Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Stochastic Processes and their Applications 120 (2010) 801–828
stochastic processes and their applications
www.elsevier.com/locate/spa

An extension of a logarithmic form of Cramér's ruin theorem to some FARIMA and related processes

Ph. Barbe[a], W.P. McCormick[b,*]

[a] CNRS (UMR8088), 90 rue de Vaugirard, 75006 Paris, France
[b] University of Georgia, Department of Statistics, Athens, GA 30602, United States

Received 9 January 2009; received in revised form 14 October 2009; accepted 9 January 2010
Available online 6 February 2010

Abstract

Cramér's theorem provides an estimate for the tail probability of the maximum of a random walk with negative drift and increments having a moment generating function finite in a neighborhood of the origin. The class of (g, F)-processes generalizes in a natural way random walks and fractional ARIMA models used in time series analysis. For those (g, F)-processes with negative drift, we obtain a logarithmic estimate of the tail probability of their maximum, under conditions comparable to Cramér's. Furthermore, we exhibit the most likely paths as well as the most likely behavior of the innovations leading to a large maximum.

MSC: primary 60F99; secondary 60G70; 60G50; 62P05; 60F10; 60K25; 60K99

Keywords: Maximum of random walk; Cramér's theorem; Fractional ARIMA process; Ruin probability; Large deviations

1. Introduction

Cramér's theorem on the maximum of a random walk with negative drift provides an estimate for the tail probability of this maximum when the moment generating function of the increments is finite in a neighborhood of the origin. Specifically, writing M for the maximum of the random walk, it asserts that there are constants c and θ such that

$$\mathrm{P}\{M > t\} \sim c\mathrm{e}^{-\theta t} \tag{1.1}$$

as t tends to infinity; the constants c and θ are explicit, but their formulas are irrelevant to the current discussion. We refer the reader to [22, Section XI.7] for a proof of Cramér's theorem.

* Corresponding author.
E-mail address: bill@stat.uga.edu (W.P. McCormick).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2010.02.001

The purpose of this paper is to make a first step toward an extension of Cramér's result to a wider class of stochastic processes which encompass some fractional ARIMA ones. As explained in [4] where we dealt with the analogous problem in the heavy tail context, the motivations are manifold. To summarize, besides the original application to insurance mathematics which motivated Cramér, other areas of applications exist, such as queueing theory — the connection between risk and queueing theory was pointed out in [41] (see e.g. [31] for an account on this connection); furthermore, on a more fundamental level, a certain analogy, described in Barbe and McCormick [4], has been developed between the asymptotic theory of the usual random walk and that of some FARIMA processes, and it is natural to investigate to what extent this analogy carries over in the context of Cramér's theorem.

Previous authors have considered extensions of Cramér's result to processes with dependent innovations. For instance, using a martingale technique, Gerber [25] considered bounded ARMA increments. His result was extended by Promislow [42] who removed the boundedness assumption and dealt with a larger class of increments. In contrast, using large deviations theory, building upon the work of Burton and Dehling [11] as well as Iscoe et al. [30], Nyrhinen [38–40] considered increments following a stationary linear process with some having a Markovian structure. Müller and Pflug [37] extended some of Nyrhinen's results by relating the asymptotic behavior of the moment generating function of the ruin process at time n, as n tends to infinity, to the behavior of its maximum, hence showing that the Gärtner–Ellis [24,21] approach in large deviations leads to a ruin probability estimate. A common feature of these works is that the processes under consideration exhibit short range dependence in order to have an explicit behavior of some moment generating functions. In contrast, the study of ruin probability associated with continuous time processes has recently focused on long range dependent models. For instance, combining Duffield and O'Connell [19] with Chang et al. [12] yields ruin probability estimates for some nonnecessarily Gaussian long range memory processes modeled after the fractional Brownian motion. More precise results were obtained by Hüsler and Piterbarg [29] for some Gaussian processes. Our results may be viewed as non-Gaussian and discrete analogues of those continuous ones, in the sense that we are interested in processes exhibiting long range dependence. Interestingly, for some values of their parameters, the processes considered in this paper, suitably rescaled and normalized, converge to some fractional Brownian motions.

A true extension of Cramér's theorem to FARIMA processes seems beyond what one can achieve at the present, and we will only consider a logarithmic form of it, namely, after taking the logarithm in (1.1),

$$\lim_{t\to\infty} t^{-1}\log \mathrm{P}\{M > t\} = -\theta.$$

The paper is organized as follows. The class of stochastic processes which we will consider and the main result are described in the next section. In Section 3 we describe the most likely scenario leading to a ruin, that is to a large maximum of the processes under consideration. Section 4 contains a broad outline of the proof. In Section 5, we prove some large deviations results which are of independent interest and lead to the proof – inspired by Collamore [13] – of the results of Section 2. The result of Section 3 is proved in Section 6. Some of the easier proofs are omitted or only outlined; their full version can be found in the preprint version of this paper available at arxiv.org/abs/0811.3460.

Notation. Throughout this paper, if (a_n) and (b_n) are two positive sequences, we say that 'a_n is lower bounded from above by an equivalent of b_n' and write $a_n \lesssim b_n$ if $a_n \le b_n\big(1+o(1)\big)$ as n tends to infinity. The symbol $\gtrsim$ is defined in an analogous way.

2. Main result

Barbe and McCormick [4] introduced (g, F)-processes as a natural extension of FARIMA processes. To define such a process, we start with a function g which is real analytic on $(-1, 1)$ and a distribution function F. The function g has a Taylor series expansion

$$g(x) = \sum_{i \geq 0} g_i x^i .$$

Considering a sequence $(X_i)_{i \geq 1}$ of independent and identically distributed random variables with common distribution function F, we define the (g, F)-process $(S_n)_{n \geq 0}$ by $S_0 = 0$ and

$$S_n = \sum_{0 \leq i < n} g_i X_{n-i} .$$

When $g(x) = (1-x)^{-1}$, the corresponding process is the random walk associated with the sequence $(X_n)_{n \geq 1}$. Some nonstationary ARMA processes are obtained when g is a rational function, and FARIMA processes are obtained when $g(x)$ is the product of some negative power of $1-x$ and a rational function in x.

For the process to drift toward minus infinity and mimic the behavior of the random walk involved in Cramér's theorem, it is natural to impose that the mean μ of F is negative and that

$$\lim_{n \to \infty} \sum_{0 \leq i < n} g_i = +\infty. \tag{2.1}$$

Indeed, in this case, the expectation of S_n drifts toward minus infinity. A consequence of (2.1) is that g has a singularity at 1. To obtain a satisfactory theory, we need to restrict the type of singularity by assuming that g is regularly varying at 1 of positive index γ, meaning, as explained for instance in [7], that there exists a positive γ such that for any positive λ,

$$\lim_{t \to \infty} \frac{g\big(1 - 1/(\lambda t)\big)}{g(1 - 1/t)} = \lambda^\gamma .$$

This assumption is satisfied by ARIMA processes.

Let Id be the identity function on the real line. We then consider a function U, defined up to asymptotic equivalence by the requirement

$$g(1 - 1/U) \sim \mathrm{Id}$$

at infinity. This function, which plays a key role in our result, is necessarily regularly varying at infinity of index $1/\gamma$. However, for notational simplicity, writing $\Gamma(\cdot)$ for the gamma function, it will be better to use the function

$$V = \Gamma(1+\gamma)^{1/\gamma} U,$$

which could alternatively be defined by the requirement $g(1 - 1/V) \sim \Gamma(1+\gamma)\mathrm{Id}$ at infinity.

In order to concentrate on the principles and the key arguments, we assume throughout this paper that the coefficients g_i are nonnegative. This restriction can be overcome with the introduction of the proper tail balance condition.

To have a compact notation, we introduce the kernel

$$k_\gamma(u) = \begin{cases} \gamma(1-u)^{\gamma-1} & \text{if } 0 \leq u < 1, \\ 0 & \text{if } u \geq 1, \end{cases}$$

defined on the nonnegative half-line.

Further notation related to large deviations theory is needed in order to state our main result. As the proof shows, the appearance of some large deviations formalism is not coincidental. Cramér's theorem assumes that the moment generating function

$$\varphi(\lambda) = \mathrm{E}e^{\lambda X_1}$$

is finite in a neighborhood of the origin. A classical consequence of Hölder's inequality is that $\log\varphi$ is convex. This implies that the function

$$\lambda \mapsto \int_0^1 \log\varphi\big(\lambda k_\gamma(u)\big)\mathrm{d}u \tag{2.2}$$

is convex as well on its domain. This function will be of importance in our results. It is not clear a priori that this function is nontrivial in the sense that if γ is less than 1 it could be infinite for all nonvanishing λ. This suggests that we should consider two cases, according to the finiteness of the integral involved in (2.2).

The convex conjugate (see e.g. [43]) of the function involved in (2.2), at a nonnegative argument a, is

$$J(a) = \sup_{\lambda>0}\Big(a\lambda - \int_0^1 \log\varphi\big(\lambda k_\gamma(u)\big)\mathrm{d}u\Big).$$

With a moment generating function φ one also associates the corresponding mean function m, which is the derivative $(\log\varphi)'$ — see [5,10] or [34].

The following convention will be convenient. We say that a (g, F)-process satisfies the standard assumption if it satisfies the following:

Standard assumption. The function g is regularly varying of positive index at 1 and its coefficients $(g_i)_{i\geq 0}$ are nonnegative. Moreover g_0 does not vanish. If the sequence $(g_n)_{n\geq 0}$ converges to 0, it is asymptotically equivalent to a monotone sequence. The distribution function F has a moment generating function finite on the nonnegative half-line. The image of the mean function contains the half-line $[0, \infty)$.

With respect to the monotonicity requirement for the sequence $(g_n)_{n\geq 0}$ involved in the standard assumption, it will follow from Proposition 1.5.3 in [7] and Lemma 5.1.1 that regular variation of g implies that $(g_n)_{n\geq 1}$ is asymptotically equivalent to a monotone sequence whenever the index of regular variation of g is different from 1.

Let $(S_n)_{n\geq 0}$ be a (g, F)-process. If the first k coefficients $g_0, g_1, \ldots, g_{k-1}$ vanish and g_k does not, then $(S_{n+k})_{n\geq 1}$ is a $(g/\mathrm{Id}^k, F)$-process, and the first Taylor coefficient of g/Id^k does not vanish. Thus, in the standard assumption, the condition that g_0 does not vanish bears no restriction.

Note that in the standard assumption, the condition on the moment generating function is stronger than in Cramér's theorem. The assumption on the mean function is a rather standard one in large deviations theory. Hölder's inequality implies that $\log\varphi$ is convex and the mean function is nondecreasing. Our assumption ensures that the equation $m(\lambda) = x$ has a solution for every positive x.

We also say that a (g, F)-process satisfying the standard assumption has a negative mean if its expectation is negative at all times. Since the innovations are independent and identically distributed, considering the expectation of the process at time 1, this is equivalent to requiring that the innovations have negative mean.

Our first result treats the case where the integral (2.2) is finite. It calls for many remarks, stated after the theorem, which will clarify both the hypotheses and the conclusion.

Theorem 2.1. *Consider a negative mean (g, F)-process which satisfies the standard assumption. Assume that either one of the following conditions holds:*

(i) $\limsup_{n\to\infty} \max_{0\le i<n} g_i/g_n$ *is finite;*
(ii) $\lim_{n\to\infty} \max_{0\le i\le n} g_i/g_n = +\infty$ *and* $-\log \overline{F}$ *is regularly varying with index* α *such that* $\alpha\gamma > 1$*; moreover,* m' *is regularly varying.*

Then, the function J is defined and finite on the nonnegative half-line and the maximum M of the (g, F)-process satisfies

$$\lim_{t\to\infty} V(t)^{-1} \log \mathrm{P}\{M > t\} = -\inf_{x>0} xJ(x^{-\gamma}).$$

We now make some remarks on the conclusion of the theorem, which will be followed by remarks on its assumptions.

