660, 1971.

[31] Kalashnikov, V. Bounds for ruin probabilities in the presence of large claims and their comparison. *North Amer. Act. J.*, 3:116–129, 1999.

[32] Kiefer, J. and Wolfowitz, J. On the characteristics of the general queueing process with applications to random walks. *Trans. Amer. Math. Soc.*, 78:1-18, 1956.

[33] Kingman, J. The single server queue in heavy traffic. *Proc. Camb. Phil. Soc.*, 57:902-904, 1961.

[34] Klüppelberg, C. Subexponential distributions and integrated tails. *J. Appl. Probab.*, 25:132–141, 1988.

[35] Korshunov, D. The critical case of the Cramér-Lundberg theorem on the asymptotics of the distribution of the maximum of a random walk with negative drift. *Siberian Math. J.*, 46:1077–1081, 2005.

[36] Kugler, J. and Wachtel, V. Upper bounds for the maximum of a random walk with negative drift. *J. Appl. Probab.*, 50:1131–1146, 2013.

[37] Kugler, J. Local limit theorem for the maximum of a random walk. *arXiv:* 1403.7372.

[38] Lorden, G. On excess over the boundary. *Ann. Math. Statist.,* 41:520–527, 1970.

[39] Mikosch, T. and Nagaev, A.V. Large deviations of heavy-tailed sums with applications in insurance. *Extremes,* 1:81–110, 1998.

[40] Mogul'skii, A.A. Absolute estimates for moments of certain boundary functionals. *Theory Probab. Appl.,* 18:340–347, 1973.

[41] Nagaev, S.V. Some renewal theorems. *Theory Probab. Appl.,* 13:547–563, 1983.

[42] Nagaev, S.V. Large deviations of sums of independent random variables. *Ann. Probab.,* 5:745–789, 1979.

[43] Olvera-Cravioto, M., Blanchet, J. and Glynn, P. On the transition from heavy traffic to heavy tails for the $M/G/1$ queue: the regularly varying case. *Ann. Appl. Prob.*, 21:645–668, 2011.

[44] Olvera-Cravioto, M. and Glynn, P. Uniform approximations for the $M/G/1$ queue with subexponential processing times. *Queueing Syst.*, 68:1–50, 2011.

[45] Prohorov, Ju.V. Transition phenomena in queueing processes. I. (Russian) *Litovsk. Mat. Sb.*, 3:199-205, 1963.

[46] Richards, A. On upper bounds for the tail distribution of geometric sums of subexponential random variables. *Queueing Syst.*, 62:229–242, 2009.

[47] Taibleson, M. Fourier coefficients of functions of bounded variation. *Proc. of the Am. Math. Soc.,* Vol. 18, 4: 766, 1967.

[48] Veraverbeke, N. Asymptotic behaviour of Wiener-Hopf factors of a random walk. *Stochastic Process. Appl.,* 5:27–37, 1977.

[49] Wachtel, V. and Shneer, S. Heavy traffic analysis of the maximum of an asymptotically stable random walk. *Theory Probab. Appl.,* 55: 332–341, 2011.