Writing θ for the negative of the limit involved in its statement, this theorem asserts that

$$\mathrm{P}\{M > t\} \sim \mathrm{e}^{-\theta V(t)(1+o(1))}$$

as t tends to infinity. This leads to the following observation which may constitute a caveat of pedagogical value. Fix the distribution function F and consider the analytic function g as a parameter. As we increase its singularity at 1, the process drifts toward minus infinity at a faster rate, for its mean at time n is $\mu \sum_{0\le i<n} g_i$. One might guess that this makes it harder for the process to reach a high threshold. However, our theorem asserts that the logarithmic order of this probability is $-V(t)$, which becomes larger with g. So, making the mean diverge to minus infinity faster makes it more likely for the process to reach a high level! This phenomenon will be explained in the next section.

In the same spirit, it follows that multiplying the X_i by a scale factor σ divides θ by $\sigma^{1/\gamma}$. Thus, increasing the drift toward minus infinity through a scaling increases the likelihood for M to take very large values.

On a different note, we see that as in Cramér's theorem, the tail of the distribution function of the increments is involved in the conclusion of Theorem 2.1 only in the constant θ and not in the logarithmic decay V.

It is also of interest to note that if γ is greater than 1, then $V \ll \mathrm{Id}$ at infinity. In this case, Theorem 2.1 shows that the distribution of the maximum of the process is subexponential, even though the innovations are superexponential. Such a possibility was observed in a different context by Kesten [33].

Regarding the assumptions of Theorem 2.1, note that in case (i) we must have γ at least 1. In case (ii), the condition that $\max_{0\le i\le n} g_i/g_n$ diverges to infinity is equivalent to the convergence of $(g_n)_{n\ge 0}$ to 0, which forces γ to be at most 1.

Let β be the conjugate exponent of α, that is such that $\alpha^{-1} + \beta^{-1} = 1$. It follows from Kasahara's theorem [7, Theorem 4.12.7] that $-\log \overline{F}$ is regularly varying of index α if and only if $\log \varphi$ is regularly varying of index β. Since $\log \varphi$ is convex, its derivative is monotone, and the monotone density theorem combined with Kasahara's theorem implies that $-\log \overline{F}$ is regularly varying of index α if and only if m is regularly varying of index $\beta - 1$. The assumption on m' in Theorem 2.1 is stronger. This assumption is not completely satisfactory since its meaning in terms of the distribution function is not clear.

Under the assumptions of Theorem 2.1, we must have $\beta - 2 > -1$. Hence, using Karamata's theorem in addition to the previous paragraph, we see that the assumption of Theorem 2.1 on $-\log \overline{F}$ and m' is equivalent to the single assumption that m' is regularly varying of index $\beta - 2$ with $\beta(1-\gamma) < 1$.

Our second result considers the case where the integral involved in (2.2) is infinite, and hence the function J in Theorem 2.1 is not defined. This essentially occurs when γ is less than 1 and $\alpha\gamma$ is at most 1. If γ is less than $1/2$ then the centered process $S_n - \mathrm{E}S_n$ converges in L^2. For γ less than $1/2$, let Z_n be the linear process

$$Z_n = \sum_{i\geq 0} g_i(X_{n-i} - \mu).$$

In this case, we see that the ruin problem for S_n is rather similar to that of determining the probability that the process $(Z_n)_{n\geq 1}$ crosses the moving boundary $t - ES_n$. This problem is of somewhat different nature than what is the focus of this paper, for the centered process is well approximated by a stationary one. Therefore, we will limit ourselves to the case where γ is greater than $1/2$.

We write $|g|_\beta$ for the ℓ_β-norm of the sequence of its coefficients, that is for $\left(\sum_{i\geq 0} g_i^\beta\right)^{1/\beta}$.

Theorem 2.2. *Consider a (g, F)-process which satisfies the standard assumption and with $1/2 < \gamma < 1$. Assume furthermore that $-\log \overline{F}$ is regularly varying of index α greater than 1 and that $\alpha\gamma < 1$. Let β be the conjugate exponent of α. Then, the maximum M of the process satisfies*

$$\log \mathrm{P}\{M > t\} \sim |g|_\beta^{-\alpha} \log \overline{F}(t).$$

as t tends to infinity.

Comparing Theorems 2.1 and 2.2, we see that in Theorem 2.2, the condition $\alpha\gamma < 1$ forces the rate of growth of $-\log F(t)$, regularly varying of index α, to be much slower than that of $U(t)$, regularly varying of index $1/\gamma$.

3. How to go bankrupt?

The purpose of this section is to determine the most likely paths which lead to the maximum of our (g, F)-processes to reach a high threshold. Beyond its relevance to choosing interesting alternatives in change point problems, in the context of ruin probability, this amounts to finding the most likely way of becoming bankrupted. In a different context, high risk scenarios have been the subject of Balkema and Embrechts [2] monograph where further discussion of the topic may be found. More closely related to the topic of this paper is the work of Chang et al. [12] in the continuous setting, who consider the analogous problem for fractional integrals of continuous time processes. In fact we are seeking more information. Not only are we interested in the most likely paths, but also we would like to understand how they arise, and, therefore, have a description of the innovations as well. For the heavy tail case, it is shown in [4] that a large value of the maximum of the process is most likely caused by a large value of an innovation. In contrast, in a setting slightly different than that of the current paper, but nonetheless related, for the usual random walk, [15, Theorem 1] shows that a large deviation is likely caused by a cooperative behavior of the increments which pushes the sum upward. More precisely, Csiszár's result implies that the conditional distribution of the first increment, given that the sum S_n exceeds an unlikely threshold nu, converges to the distribution $\mathrm{d}F_u(x) = \mathrm{e}^{m^{\leftarrow}(u)x}\mathrm{d}F(x)/\varphi \circ m^{\leftarrow}(u)$.

The distribution $\mathrm{d}F_u$ has mean u. For the usual random walk, since the increments are exchangeable given their sum, Csiszár's result asserts that, loosely, a randomly chosen increment, or a typical increment, has a conditional distribution about $\mathrm{d}F_u$. Thus, asymptotically, the bulk of the increments behave like a random variable of mean u under the conditional distribution that the random walk at time n exceeds nu. We refer the reader to [17] for a refined result in the framework of exponential families.

In general, for (g, F)-processes, the innovations are not exchangeable given the value of the process at time n, and, paralleling what has been done for the random walk, it is of interest to identify the cooperative behavior of the increments, if any, which makes the process reach a high level.

Besides a theoretical understanding, this type of conditional limiting result has some bearing on techniques of simulation of rare events by importance sampling [27]. Indeed, when specialized to the regular random walk, the work of Sadowsky [45] gives a rationale for using the limiting conditional distribution of the increments to simulate unlikely paths of random walks using importance sampling; see also [18]. Our result is a key building block for extending this technique to some FARIMA processes, and, more generally, to (g, F)-processes.

To investigate these questions, we consider first the rescaled trajectory

$$\mathcal{S}_t(\lambda) = S_{\lfloor \lambda V(t) \rfloor}/t, \quad \lambda \geq 0.$$

Next, to study the behavior of the innovation, we consider the sequential measure

$$\mathcal{M}_t = \frac{1}{V(t)} \sum_{i \geq 1} \delta_{(i/V(t), X_i)}$$

which puts mass $1/V(t)$ at each pair $\big(i/V(t), X_i\big)$. In contrast with a standard empirical measure which would put equal mass on each innovation up to some fixed time, the sequential measure keeps track of the sequential ordering of the innovation through the first component $i/V(t)$.

Of further interest is also the normalized first time that the process reaches the level t,

$$\mathcal{N}_t = \frac{1}{V(t)} \min\{n : S_n > t\}.$$

In order to speak of convergence of the stochastic process $\mathcal{S}_t$, we view it in the Skorohod space $\mathrm{D}[0, \infty)$ equipped with the Skorohod topology [6,35].

In what follows, we call $[0, \infty) \times \mathbb{R}$ the right half-space. A subset of the right half-space of the form $[a, b] \times \mathbb{R}$ is called a vertical strip.

The measure $\mathcal{M}_t$ belongs to the space $\mathcal{M}([0, \infty) \times \mathbb{R})$ of σ-finite measures on the right half-space. We consider this space equipped with a topology between those of vague and weak$*$ convergences defined as follows. Let $\mathrm{C}_{\mathrm{K,b}}([0, \infty) \times \mathbb{R})$ be the space of all real-valued continuous and bounded functions on the right half-space, supported on a vertical strip. A basis for the topology on $\mathcal{M}([0, \infty) \times \mathbb{R})$ is defined by the sets

$$\left\{ \mu \in \mathcal{M}([0, \infty) \times \mathbb{R}) : \forall i = 1, \ldots, k, \left| \int f_i \,\mathrm{d}(\mu - \nu) \right| < \epsilon \right\},$$

indexed by

$$\nu \in \mathcal{M}([0, \infty) \times \mathbb{R}), \qquad f_i \in \mathrm{C}_{\mathrm{K,b}}([0, \infty) \times \mathbb{R}), \quad \epsilon > 0.$$

In this paper, except when specified otherwise, all convergences of measures on the right half-space are for this topology.

Our next result gives the limit in probability of the various quantities introduced, conditionally on having the process reaching the level t, and under the assumptions of Theorem 2.1. We assume that

$$\tau = \arg\min_{x>0} x J(x^{-\gamma}) \text{ is unique.} \tag{3.1}$$

Furthermore, we define the constant A to be the solution of

$$\tau^{-\gamma} = \int_0^1 k_\gamma(u) m\bigl(A k_\gamma(u)\bigr) \mathrm{d}u. \tag{3.2}$$

Let L be the Lebesgue measure. We define the measure $\mathcal{M}$ by its density with respect to the product measure $L \otimes F$,

$$\frac{\mathrm{d}\mathcal{M}}{\mathrm{d}(L \otimes F)}(v, x) = \frac{\exp\bigl(A k_\gamma(v/\tau)x\bigr)}{\varphi\bigl(A k_\gamma(v/\tau)\bigr)}. \tag{3.3}$$

In particular, since k_γ vanishes on $[\,1, \infty)$, the measure $\mathcal{M}$ coincides with $L \otimes F$ on $[\,\tau, \infty) \times \mathbb{R}$. We also define the function

$$\mathcal{S}(\lambda) = \int_0^\lambda \gamma(\lambda - v)^{\gamma-1} m\bigl(A k_\gamma(v/\tau)\bigr) \mathrm{d}v. \tag{3.4}$$

Writing

$$\mathcal{S}(\lambda) = \lambda^\gamma \int_0^1 k_\gamma(v) m\bigl(A k_\gamma(v\lambda/\tau)\bigr) \mathrm{d}v$$

and using (3.2), we see that $\mathcal{S}(\tau) = 1$.

The following result describes the most likely ruin scenario.

Theorem 3.1. *Under the assumptions of Theorem 2.1, the following hold in probability under the conditional probability given $M > t$ as t tends to infinity:*

(i) *$\mathcal{N}_t$ converges to τ;*
(ii) *$\mathcal{M}_t$ converges to $\mathcal{M}$;*
(iii) *moreover, if the moment generating function of $|X_1|$ is finite in a neighborhood of the origin, then $\mathcal{S}_t$ converges locally uniformly to $\mathcal{S}$.*

Loosely speaking, the meaning of assertion (ii) is that the conditional distribution of $X_{\lfloor vV(t)\rfloor}$ given $M > t$ converges to the measure

$$\begin{cases} \dfrac{\mathrm{e}^{A\gamma(1-v/\tau)^{\gamma-1}x}}{\varphi\bigl(A\gamma(1-v/\tau)^{\gamma-1}\bigr)} \mathrm{d}F(x) & \text{if } v \le \tau \\ \mathrm{d}F(x) & \text{if } v > \tau, \end{cases}$$

with mean $m\bigl(A\gamma(1-v/\tau)^{\gamma-1}\bigr)$ if $v < \tau$, and μ if $v \ge \tau$. Thus it asserts that a large value of M is likely caused by a cooperative behavior of the random variables up to a time $\tau V(t)\bigl(1 + o(1)\bigr)$, while the remainder of the innovations keep their original distribution. This somewhat confirms that the Cramér ruin model might be unrealistic in some situations. Indeed, Theorem 3.1 shows that for the process to reach the large level t, both the increments and the process, from the very beginning, have to follow a very unlikely path. One would think that seeing such a strange path unfolding, a careful insurer would quickly re-examine the model and raise the premium accordingly.

Theorem 3.1 also explains why Theorem 2.1 implies that for those (g, F)-processes, adding more drift toward minus infinity may increase the likelihood of a large maximum. Indeed Theorem 3.1 indicates that a large value of the maximum is likely to be caused by many innovation being large; but if the weights $(g_n)_{n\geq 0}$ are made larger, then comparatively smaller innovation suffices for the maximum of the process to reach a large value, because the coefficients $(g_n)_{n\geq 0}$ amplify the innovations.

We now consider an example of processes of interest and for which the limit involved in Theorems 2.1 or 2.2 can be made explicit. In general this limit must be evaluated by numerical methods.

We consider a Gaussian FARIMA process. More specifically, we consider F to be the Gaussian distribution function with mean μ and variance σ^2, and we introduce two polynomials Θ and Φ, neither of which vanishes at 1. We consider the function $g(x) = (1-x)^{-\gamma}\,\Theta(x)/\Phi(x)$, so the corresponding (g, F)-process is a FARIMA(Φ, γ, Θ) process whose innovations have a common distribution function F. For this specific function g we may take $U(t) = \big(t\,\Phi(1)/\Theta(1)\big)^{1/\gamma}$.

The moment generating function of the innovations is

$$\varphi(\lambda) = \mathrm{e}^{\lambda\mu+\sigma^2\lambda^2/2}.$$

The function involved in (2.2) is then

$$\gamma \int_0^1 \mu\lambda\gamma u^{\gamma-1} + \frac{\sigma^2}{2}(\lambda\gamma u^{\gamma-1})^2 \mathrm{d}u = \lambda\mu + \frac{\sigma^2}{2}\lambda^2 \frac{\gamma^2}{2\gamma-1}.$$

This implies that

$$J(a) = \sup_\lambda\Big(a\lambda - \lambda\mu - \frac{\sigma^2}{2}\lambda^2\frac{\gamma^2}{2\gamma-1}\Big) = \frac{(a-\mu)^2(2\gamma-1)}{2\sigma^2\gamma^2}.$$

Using standard calculus one more time, we obtain

$$\inf_{x>0} xJ(x^{-\gamma}) = \frac{2(2\gamma-1)^{1/\gamma-1}}{\sigma^2}(-\mu)^{2-1/\gamma}.$$

Therefore, the conclusion of Theorem 2.1 is that

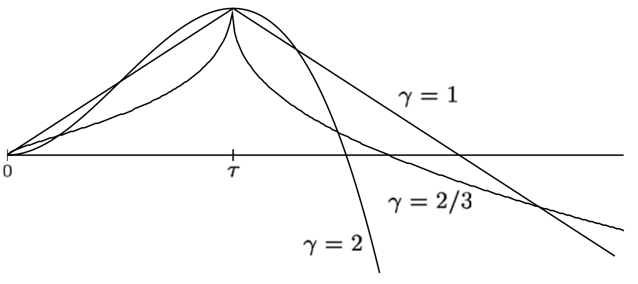
$$\log \mathrm{P}\{M > t\} \sim -t^{1/\gamma}\Big(\frac{\Phi(1)}{\Theta(1)}\Big)^{1/\gamma}\Gamma(1+\gamma)^{1/\gamma}2(2\gamma-1)^{(1/\gamma)-1}\Big(\frac{\mu}{\sigma}\Big)^2(-\mu)^{-1/\gamma}$$

as t tends to infinity.

To calculate the limiting process $\mathcal{S}$, for simplicity we restrict ourselves to the case where the mean μ is -1 and the standard deviation σ is 1. Then, $m(\lambda) = \lambda - 1$, and

$$\mathcal{S}(\lambda) = A\int_0^{\lambda\wedge\tau} \gamma(\lambda-v)^{\gamma-1}\gamma\Big(1-\frac{v}{\tau}\Big)^{\gamma-1}\mathrm{d}v - \lambda^\gamma.$$

The following graphic shows the shape of the limiting function when γ is $2/3$, 1 and 2.



We conclude this section by some remarks concerning Theorem 3.1 and its proof. A close look at the proofs of Theorems 2.1 and 3.1 reveals that the same technique allows one to derive a large deviations principle for the process $\mathcal{S}_t$ and the measure $\mathcal{M}_t$ under the conditional distribution of M exceeding t, as t tends to infinity, in the spirit of Collamore [14].

One can also see that under the assumptions of Theorem 2.1, the large deviations principle for FARIMA processes proved in [3] remains true when the order of differentiation γ is between $1/2$ and 1 and that the logarithm of the tail of the distribution function of the innovation is regularly varying with index greater than $1/\gamma$. This has the following interesting consequence as regards the standard partial sum process, $\Pi_n(\lambda) = n^{-1} \sum_{1 \le i \le n\lambda} X_i$, $0 \le \lambda \le 1$. Consider the Cramér transform of the increment, $I(x) = \sup_\lambda \bigl(\lambda x - \log \varphi(\lambda)\bigr)$. Mogulskii's theorem [36] (see also [16], Section 5.1) asserts that the partial sum process obeys a large deviations principle, in the supremum norm topology. We can write the partial sum process at time t as $\int_0^\lambda \mathrm{d}\Pi_n(v) = \int \mathbb{1}_{[0,\lambda)}(v) \mathrm{d}\Pi_n(v)$. One could then wonder whether some fractional integral of Π_n still obeys a large deviations principle. While an integration by parts shows that for γ greater than 1, the process $\lambda \in [0, 1] \mapsto \int_0^\lambda (\lambda - v)^{\gamma-1} \mathrm{d}\Pi_n(v)$ obeys a large deviation, the proof of Theorem 3.1 shows that such a large deviations principle still holds if $1/2 < \gamma < 1$, provided that $\log \overline{F}$ is regularly varying of index greater than $1/\gamma$. The Gaussian case, $\alpha = 2$, appears to be a boundary one corresponding to $\gamma = 1/2$; and this matches the fact that the Brownian motion belongs to any set of functions with Hölder exponent less than $1/2$.

4. Generalities

The study of first-passage times using large deviations is now a classical topic which has been presented in book form by Freidlin and Wentzell [23]. The purpose of this section is to give another short variation on this theme, with a formalism more suitable for the problems considered in this paper. What follows is inspired by the work of Collamore [14] as well as Duffield and Whitt [20]. However, in contrast to those authors, we are interested in processes which are not Markovian, not mixing and not monotone.

Some notation will purposely be identical to those used in the previous sections, the reason being that they have the same meaning when specialized to the context of the previous sections; this will be clear during the proofs of Theorems 2.1, 2.2 and 3.1.

In what follows, sequences are viewed as functions defined on the nonnegative half-line and evaluated at the integers. Therefore, if we write $(a_n)_{n \ge 1}$ for a sequence, we will also speak of the function a, meaning that $a_n = a(n)$ for every positive integer n. If we are given the sequence, it is understood that the function a is obtained by a linear interpolation say; other 'reasonable' interpolation procedures would do just as well.

In this section we consider a stochastic process $(S_n^0)_{n\geq 1}$ and a sequence $(s_n)_{n\geq 1}$ which diverges to infinity. We are interested in evaluating the probability that the process $(S_n^0)_{n\geq 1}$ crosses the moving boundary $(t + s_n)_{n\geq 1}$ for large values t. In other words, assuming that $M = \max_{n\geq 1} S_n^0 - s_n$ is well defined, we are interested in finding an estimate of

$$\mathrm{P}\{\exists n \geq 1 \,:\, S_n^0 > t + s_n\} = \mathrm{P}\{M > t\}$$

as t tends to infinity. Assuming that the function

$$s \text{ is regularly varying of positive index } \gamma, \tag{4.1}$$

there exists a function V, defined, up to asymptotic equivalence, by the relation $s \circ V \sim \mathrm{Id}$ at infinity. Also of interest is the normalized first-passage time at which the process crosses the moving boundary,

$$\mathcal{N}_t = \frac{1}{V(t)} \min\{n \geq 1 \,:\, S_n^0 > t + s_n\}.$$

Suppose that $(S_n^0)_{n\geq 1}$ obeys a large deviations principle in the sense that there exist two functions r and I such that for any positive x,

$$\log \mathrm{P}\{S_n^0 > s_n x\} \sim -r_n I(x) \tag{4.2}$$

as n tends to infinity. Since the left hand side of (4.2) is monotone in x, so is the right hand side, and, necessarily, I is monotone as well as continuous almost everywhere. If we assume more, namely that

$$I \text{ is continuous}, \tag{4.3}$$

then the asymptotic equivalence in (4.2) holds locally uniformly in x over the nonnegative half-line, because a pointwise convergent sequence of nondecreasing functions whose limit is continuous converges locally uniformly (see [44], chapter 7, exercise 13).

For our problem, we will be able to assume that

$$r \text{ is regularly varying of positive index } \rho. \tag{4.4}$$

In this case, r is asymptotically equivalent to a nondecreasing function, and we will consider, without any loss of generality, that r is nondecreasing. We define θ as

$$\theta = \inf_{x>0} x^\rho I(x^{-\gamma} + 1). \tag{4.5}$$

We will also assume that the process is unlikely to reach the moving boundary $t + s_n$ before a time of order $V(t)$, in the sense that

$$\lim_{\epsilon\to 0} \limsup_{t\to\infty} \frac{1}{r \circ V(t)} \log \mathrm{P}\{\exists n \,:\, 1 \leq n \leq \epsilon V(t)\,;\; S_n^0 > t + s_n\} \leq -\theta. \tag{4.6}$$

Equipped with these perhaps drastically looking – but to be proved useful – conditions, we have the following.

Proposition 4.1. *If (4.1)–(4.6) hold, then*

$$\log \mathrm{P}\{M > t\} \sim -\theta r \circ V(t)$$

as t tends to infinity. Moreover, if

$$\tau = \arg\min_{x>0} x^{\rho} I(x^{-\gamma} + 1) \text{ exists and is unique,} \tag{4.7}$$

then $\mathcal{N}_t$ converges to τ in probability given $M > t$, as t tends to infinity.

Remark. If we replace assumption (4.2) by

$$\log \mathrm{P}\{S_n^0 > x s_n\} \lesssim -r_n I(x) \tag{4.8}$$

as n tends to infinity, the proof of Proposition 4.1 shows that

$$\log \mathrm{P}\{M > t\} \lesssim -\theta r \circ V(t)$$

as t tends to infinity. This remark will be useful for proving Theorem 2.2.

In order to prove Proposition 4.1, we recall that if r is a nondecreasing regularly varying function of positive index, then

$$\sum_{n \geq k} \mathrm{e}^{-r_n} \lesssim \mathrm{e}^{-r_k(1+o(1))} \tag{4.9}$$

as k tends to infinity (see Theorem 4.12.10 in [7])

Proof of Proposition 4.1. The proof of the first assertion consists in establishing the proper upper and lower bounds.

Upper bound. Let ϵ be a positive real number less than 1. For t large enough and uniformly in n between $\epsilon V(t)$ and $V(t)/\epsilon$,

$$r_n = r\Big(V(t)\frac{n}{V(t)}\Big) \sim r \circ V(t)\Big(\frac{n}{V(t)}\Big)^{\rho}$$

and

$$\frac{t}{s_n} = \frac{t}{s\Big(V(t)\frac{n}{V(t)}\Big)} \sim \Big(\frac{n}{V(t)}\Big)^{-\gamma}$$

as t tends to infinity. In particular,

$$\begin{aligned} r_n I\Big(\frac{t}{s_n} + 1\Big) &\sim r \circ V(t)\Big(\frac{n}{V(t)}\Big)^{\rho} I\Big(\Big(\frac{n}{V(t)}\Big)^{-\gamma} + 1\Big) \\ &\gtrsim r \circ V(t)\theta. \end{aligned}$$

Combining this lower bound with the large deviations assumption (4.2) yields, in the range of n between $\epsilon V(t)$ and $V(t)/\epsilon$ and for t large enough,

$$\mathrm{P}\{S_n^0 > t + s_n\} \leq \exp\big(-r \circ V(t)\theta(1-\epsilon)\big). \tag{4.10}$$

It follows that for t large enough,

$$\mathrm{P}\{\exists n : \epsilon V(t) \leq n \leq V(t)/\epsilon\,;\ S_n^0 > t + s_n\} \leq \frac{V(t)}{\epsilon}\exp\big(-r \circ V(t)\theta(1-\epsilon)\big).$$

Still using the large deviations assumption (4.2), for n at least $V(t)/\epsilon$ and t large enough, we have

$$\mathrm{P}\{S_n^0 > t + s_n\} \le \mathrm{P}\{S_n^0 > s_n\} \le \mathrm{e}^{-r_n I(1)/2}.$$

Thus, for t large enough, using (4.9),

$$\begin{aligned}\mathrm{P}\{\exists n\,:\, n \ge V(t)/\epsilon\,;\; S_n^0 > t + s_n\} &\le \sum_{n \ge V(t)/\epsilon} \mathrm{e}^{-r_n I(1)/2} \\ &\lesssim \exp\Bigl(-r \circ V(t) \frac{I(1)}{2\epsilon^\rho}\bigl(1 + o(1)\bigr)\Bigr).\end{aligned}$$

Taking ϵ small enough, it follows that

$$\log \mathrm{P}\{\exists n\,:\, n \ge V(t)/\epsilon\,;\; S_n^0 > t + s_n\} \lesssim -2r \circ V(t)\theta$$

as t tends to infinity. Using assumption (4.6), we conclude that

$$\log \mathrm{P}\{M > t\} \lesssim -r \circ V(t)\theta$$

asymptotically.

Lower bound. Let ϵ be a positive real number and let x be a positive real number such that $x^\rho I(x^{-\gamma} + 1) \le \theta + \epsilon$. Let n be the integer part of $xV(t)$. Then

$$\mathrm{P}\{M > t\} \ge \mathrm{P}\{S_n^0 > t + s_n\}.$$

From the large deviations hypothesis (4.2), we then deduce

$$\log \mathrm{P}\{M > t\} \gtrsim -r \circ V(t)(\theta + \epsilon). \tag{4.11}$$

Since ϵ is arbitrary, the first assertion of Proposition 4.1 follows.

To prove the second assertion, note that estimate (4.2), with (4.7), implies that

$$\mathrm{P}\{|\mathcal{N}_t - \tau| > \eta \mid M > t\} \le \frac{\mathrm{P}\{\exists n\,:\, |n - \tau V(t)| > \eta V(t)\,;\; S_n^0 > t + s_n\}}{\mathrm{P}\{M > t\}}$$

tends to 0 as t tends to infinity. The second assertion follows. ■

5. Proof of results of Section 2

We will assume that the mean of the innovations, μ, is -1. Other values of μ will be dealt with by a scaling argument.

To obtain pleasing expressions, for every positive real number r we write $g_{[0,r)}$ for $\sum_{0 \le i < r} g_i$ and we also write s_n for the negative of the mean of S_n, that is $s_n = g_{[0,n)}$ — recall our assumption that μ is -1 until further notice. With the notation of the previous section, S_n^0 is the centered process $S_n - ES_n = S_n + s_n$. Moreover, V is defined by $s_{\lfloor V(t) \rfloor} \sim t$ as t tends to infinity.

5.1. Preliminary

The following lemma, relating g_n and $g_{[0,n)}$ to $g(1 - 1/n)$, will be very useful. It essentially restates Karamata's Tauberian theorem for power series ([7], Corollary 1.7.3) and is proved in Lemma 5.1.1 in [4]. We state it here for the sake of making the proof easier to read, for it is fundamental in our problem and we will refer to it often.

Lemma 5.1.1. *The following asymptotic equivalences hold as n tends to infinity, uniformly in x in any compact subset of the positive half-line:*

(i) $g_{\lfloor nx \rfloor} \sim \frac{x^{\gamma-1}}{\Gamma(\gamma)} \frac{g(1-1/n)}{n}$,

(ii) $g_{[0,nx)} \sim \frac{x^{\gamma}}{\Gamma(1+\gamma)} g(1-1/n)$.

In particular, this implies that $g_n \sim \gamma g_{[0,n)}/n$ as n tends to infinity, so locally uniformly in any positive c,

$$g_{\lfloor cV(t) \rfloor} \sim \gamma c^{\gamma-1} \frac{t}{V(t)} \tag{5.1.1}$$

as t tends to infinity.

We introduce the notation $g_{i/n}$ for $\gamma g_i/g_n$ in which the subscript i/n has clearly nothing to do with the division of i by n but serves as a mnemonic for the division of g_i by g_n. In particular, $g_{n-i/n}$ is $\gamma g_{n-i}/g_n$. Lemma 5.1.1 asserts that $g_{n-i/n} \sim k_\gamma(i/n)$ as n tends to infinity and i/n stays bounded away from 1.

The following easy lemma, whose proof is omitted, is recorded for further reference.

Lemma 5.1.2. *Let*

$$c_1 = \liminf_{n\to\infty} g_n \quad \textit{and} \quad c_2 = \limsup_{n\to\infty} \max_{0\le i\le n} g_{i/n}.$$

(i) *If c_1 is positive, then $U \lesssim \mathrm{Id}/\big(c_1\Gamma(\gamma)\big)$ at infinity.*
(ii) *Assume that c_2 is finite. If the sequence $(g_n)_{n\ge0}$ is bounded, then $U \lesssim c_2\mathrm{Id}/\gamma \max_{i\ge0} g_i$; otherwise $U = o(\mathrm{Id})$ at infinity.*

Our next lemma is perhaps the heart of the proof, which ultimately relies on approximation of Riemann sums by a Riemann integral, a modicum of regular variation, and the exponential form of Markov's inequality.

We define the sequence of probabilities measures

$$\Gamma_n = n^{-1} \sum_{1\le i\le n} \delta_{(i/n, g_{n-i/n})}, \qquad n \ge 1.$$

Lemma 5.1.3. *The sequence of probability measures $(\Gamma_n)_{n\ge1}$ converges weakly $*$ to the measure $\int_0^1 \delta_{(u,k_\gamma(u))}\mathrm{d}u$.*

Proof. Let f be a nonnegative continuous and bounded function on $[0,1] \times \mathbb{R}$. We write $|f|_{[0,1]\times\mathbb{R}}$ for its supremum on the strip $[0,1]\times\mathbb{R}$. Let ϵ be a positive real number less than 1. Note that

$$n^{-1} \sum_{(1-\epsilon)n<i\le n} f(i/n, g_{n-i/n}) \le \epsilon |f|_{[0,1]\times\mathbb{R}}.$$

The result then follows from Lemma 5.1.1, which implies that $g_{n-i/n} \sim k_\gamma(i/n)$ whenever i/n stays bounded away from 1, and from the approximation of a Riemann sum by the corresponding integral. ■

In order to simplify the notation during the proof and make a later scaling argument easier to follow, we write φ_0 for the moment generating function of the centered random variable

$\big(X_1/(-\mu)\big)+1$. Furthermore, we write

$$J_0(a)=\sup_{\lambda>0}\Big(a\lambda-\int_0^1\log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u\Big). \tag{5.1.2}$$

The equality $\varphi_0(\lambda)=\mathrm{e}^\lambda\varphi(\lambda/-\mu)$, valid for all λ positive, yields

$$J_0(a+1)=J(-\mu a). \tag{5.1.3}$$

5.2. *Proof of* Theorem 2.1

Recall that except if specified otherwise, we consider the mean μ to be -1. Also, throughout this subsection, we assume that the hypotheses of Theorem 2.1 hold, even if this is not specified.

The proof is based on a large deviations estimate which is the analogue for the (g,F)-process of the classical estimate of Chernoff for the sample mean. The proof requires several lemmas.

Our first lemma will be useful in taking limits in various sums involving the moment generating function.

Lemma 5.2.1. *Let h be a continuous function on the nonnegative half-line. Assume g satisfies the assumption of Theorem 2.1. If $\lim_{n\to\infty}\max_{0\le i\le n}g_{i/n}=\infty$, assume further that h is regularly varying of index β less than $1/(1-\gamma)$. Then, locally uniformly in λ in $(0,\infty)$,*

$$\lim_{n\to\infty}n^{-1}\sum_{1\le i\le n}h(\lambda g_{n-i/n})=\int_0^1 h\big(\lambda k_\gamma(u)\big)\mathrm{d}u,$$

and this limit is finite.

When $\lim_{n\to\infty}\max_{0\le i\le n}g_{i/n}=\infty$ and γ is 1, the condition on h should simply be read as h is regularly varying of some positive index.

Proof. Note first that in both cases, $\lim_{\epsilon\to0}\int_0^\epsilon h(\lambda u^{\gamma-1})\mathrm{d}u=0$ and the integral involved in the limit in the lemma is indeed finite.

Once the limit is established for a fixed λ, it will be clear that using the uniform convergence theorem for regularly varying functions, the limit is locally uniform in λ. Thus, up to changing the function h, it suffices to prove the result only when λ is 1.

For any positive real number c we define the function $h_c=h(\cdot\wedge c)$. These functions are continuous and bounded.

If $\limsup_{n\to\infty}\max_{0\le i\le n}g_{i/n}$ is finite, we take c to be twice this limit, so that for n large enough, $\max_{0\le i\le n}g_{i/n}$ is at most c. Then the result follows from Lemma 5.1.3 and the local uniform continuity in λ of the functions $h_c(\lambda\,\cdot)$.

If $\lim_{n\to\infty}\max_{0\le i\le n}g_{i/n}$ is infinite, let ϵ be a positive real number less than 1. Since $\lim_{n\to\infty}\max_{\epsilon n\le i\le n}g_{i/n}=\epsilon^{\gamma-1}$ is finite, it suffices to prove that

$$\lim_{\epsilon\to0}\limsup_{n\to\infty}n^{-1}\sum_{0\le i\le\epsilon n}h(g_{i/n})=0. \tag{5.2.1}$$

This follows from Lemma 5.1.1 and standard regular variation theoretic arguments, using Potter's bound ([7], Theorem 1.5.3). ■

Equipped with Lemma 5.2.1, we can prove the following large deviation principle. Recall that we assume for the time being that the distribution function F has mean -1. We write F_0 for the cumulative distribution function $F(\cdot-1)$. As the subscript indicates, its mean is 0.

Proposition 5.2.2. *Let $(S_n^0)_{n\geq 0}$ be a centered (g, F_0)-process. Under the assumptions of Theorem 2.1, for any nonnegative x,*

$$\lim_{n\to\infty} n^{-1}\log \mathrm{P}\{S_n^0 > x g_{[0,n)}\} = -J_0(x).$$

Moreover, the limit is locally uniform in x on the set where J_0 is finite.

Proof. The proof is modeled after the standard one for the mean. We will concentrate on proving a pointwise version in x because the following purely analytical argument gives the local uniformity. If the pointwise result holds, it asserts that the sequence of nonincreasing functions $(n^{-1}\log \mathrm{P}\{S_n > g_{[0,n)}\cdot\})_{n\geq 1}$ converges to the function $-J_0$; since the limit is continuous, and monotone as a limit of monotone functions, the convergence is locally uniform (see [44], chapter 7, exercise 13).

Lemma 5.1.1 implies that $\gamma g_{[0,n)} \sim n g_n$ while Lemma 5.2.1 yields, in view of the fact that $\log\varphi_0$ is regularly varying with index β, the conjugate exponent to α, and $\alpha\gamma > 1$, that

$$n^{-1}\sum_{1\leq i\leq n}\log\varphi_0(\lambda g_{n-i/n}) \sim \int_0^1 \log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u$$

as n tends to infinity. The result follows from the Gartner–Ellis theorem (see e.g. [16], Theorem 2.3.6). ■

Proving Theorem 2.1 requires a couple more lemmas related to the function J_0.

Lemma 5.2.3. *The function J_0 is positive on the positive half-line.*

Proof. Let J_0^* be the function

$$J_0^*(\lambda) = \int_0^1 \log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u.$$

Both J_0^* and $J_0^{*\prime}$ vanish at the origin, while $J_0^{*\prime\prime}(0)$ is positive. In particular, referring to (5.1.2), taking λ to be $x/J_0^{*\prime\prime}(0)$, we see that as x tends to 0,

$$J_0(x) \geq \frac{x^2}{2J_0^{*\prime\prime}(0)} + o(x^2).$$

Thus, J_0 is positive on an open interval with left endpoint the origin. Since J_0 is a supremum of nondecreasing functions of x it is also nondecreasing and the result follows. ■

Lemma 5.2.4. *For any positive real number c, the function $x\in[0,\infty)\mapsto xJ_0(x^{-\gamma}+c)$ tends to infinity at 0 and infinity. Moreover, it reaches its minimum at a positive argument.*

Proof. Let c be a positive real number. Lemma 5.2.3 ensures that $J_0(c)$ is positive. It follows that $xJ_0(x^{-\gamma}+c)$ tends to infinity with x.

Assume that γ is at least 1. Since for any positive θ the inequality $J_0(x)\geq x\theta - J_0^*(\theta)$ holds, we see that J_0 ultimately grows faster than any multiple of the identity. Thus, $xJ_0(x^{-\gamma}+c)$ tends to infinity as x tends to 0, and this proves the lemma in this case.

Assume that γ is less than 1. The assumption $\alpha\gamma > 1$ ensures that $-\log\overline{F}$ is regularly varying of index α greater than 1. By Kasahara's [32] Tauberian theorem ([7], Theorem 4.12.7), $\log\varphi_0$ is regularly varying of index β, the exponent conjugate to α. This implies that J_0^* is also regularly varying of index β at infinity. By Bingham and Teugels' [8] theorem (see [7], Theorem 1.8.10), this implies that J_0 is regularly varying of index α. Since $\alpha\gamma$ is greater than 1, it then follows that $xJ_0(x^{-\gamma}+c)$ tends to infinity as x tends to 0 ([7], Proposition 1.3.6). This proves the first part of the lemma.

The second part of the lemma follows, because the function $xJ_0(x^{-\gamma}+c)$ is continuous on the positive half-line. ■

Our next lemma shows that the process S_n^0 is unlikely to reach a high threshold t before a time of order $V(t)$.

Lemma 5.2.5. *The following holds:*

$$\lim_{\epsilon\to 0}\limsup_{t\to\infty}\frac{1}{V(t)}\log \mathrm{P}\{\exists n\ :\ n\leq \epsilon V(t)\,;\ S_n^0>t\}=-\infty.$$

Proof. We distinguish according to whether $\max_{0\leq i\leq n} g_{i/n}$ remains bounded or not.

Assume first that $\limsup_{n\to\infty}\max_{0\leq i\leq n} g_{i/n}$ is some finite positive number c. Necessarily, γ is at least 1. In that case, (5.1.1) implies

$$\max_{0\leq i\leq \epsilon V(t)} g_i \lesssim \frac{c}{\gamma}g_{\lfloor \epsilon V(t)\rfloor}\sim c\epsilon^{\gamma-1}\frac{t}{V(t)}$$

as t tends to infinity. In particular, uniformly in i nonnegative and at most $\epsilon V(t)$, and as t tends to infinity, $t/g_i \gtrsim V(t)\epsilon^{1-\gamma}/c$. Moreover, Lemma 5.2.1 shows that

$$\sum_{1\leq i\leq n}\log\varphi_0(\lambda g_{n-i/n})\lesssim n\int_0^1\log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u$$

as n tends to infinity. Then, using the Markov exponential inequality, for any fixed positive λ, for any n large enough and at most $\epsilon V(t)$,

$$\log \mathrm{P}\{S_n^0>t\}\lesssim -\lambda\gamma\frac{V(t)}{2c}\epsilon^{1-\gamma}+2n\int_0^1\log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u$$

provided t is large enough, n is large enough and less than $\epsilon V(t)$.

Since γ is at least 1, for n at most $\epsilon V(t)$, this upper bound is at most

$$-V(t)\Big(\lambda\gamma\frac{\epsilon^{1-\gamma}}{2c}-2\epsilon\int_0^1\log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u\Big).$$

It can be made smaller than any negative multiple of $V(t)$ by first taking λ positive and then ϵ small enough. Hence, there exists n_0 such that

$$\lim_{\epsilon\to 0}\limsup_{t\to\infty}\max_{n_0\leq n\leq \epsilon V(t)}\frac{1}{V(t)}\log \mathrm{P}\{S_n^0>t\}=-\infty. \tag{5.2.2}$$

For n at most n_0, recalling that the mean of X_i is -1, we have, since t is positive,

$$P\{S_n^0>t\}\leq n_0\overline{F}_0\Big(\frac{t}{n_0\max_{0\leq i\leq n_0} g_i}\Big).$$

Since the moment generating function of F_0 is finite on the nonnegative half-line, Chernoff's inequality implies that $-\log\overline{F}\gg \mathrm{Id}$ at infinity. Lemma 5.1.2 shows that in the present case, U grows at most like a multiple of the identity at infinity. This implies that the function $U^{-1}\log\overline{F}$ tends to minus infinity at infinity. We conclude that (5.2.2) holds with n_0 being 1.

We now consider the case where $\max_{0\leq i\leq n} g_{i/n}$ tends to infinity with n. In this case, the sequence $(g_n)_{n\geq 0}$ converges to 0, and, for any η positive, $\log\varphi_0\lesssim \mathrm{Id}^{\beta+\eta}$ at infinity. Again, we use the exponential Markov inequality

$$\log \mathrm{P}\{S_n^0 > t\} \le -\lambda t + \sum_{0\le i<n} \log \varphi_0(\lambda g_i), \tag{5.2.3}$$

taking now λ of the form $cV(t)/t$ for some positive constant c to be determined.

Since the standard assumption ensures that $(g_n)_{n\ge 0}$ is asymptotically equivalent to a monotone sequence, $\min_{0\le i\le \epsilon V(t)} g_i \gtrsim g_{\lfloor \epsilon V(t)\rfloor}$ as t tends to infinity. Using (5.1.1), it follows that $\lambda g_i \gtrsim c\gamma\epsilon^{\gamma-1}$ is large whenever c is large and ϵ is small. Thus, provided c is large enough, ϵ is small enough and n is at most $\epsilon V(t)$,

$$\sum_{0\le i<n} \log \varphi_0(\lambda g_i) \le 2 \sum_{0\le i<n} (\lambda g_i)^{\beta+\eta} \le 2\Big(c\frac{V(t)}{t}\Big)^{\beta+\eta} \sum_{0\le i\le \epsilon V(t)} g_i^{\beta+\eta}. \tag{5.2.4}$$

Since $\beta(\gamma-1)+1$ is positive,

$$\sum_{0\le i<n} g_i^{\beta+\eta} \sim \frac{n}{1+(\gamma-1)(\beta+\eta)}\Big(\frac{g(1-1/n)}{\Gamma(\gamma)n}\Big)^{\beta+\eta}$$

as n tends to infinity, and the bound (5.2.4) is at most

$$2(\gamma c)^{\beta+\eta}\epsilon^{(\gamma-1)(\beta+\eta)+1}\frac{V(t)}{1+(\gamma-1)(\beta+\eta)}.$$

For any fixed ϵ, the upper bound (5.2.4) can be made less than any a priori given negative number times $V(t)$ by taking c large enough. This proves the lemma. ■

Proof of Theorem 2.1. Comparing (4.2) with Proposition 5.2.2, we may take r to be the identity, so that ρ is 1; furthermore, still referring to assumption (4.2) and Proposition 5.2.2, we see that $I(x) = J_0(x)$. Using (5.1.3), and since the mean μ is -1, it follows that θ, as defined in (4.5), is

$$\theta = \inf_{x>0} xJ_0(x^{-\gamma}+1) = \inf_{x>0} xJ(x^{-\gamma}). \tag{5.2.5}$$

The assumptions needed to apply the first part of Proposition 4.1 are satisfied thanks to Lemma 5.1.1, Proposition 5.2.2 and Lemma 5.2.5. Thus Proposition 4.1 yields Theorem 2.1 when μ is -1.

The result when μ is different than -1 is obtained by a simple rescaling. ■

5.3. *Proof of* Theorem 2.2

As for Theorem 2.1, we first prove Theorem 2.2 when μ is -1, which we assume from now on.

Our first lemma is an analogue of Lemma 5.2.1 but in the context of Theorem 2.2.

Note that in the context of Theorem 2.2, the conditions $\alpha\gamma < 1$ and $\gamma > 1/2$ force α to be less than 2. Therefore, its conjugate exponent, β, is greater than 2.

Lemma 5.3.1. *Let λ be a regularly varying function of index greater than $(2\gamma-1)/(\beta-2)$ and set $\lambda_n = \lambda(n)$. Under the assumptions of Theorem 2.2,*

$$\sum_{0\le i<n} \log \varphi_0(\lambda_n g_i) \sim \log \varphi_0(\lambda_n)|g|_\beta^\beta$$

as n tends to infinity.

Proof. Let σ^2 be the variance of X_i. Since $\log\varphi_0 \sim \mathrm{Id}^2\sigma^2/2$ at the origin, for any positive R there exists a positive c such that $\log\varphi_0 \le c\mathrm{Id}^2$ on $[\,0, R\,]$. Using Lemma 5.1.1, this implies

$$\sum_{0\le i<n} \mathbb{1}\{\lambda_n g_i \le R\}\log\varphi_0(\lambda_n g_i) \lesssim \lambda_n^2 \frac{g(1-1/n)^2}{n}. \tag{5.3.1}$$

Let ϵ be a positive real number. Using Potter's bound, we see that provided λ_n and $\lambda_n g_i$ are large enough, and provided that i is large enough for g_i to be less than 1,

$$\frac{1}{2}g_i^{\beta+\epsilon} \le \frac{\log\varphi_0(\lambda_n g_i)}{\log\varphi_0(\lambda_n)} \le 2g_i^{\beta-\epsilon}.$$

By a standard regular variation theoretic argument, this implies

$$\begin{aligned}\sum_{0\le i<n} \mathbb{1}\{\lambda_n g_i > R\}\log\varphi_0(\lambda_n g_i) &\sim \log\varphi_0(\lambda_n)\sum_{0\le i<n} g_i^{\beta}\mathbb{1}\{\lambda_n g_i > R\}\\ &\sim \log\varphi_0(\lambda_n)|g|_\beta^\beta \end{aligned} \tag{5.3.2}$$

as n tends to infinity — recall that $|g|_\beta$ is finite here, since $\beta(1-\gamma) > 1$.

Write ρ for the index of regular variation of λ. Since $\log\varphi_0\circ\lambda$ is regularly varying of index $\beta\rho$, and $\lambda^2 g(1-1/\mathrm{Id})^2/\mathrm{Id}$ is regularly varying of index $2\rho+2\gamma-1$, our assumption that ρ is greater than $(2\gamma-1)/(\beta-2)$ ensures that the right hand side of (5.3.2) dominates the right hand side of (5.3.1), and the result holds. ■

We define the Cramér transform of the centered random variables,

$$I_0(x) = \sup_{\lambda>0}\big(\lambda x - \log\varphi_0(\lambda)\big).$$

Recall that we assume that μ is -1. We then write F_0 for the distribution of the centered random variable $X_i + 1$.

We can now state and prove the following large deviations inequality.

Proposition 5.3.2. *Let $(S_n^0)_{n\ge0}$ be a centered (g, F_0)-process. Under the assumptions of Theorem 2.2, for any positive x,*

$$\log \mathrm{P}\{S_n^0 > x g_{[0,n)}\} \lesssim -\frac{x^\alpha}{|g|_\beta^\alpha} I_0(g_{[0,n)})$$

as n tends to infinity.

Proof. Recall that under the assumptions of Theorem 2.2, $\log\varphi_0$ is regularly varying of index β and I_0 is regularly varying of index α — see the proof of Lemma 5.2.4. Define

$$\lambda(t) = \frac{x^{1/(\beta-1)}}{|g|_\beta^\alpha} m_0^{\leftarrow}(g_{[0,t)}).$$

This function is regularly varying of positive index $\gamma/(\beta-1)$. We define λ_n as $\lambda(n)$. Using the exponential form of Markov's inequality and Lemma 5.3.1 – applicable since the inequality $\alpha\gamma < 1$ implies $\gamma/(\beta-1) > (2\gamma-1)/(\beta-2)$,

$$\begin{aligned}\log \mathrm{P}\{S_n^0 > x g_{[0,n)}\} &\le -\lambda_n x g_{[0,n)} + \sum_{0\le i<n}\log\varphi_0(\lambda_n g_i)\\ &\le -\frac{x^\alpha}{|g|_\beta^\alpha} g_{[0,n)} m_0^{\leftarrow}\circ g_{[0,n)} + |g|_\beta^\beta \frac{x^\alpha}{|g|_\beta^{\alpha\beta}}\log\varphi_0\circ m_0^{\leftarrow}(g_{[0,n)})\big(1+o(1)\big).\end{aligned}$$

Since by regular variation $\log\varphi_0 \sim \mathrm{Id}m_0/\beta$ at infinity, the right hand side of the above upper bound is asymptotically equivalent to

$$\frac{x^\alpha}{|g|_\beta^\alpha}(\mathrm{Id}m_0^{\leftarrow})(g_{[0,n)})(-1+1/\beta). \tag{5.3.3}$$

Upon noting that the maximizing value of λ in the definition of $I_0(\cdot)$ is $m_0^{\leftarrow}(\cdot)$, the chain rule yields $I_0' = m_0^{\leftarrow}$. Therefore, $\mathrm{Id}m_0^{\leftarrow} \sim \alpha I_0$ at infinity. We obtain that (5.3.3) is asymptotically equivalent to $-x^\alpha|g|_\beta^{-\alpha}I_0(g_{[0,n)})$ as n tends to infinity. This proves Proposition 5.3.2. ■

Our next result is yet another large deviations inequality. Its statement is suitable for our application, though its proof gives a somewhat more precise estimate.

Proposition 5.3.3. *For any positive real number ζ,*

$$\max_{1\le n<\zeta V(t)} \log \mathrm{P}\{S_n^0 > t\} \lesssim -\frac{I_0(t)}{|g|_\beta^\alpha}$$

as t tends to infinity.

Proof. Let $\lambda(t) = m_0^{\leftarrow}(t)|g|_\beta^{-\alpha}$. The exponential Markov inequality implies

$$\log \mathrm{P}\{S_n^0 > t\} \le -\lambda(t)t + \sum_{0\le i<n} \log\varphi_0\big(\lambda(t)g_i\big). \tag{5.3.4}$$

Using Potter's bound and regular variation of $\log\varphi_0$, there exists a positive R such that uniformly in n positive and less than $\zeta V(t)$,

$$\sum_{0\le i<n} \log\varphi_0\big(\lambda(t)g_i\big)\mathbb{1}\{\lambda(t)g_i > R\} \lesssim \log\varphi_0\big(\lambda(t)\big)|g|_\beta^\beta. \tag{5.3.5}$$

Moreover, as was shown in the proof of Lemma 5.3.1, there exists a positive real number c such that for any n less than $\zeta V(t)$,

$$\begin{aligned}\sum_{0\le i<n} \log\varphi_0\big(\lambda(t)g_i\big)\mathbb{1}\{\lambda(t)g_i \le R\} &\le c\sum_{0\le i<n}\lambda(t)^2 g_i^2\\ &\le c\lambda(t)^2 \sum_{0\le i<\zeta V(t)} g_i^2\\ &= O\Big(\lambda(t)^2\frac{t^2}{V(t)}\Big).\end{aligned}$$

As a function of t, this asymptotic upper bound is regularly varying of index $\frac{2}{\beta-1}+2-\frac{1}{\gamma} = 2\alpha-\frac{1}{\gamma}$. The upper bound (5.3.5) is regularly varying of index $\beta/(\beta-1)=\alpha$. Since $\alpha\gamma$ is less than 1, we see that $2\alpha-1/\gamma$ is less than α, and, consequently, for n less than $\zeta V(t)$,

$$\sum_{0\le i<n} \log\varphi_0\big(\lambda(t)g_i\big) \lesssim \log\varphi_0\big(\lambda(t)\big)|g|_\beta^\beta \sim (\log\varphi_0)\circ m_0^{\leftarrow}(t)|g|_\beta^{-\alpha}.$$

This implies that the exponent in the upper bound (5.3.4) is asymptotically bounded by an equivalent of

$$-|g|_\beta^{-\alpha}m_0^{\leftarrow}(t)t + |g|_\beta^{-\alpha}\log\varphi_0\circ m_0^{\leftarrow}(t) = -|g|_\beta^{-\alpha}I_0(t).$$

The result follows. ■

We now prove a trivial lower bound.

Lemma 5.3.4. *For any positive n,*

$$\log \mathrm{P}\{S_n > t\} \gtrsim \log \overline{F}(t)\Big(\sum_{0\le i<n} g_i^{\beta}\Big)^{-\alpha/\beta}$$

as t tends to infinity.

Proof. Let x_i be $g_i^{1/(\alpha-1)}/\sum_{0\le i<n} g_i^{\beta}$, so that $\sum_{0\le i<n} g_i x_i = 1$. We have

$$\log \mathrm{P}\{S_n > t\} = \sum_{0\le i<n} \log \overline{F}(tx_i) \sim \log \overline{F}(t) \sum_{0\le i<n} x_i^{\alpha}$$

as t tends to infinity. The result follows. ■

Proof of Theorem 2.2. *Lower bound.* Applying Lemma 5.3.4, for any positive integer n,

$$\begin{aligned}\log \mathrm{P}\{M > t\} &\ge \log \mathrm{P}\{S_n > t\}\\ &\gtrsim \log \overline{F}(t)\Big(\sum_{0\le i<n} g_i^{\beta}\Big)^{-\alpha/\beta}.\end{aligned}$$

Consequently, as t tends to infinity,

$$\log \mathrm{P}\{M > t\} \gtrsim \log \overline{F}(t)|g|_{\beta}^{-\alpha}.$$

Upper bound. We apply the remark following Proposition 4.1. In the present context, Proposition 5.3.2 shows that (4.8) holds with $r_n = I_0(g_{[0,n)})$ and $I(x) = |g|_{\beta}^{-\alpha} x^{\alpha}$. Note that $\log \overline{F}_0 \sim \log \overline{F}$ at infinity. Since Broniatowski and Fuchs' [9] Theorem 3.1 implies that, under the assumption of Theorem 2.2, $-\log \overline{F}_0 \sim I_0$ at infinity, the function r is regularly varying of index $\alpha\gamma$. Referring to Proposition 4.1, we see that $\theta = |g|_{\beta}^{-\alpha}$ for

$$\inf_{x\ge 0} x^{\alpha\gamma}(x^{-\gamma}+1)^{\alpha} = \inf_{x\ge 0}(1+x^{\gamma})^{\alpha} = 1.$$

Since $g_{[0,V(t))} \sim t$,

$$-r\circ V(t) = -I_0(g_{[0,V(t))}) \sim -I_0(t) \sim \log \overline{F}(t)$$

as t tends to infinity. Therefore, in view of this and Proposition 5.3.3, we see that condition (4.6) holds. This proves Theorem 2.2 when μ is -1. The same scaling argument as in the end of the proof of Theorem 2.1 allows for the extension to other values of μ. ■

6. Proof of Theorem 3.1

As for the proof of the results of Section 2, we will prove the result when the mean μ is -1. In the first two subsections, we prove assertions (i) and (ii). Assertion (iii) follows from assertion (ii) with standard arguments which are outlined in the last subsection. A scaling argument, which we omit, similar to that used to prove Theorem 2.1 yields Theorem 3.1 when the mean μ is arbitrary.

Throughout this section we will use the following obvious fact. Let E_t be an event indexed by t. To prove that $\mathrm{P}(E_t \mid M > t)$ tends to 0 as t tends to infinity, it suffices to prove that $\mathrm{P}(E_t) = o(\mathrm{P}\{M > t\})$ as t tends to infinity.

6.1. *Proof of* Theorem 3.1.*i*

Assume that μ is -1. The assumptions of Proposition 4.1 are satisfied by virtue of Proposition 5.2.2 and Lemma 5.2.5. From the second assertion of Proposition 4.1 and equality (5.1.3), we deduce that $\mathcal{N}_t$ converges to τ in probability as t tends to infinity and conditionally on M exceeding t. This is assertion (i) when the mean is -1.

6.2. *Proof of* Theorem 3.1.*ii*

We assume that μ is -1. Our next lemma is the analogue of Lemma 5.2.1 specialized to the context of the proof of Theorem 3.1. Recall that X_i has mean -1 for the time being, and that φ_0 is the moment generating function of the centered random variable $X_i + 1$.

Lemma 6.2.1. *Let f be a continuous real-valued and bounded function on $[\,0,1\,] \times \mathbb{R}$. For any fixed λ,*

$$\lim_{n\to\infty} n^{-1} \sum_{1\le i\le n} E\Big(f\Big(\frac{\mathrm{i}}{n}, X_i\Big) \frac{\mathrm{e}^{\lambda g_{n-i/n}(X_i+1)}}{\varphi_0(\lambda g_{n-i/n})}\Big)$$
$$= \int f(v,x) \frac{\mathrm{e}^{\lambda k_\gamma(v)(x+1)}}{\varphi_0\big(\lambda k_\gamma(v)\big)} \mathbb{1}_{[0,1)}(v)\mathrm{d}(L\otimes F)(v,x).$$

Proof. Let c be a number larger than $\lim_{n\to\infty} \max_{0\le i\le n} g_{i/n}$. Consider the function

$$\psi(v,y,x) = f(v,x)\frac{\mathrm{e}^{\lambda(y\wedge c)(x+1)}}{\varphi_0\big(\lambda(y\wedge c)\big)}.$$

For n large enough and with Γ_n the measure defined prior to Lemma 5.1.3,

$$n^{-1} \sum_{1\le i\le n} \mathrm{E}\Big(f\Big(\frac{\mathrm{i}}{n}, X_i\Big) \frac{\mathrm{e}^{\lambda g_{n-i/n}(X_i+1)}}{\varphi_0\big(\lambda g_{n-i/n}\big)}\Big) = \mathrm{E}\int \psi(v,y,X_1)\mathrm{d}\Gamma_n(v,y).$$

For any fixed x the function $\psi(v,y,x)$ is a continuous and bounded function of (v,y) in $[\,0,1\,]\times\mathbb{R}$. By Lemma 5.1.3, the sequence of functions

$$\psi_n(x) = \int \psi(v,y,x)\mathrm{d}\Gamma_n(v,y), \quad n\ge 1,$$

converges pointwise to the function

$$\psi(x) = \int_0^1 \psi\big(u, k_\gamma(u), x\big)\mathrm{d}u.$$

Since

$$|\psi(v,y,x)| \le |f|_{[0,1]\times\mathbb{R}} \mathrm{e}^{\lambda c|x+1|} \left|\frac{1}{\varphi_0}\right|_{[0,\lambda c]},$$

the dominated convergence theorem implies that $\mathrm{E}\psi_n(X_1)$ tends to $\mathrm{E}\psi(X_1)$ as n tends to infinity, which is what the lemma asserts. ■

Recall that in Section 2 we used the notation θ for

$$\theta = -\lim_{t\to\infty} V(t)^{-1}\log \mathrm{P}\{M>t\}.$$

Considering the definition of τ in (3.1), that of J_0 in (5.1.2), equality (5.1.3), and how θ was obtained in (5.2.5),

$$\begin{aligned}\theta &= \tau J_0(\tau^{-\gamma}+1)\\ &= \tau \sup_{\lambda>0}\Big((\tau^{-\gamma}+1)\lambda - \int_0^1 \log\varphi_0\big(\lambda k_\gamma(u)\big)\mathrm{d}u\Big).\end{aligned} \tag{6.2.1}$$

Since m_0 is onto the nonnegative half-line, the supremum in λ in the above formula is achieved for some value A. By considering the derivative in λ, which must vanish at the maximizer A, we obtain

$$\tau^{-\gamma}+1 = \int_0^1 k_\gamma(u) m_0\big(Ak_\gamma(u)\big)\mathrm{d}u. \tag{6.2.2}$$

When μ is -1 as currently, we have $\varphi(\lambda) = \mathrm{e}^{-\lambda}\varphi_0(\lambda)$ and, consequently, $m = -1 + m_0$. Therefore, the definition of A in (6.2.2) matches that in (3.2).

As will be apparent in the bound (6.2.8) to come and in its evaluation, the following result is strongly related to Proposition 5.2.2 if one takes n to be about $\tau V(t)$ and x to be about τ in that proposition.

Lemma 6.2.2. *The following holds:*

$$\lim_{\epsilon\to 0}\limsup_{t\to\infty}\sup_{n\,:\,|\frac{n}{V(t)}-\tau|<\epsilon}\left|A\gamma\frac{t+s_n}{g_n V(t)} - \frac{1}{V(t)}\sum_{1\le i\le n}\log\varphi_0(Ag_{n-i/n}) - \theta\right| = 0.$$

Proof. Write $n = \nu V(t)$. Since s is regularly varying, $s_n \sim \nu^\gamma t$. Moreover, (5.1.1) shows that

$$g_n \sim \gamma\nu^{\gamma-1}\frac{t}{V(t)},$$

and those equivalences hold locally uniformly in ν thanks to the uniform convergence Theorem ([7], Theorem 1.2.1). In particular,

$$\gamma\frac{t+s_n}{g_n V(t)} \sim \frac{1+\nu^\gamma}{\nu^{\gamma-1}}, \tag{6.2.3}$$

as t tends to infinity. Applying Lemma 5.2.1, we also have

$$\begin{aligned}\frac{1}{V(t)}\sum_{1\le i\le n}\log\varphi_0(Ag_{n-i/n}) &\sim \frac{n}{V(t)}\int_0^1 \log\varphi_0\big(Ak_\gamma(v)\big)\mathrm{d}v\\ &\sim \nu\int_0^1 \log\varphi_0\big(Ak_\gamma(v)\big)\mathrm{d}v,\end{aligned} \tag{6.2.4}$$

again locally uniformly in ν positive and as t tends to infinity. Combining (6.2.3) and (6.2.4), we obtain that

$$\begin{aligned}&A\gamma\frac{t+s_n}{g_n V(t)} - \frac{1}{V(t)}\sum_{1\le i\le n}\log\varphi_0(Ag_{n-i/n})\\ &\quad= \nu\Big(A(\nu^{-\gamma}+1) - \int_0^1 \log\varphi_0\big(Ak_\gamma(v)\big)\mathrm{d}v\Big) + o(1)\end{aligned} \tag{6.2.5}$$

as t tends to infinity. When ν is τ, equality (6.2.1) shows that the right hand side in (6.2.5) is θ. The result follows from the continuity in ν of the function involved in the right hand side of (6.2.5). ■

We can now prove the second assertion of Theorem 3.1. Let f be a continuous function supported by a vertical strip of the right half-space. Whenever ν is a measure on the right half-space, we write νf for $\int f \,\mathrm{d}\nu$. Let ϵ be a positive real number. Assume that we have proved that for any real number h greater than $\mathcal{M} f$,

$$\lim_{t\to\infty} \mathrm{P}\{\mathcal{M}_t f > h \mid M > t\} = 0. \tag{6.2.6}$$

If h is less than $\mathcal{M} f$, then applying the above relation to $-f$ and $-h$, we see that the conditional probability of $\mathcal{M}_t f < h$ given that M exceeds t tends to 0 as t tends to infinity. We then conclude that

$$\lim_{t\to\infty} \mathrm{P}\{|(\mathcal{M}_t - \mathcal{M}) f| > \epsilon \mid M > t\} = 0.$$

Thus, as t tends to infinity, $\mathcal{M}_t f$ converges in probability to $\mathcal{M} f$ conditionally on M exceeding t. Since f is arbitrary, this shows that $\mathcal{M}_t$ converges to $\mathcal{M}$ in probability, under the conditional probability that M exceeds t. This would prove the second assertion of Theorem 3.1, and therefore, it suffices to prove (6.2.6), which we do now.

The proof of the second assertion of Proposition 4.1 shows that for any positive ϵ,

$$\mathrm{P}\{\mathcal{M}_t f > h\,;\; M > t\} \le \sum_{n\,:\,|\frac{n}{V(t)}-\tau|\le\epsilon} \mathrm{P}\{\mathcal{M}_t f > h;\, S_n^0 > t + s_n\} + o(\mathrm{P}\{M > t\}) \tag{6.2.7}$$

as t tends to infinity.

The basic inequality for our proof is the exponential form of Markov's, which implies that for any positive λ,

$$\begin{aligned}
&\log \mathrm{P}\{\mathcal{M}_t f > h;\, S_n^0 > t + s_n\} \\
&\quad\le \log \mathrm{P}\Big\{\lambda V(t)\mathcal{M}_t f + A\gamma \frac{S_n^0}{g_n} > \lambda V(t) h + A\gamma \frac{t+s_n}{g_n}\Big\} \\
&\quad\le -V(t)\Big(\lambda h + A\gamma \frac{t+s_n}{g_n V(t)} - \frac{1}{V(t)} \log \mathrm{E} \exp\Big(\lambda V(t)\mathcal{M}_t f + A\gamma \frac{S_n^0}{g_n}\Big)\Big).
\end{aligned} \tag{6.2.8}$$

Now, we define some small – arguably, bewildering – constants. Let δ be a positive real number less than 1 such that $h > (1+2\delta)\mathcal{M} f$. Let λ be positive and small enough that

$$\lambda |f|_{[0,\infty)\times\mathbb{R}} < \sup\{x \,:\, \mathrm{e}^x < 1 + (1+\delta)x\} \wedge (1+\delta)^{-1}. \tag{6.2.9}$$

Next, let η be small enough that $\lambda\big(h - (1+2\delta)\mathcal{M} f\big) > 3\eta$. Finally, using Lemma 6.2.2, let ϵ be a positive real number so that

$$\limsup_{t\to\infty} \sup_{n\,:\,|\frac{n}{V(t)}-\tau|<\epsilon} \left| A\gamma \frac{t+s_n}{g_n V(t)} - \frac{1}{V(t)} \sum_{1\le i\le n} \log \varphi_0(A g_{n-i/n}) - \theta \right| < \eta. \tag{6.2.10}$$

To evaluate the upper bound (6.2.8), we first bound the term containing an expectation. Given the constraint (6.2.9) on λ,

$$\mathrm{e}^{\lambda f(i/V(t), X_i)} \le 1 + (1+\delta)\lambda f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big).$$

Recall that X_i is of mean -1 currently. Since

$$\lambda V(t)\mathcal{M}_t f + A\gamma S_n^0/g_n$$
$$= \sum_{1\le i\le n}\Big(\lambda f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big) + Ag_{n-i/n}(X_i+1)\Big) + \sum_{i>n}\lambda f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big),$$

the term $\mathrm{E}\exp\Big(\lambda V(t)\mathcal{M}_t f + A\gamma \frac{S_n^0}{g_n}\Big)$ in (6.2.8) is at most

$$\prod_{1\le i\le n}\mathrm{E}\Big(1+(1+\delta)\lambda f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big)\Big)\mathrm{e}^{Ag_{n-i/n}(X_i+1)}\prod_{i>n}\mathrm{E}\Big(1+(1+\delta)\lambda f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big)\Big). \tag{6.2.11}$$

Note the inequality $\log(a+b)\le \log a + b/a$, valid for any positive a and any b larger than $-a$. To apply this inequality with

$$a = \mathrm{E}\mathrm{e}^{Ag_{n-i/n}(X_i+1)}$$

and

$$b = (1+\delta)\lambda \mathrm{E} f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big)\mathrm{e}^{Ag_{n-i/n}(X_i+1)},$$

we first observe that

$$|b| \le (1+\delta)\lambda |f|_{[0,\infty)\times\mathbb{R}}\mathrm{E}\mathrm{e}^{Ag_{n-i/n}(X_i+1)}$$

and (6.2.9) ensures that $|b|$ is less than a. Therefore, $a+b$ is positive. We then have, referring to the first product of (6.2.11),

$$\log \mathrm{E}\Big(1+(1+\delta)\lambda f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big)\Big)\mathrm{e}^{Ag_{n-i/n}(X_i+1)}$$
$$\le \log \mathrm{E}\mathrm{e}^{Ag_{n-i/n}(X_i+1)} + (1+\delta)\lambda\frac{\mathrm{E} f\big(i/V(t), X_i\big)\mathrm{e}^{Ag_{n-i/n}(X_i+1)}}{\mathrm{E}\mathrm{e}^{Ag_{n-i/n}(X_i+1)}}$$
$$= \log\varphi_0(Ag_{n-i/n}) + (1+\delta)\lambda\mathrm{E}\Big(f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big)\frac{\mathrm{e}^{Ag_{n-i/n}(X_i+1)}}{\varphi_0(Ag_{n-i/n})}\Big).$$

Consequently, using the inequality $\log(1+x)\le x$ to handle the second product in the upper bound (6.2.11), we see that the logarithm of (6.2.11) is at most

$$\sum_{1\le i\le n}\log\varphi_0(Ag_{n-i/n}) + (1+\delta)\lambda\sum_{1\le i\le n}\mathrm{E}\Big(f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big)\frac{\mathrm{e}^{Ag_{n-i/n}(X_i+1)}}{\varphi_0(Ag_{n-i/n})}\Big)$$
$$+(1+\delta)\lambda\sum_{i>n}\mathrm{E} f\Big(\frac{\mathrm{i}}{V(t)}, X_i\Big).$$

Referring to the upper bound (6.2.8),

$$\lambda h + A\gamma\frac{t+s_n}{g_n V(t)} - \frac{1}{V(t)}\log \mathrm{E}\exp\Big(\lambda V(t)\mathcal{M}_t f + A\gamma\frac{S_n^0}{g_n}\Big)$$

is then at least

$$\begin{aligned}
&A\gamma\frac{t+s_n}{g_nV(t)}-\frac{1}{V(t)}\sum_{1\le i\le n}\log\varphi_0(Ag_{n-i/n})+\lambda h\\
&\quad-(1+\delta)\lambda\frac{n}{V(t)}\frac{1}{n}\sum_{1\le i\le n}\mathrm{E}\Big(f\Big(\frac{\mathrm{i}}{V(t)},X_i\Big)\frac{\mathrm{e}^{Ag_{n-i/n}(X_i+1)}}{\varphi_0(Ag_{n-i/n})}\Big)\\
&\quad-(1+\delta)\lambda\frac{n}{V(t)}\frac{1}{n}\sum_{i>n}\mathrm{E}f\Big(\frac{\mathrm{i}}{V(t)},X_i\Big).
\end{aligned}\tag{6.2.12}$$

Define ν as $n/V(t)$. Using (6.2.10), Lemma 6.2.1, and the equality $\varphi_0(\lambda)=\mathrm{e}^{\lambda}\varphi(\lambda)$ valid here since μ is -1, we obtain that (6.2.8) is at most the exponential of $-V(t)$ times

$$\begin{aligned}
&\theta-\eta+\lambda h-(1+\delta)\lambda\nu\int f(v\nu,x)\frac{\mathrm{e}^{A\gamma(1-v)^{\gamma-1}(x+1)}}{\varphi_0\big(A\gamma(1-v)^{\gamma-1}\big)}\mathbb{1}_{[0,1)}(v)\mathrm{d}(L\otimes F)(v,x)\\
&\qquad-(1+\delta)\lambda\nu\int f(v\nu,x)\mathbb{1}_{[1,\infty)}(v)\mathrm{d}(L\otimes F)(v,x)\\
&\quad=\theta-\eta+\lambda\Big(h-(1+\delta)\frac{\nu}{\tau}\int f\Big(\frac{\nu}{\tau}v,x\Big)\mathrm{d}\mathcal{M}(v,x)\Big).
\end{aligned}\tag{6.2.13}$$

If ϵ is small enough that ν/τ is close enough to 1, then

$$\Big|\int\frac{\nu}{\tau}f\Big(\frac{\nu}{\tau}v,x\Big)\mathrm{d}\mathcal{M}(v,x)-\mathcal{M}f\Big|<\eta/\lambda$$

and (6.2.13) is at least

$$\theta-\eta+\lambda\big(h-(1+\delta)\mathcal{M}f\big)-\eta,$$

which, by our choice of η, is greater than $\theta+(1-\delta)\eta$. Hence

$$\log\mathrm{P}\{\mathcal{M}_tf>h;\ S_n^0>t+s_n\}\lesssim-V(t)\big(\theta+(1-\delta)\eta\big)$$

as t tends to infinity. Since V is regularly varying, (6.2.7) shows that (6.2.6) holds, and this proves assertion (ii) of Theorem 3.1 when μ is -1.

6.3. *Proof of* Theorem 3.1.*iii*

In essence, the proof consists in writing the process $\mathcal{S}_t$ as a functional of $\mathcal{M}_t$ and showing that the convergence of $\mathcal{M}_t$ to $\mathcal{M}$ implies that of the functional of $\mathcal{M}_t$ to the functional of $\mathcal{M}$. The main difficulty is that the functional is not continuous with respect to our topology on measures. This forces us to develop various approximation results to show that $\mathcal{S}_t$ is approximable by a well behaved functional of $\mathcal{M}_t$. The proofs are rather routine in large deviation theory and are omitted.

To proceed with a sketch of the proof, for any measure ν on the right half-space for which the integrals

$$\int\mathbb{1}_{[0,\lambda)}(v)(\lambda-v)^{\gamma-1}|x|\mathrm{d}\nu(v,x),\quad\lambda>0,$$

are finite, we define the functional $\mathfrak{S}$ of ν evaluated at λ by

$$\mathfrak{S}(\nu)(\lambda)=\int\mathbb{1}_{[0,\lambda)}(v)\gamma(\lambda-v)^{\gamma-1}x\mathrm{d}\nu(v,x),$$

with the convention that $\mathfrak{S}(\nu)(0)$ is 0.

The proof then consists in showing that provided t is large, $\mathcal{S}_t$ is locally uniformly close to $\mathfrak{S}(\mathcal{M}_t)$ given $M > t$. Then one proves that Theorem 3.1.ii implies that $\mathfrak{S}(\mathcal{M}_t)$ converges locally uniformly to $\mathfrak{S}(\mathcal{M})$ under the conditional distribution given $M > t$. Since the functional $\mathfrak{S}$ is not continuous for the topology that we are using on measures, the argument consists in approximating it in a natural way by a continuous functional, using a standard technique from large deviations theory — see e.g. [1,26,28].

References

[1] R.R. Bahadur, Some Limit Theorems in Statistics, SIAM, 1971.

[2] G. Balkema, P. Embrechts, High Risk Scenarios and Extremes, a Geometric Approach, European Mathematical Society, 2007.

[3] Ph. Barbe, M. Broniatowski, Note on functional large deviation principle for fractional ARIMA processes, Statist. Inference. Stoch. Process. 1 (1998) 17–27.

[4] Ph. Barbe, W.P. McCormick, Veraverbeke's theorem at large — on the maximum of some processes with negative drifts and heavy tail innovations, 2008. Preprint arXiv:0802.3638.

[5] O.E. Barndorff-Nielsen, Information and Exponential Families in Statistical Theory, Wiley, 1978.

[6] P. Billingsley, Convergence of Probability Measures, Wiley, 1968.

[7] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, 2nd ed., Cambridge University Press, 1989.

[8] N.H. Bingham, J. Teugels, Duality for regularly varying functions, Quarterly J. Math. 26 (1975) 333–353.

[9] M. Broniatowski, A. Fuchs, Tauberian theorems, Chernoff inequality, and the tail behavior of finite convolutions of distribution functions, Adv. Math. 116 (1995) 12–33.

[10] L.D. Brown, Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory, IMS, 1986.

[11] R. Burton, H. Dehling, Large deviations for some weakly dependent random processes, Statist. Probab. Lett. 9 (1990) 397–401.

[12] C.-S. Chang, D.D. Yao, T. Zajic, Large deviations, moderate deviations, and queues with long-range dependent input, Adv. Appl. Probab. 31 (1999) 254–277.

[13] J. Collamore, Hitting probabilities and large deviations, Ann. Probab. 24 (1996) 2065–2078.

[14] J. Collamore, First passage times of general sequences of random vectors: A large deviations approach, Stochastic. Process. Appl. 78 (1998) 97–130.

[15] I. Csiszár, Sanov property, generalized I-projection and a conditional limit theorem, Ann. Probab. 12 (1984) 768–793.

[16] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Jones and Bartlett, 1993.

[17] P. Diaconis, D. Freedman, Conditional limit theorems for exponential families and finite version of de Finetti's theorem, J. Theoret. Probab. 1 (1988) 381–410.

[18] A.B. Dieker, M. Mandjes, Efficient simulation of random walks exceeding a nonlinear boundary, Stoch. Models 22 (2006) 459–481.

[19] N.G. Duffield, N. O'Connell, Large deviations and overflow probabilities for the general single-server queue, with applications, Math. Proc. Cambridge. Philos. Soc. 118 (1995) 363–374.

[20] N.G. Duffield, W. Whitt, Large deviations of inverse processes with nonlinear scalings, Ann. Appl. Probab. 4 (1998) 995–1026.

[21] R. Ellis, Large deviations for a general class of random vectors, Ann. Probab. (1984) 1–12.

[22] W. Feller, An Introduction to Probability Theory and its Applications, Wiley, 1971.

[23] M.I. Freidlin, A.D. Wentzell, Random Perturbation of Dynamical Systems, Springer, 1984.

[24] J. Gärtner, On large deviations from invariant measure, Theoret. Probab. Appl. 22 (1977) 24–39.

[25] H.U. Gerber, Ruin theory in the linear model, Insurance Math. Econo. 1 (1982) 177–184.

[26] P. Groeneboom, J. Oosterhoff, F. Ruymgaart, Large deviation theorems for empirical probability measures, Ann. Probab. 7 (1979) 553–586.

[27] J.M. Hammersley, D.G. Handscomb, Monte Carlo Methods, Chapman & Hall, 1964.

[28] A.B. Hoadley, On the probability of large deviations of functions of several empirical cdf's, Ann. Math. Statist. 38 (1967) 360–381.

[29] J. Hüsler, V. Piterbarg, On the ruin probability for physical fractional Brownian motion, Stochastic. Process. Appl. 113 (2004) 315–332.

[30] J. Iscoe, P. Ney, E. Nummelin, Large deviations of uniformly recurrent Markov additive processes, Adv. Appl. Math. 6 (1985) 373–412.
[31] J. Janssen, On the interaction between risk and queueing theories, Bl. DGVFM 15 (1982) 383–395.
[32] Y. Kasahara, Tauberian theorems of exponential type, J. Math. Kyoto Univ. 18 (1978) 209–219.
[33] H. Kesten, Random difference equations and renewal theory for product of random matrices, Acta. Math. 131 (1973) 207–248.
[34] G. Letac, Lectures on Exponential Families and their Variance Functions, IMPA, 1992.
[35] T. Lindvall, Weak convergence of probability measures and random functions in the function space $D[0, \infty)$, J. Appl. Probab. 10 (1973) 109–121.
[36] A.A. Mogulskii, Large deviations for trajectories of multi-dimensional random walks, Theoret. Probab. Appl. 21 (1976) 300–315.
[37] A. Müller, G. Pflug, Asymptotic ruin probabilities for risk processes with dependent increments, Insurance Math. Econom. 28 (2001) 381–392.
[38] H. Nyrhinen, Rough limit results for level-crossing probabilities, J. Appl. Probab. 31 (1994) 373–382.
[39] H. Nyrhinen, On the typical level crossing time and path, Stochastic. Process. Appl. 58 (1995) 11–137.
[40] H. Nyrhinen, Rough descriptions of ruin for a general class of surplus processes, Adv. Appl. Probab. 30 (1998) 1008–1026.
[41] N.U. Prabhu, On the ruin problem of collective risk theory, Ann. Math. Statist. 32 (1961) 757–764.
[42] S.D. Promislow, The probability of ruin in a process with dependent increments, Insurance Math. Econom. 10 (1991) 99–107.
[43] R.T. Rockafellar, Convex Analysis, Princeton University Press, 1970.
[44] W. Rudin, Principle of Mathematical Analysis, 3rd ed., McGraw-Hill, 1976.
[45] J. Sadowsky, On Monte Carlo estimation of large deviations probabilities, Ann. Appl. Probab. 6 (1996) 399–422.