EXACT ASYMPTOTICS FOR A MULTI-TIMESCALE MODEL

BY MARISKA HEEMSKERK & MICHEL MANDJES

Abstract. In this paper we study the probability $\xi_n(u) := \mathbb{P}(C_n \geq un)$, with $C_n := A(\psi_n B(\varphi_n))$ for Lévy processes $A(\cdot)$ and $B(\cdot)$, and $\varphi_n$ and $\psi_n$ non-negative sequences such that $\varphi_n \psi_n = n$ and $\varphi_n \to \infty$ as $n \to \infty$. Two timescale regimes are distinguished: a ‘fast’ regime in which $\varphi_n$ is superlinear and a ‘slow’ regime in which $\varphi_n$ is sublinear. We provide the exact asymptotics of $\xi_n(u)$ (as $n \to \infty$) for both regimes, relying on change-of-measure arguments in combination with Edgeworth-type estimates. The asymptotics have an unconventional form: the exponent contains the commonly observed linear term, but may also contain sublinear terms (the number of which depends on the precise form of $\varphi_n$ and $\psi_n$). To showcase the power of our results we include two examples, covering both the case where $C_n$ is lattice and non-lattice. Finally we present numerical experiments that demonstrate the importance of taking into account the doubly stochastic nature of $C_n$ in a practical application related to customer streams in service systems; they show that the asymptotic results obtained yield highly accurate approximations, also in scenarios in which there is no pronounced timescale separation.

2010 Mathematics Subject Classification. Primary: 60F10. Secondary: 60G51, 60K37.
Keywords. Lévy processes ◦ Multi-timescale model ◦ Exact asymptotics ◦ Edgeworth expansions ◦ Overdispersion

1. INTRODUCTION, PRELIMINARIES, NOTATION, LITERATURE

Consider a scalar Lévy process $A(\cdot)$, and (independently of $A(\cdot)$) an increasing scalar Lévy process $B(\cdot)$. These are uniquely characterized by their characteristic exponents, which are defined as

$$\alpha(\vartheta) := \log \mathbb{E}e^{\vartheta A(1)}, \quad \beta(\vartheta) := \log \mathbb{E}e^{\vartheta B(1)},$$

and are in fact the logarithmic moment generating functions of $A(1)$ and $B(1)$. Targeting at large-deviations asymptotics we impose the assumption that both characteristic exponents are finite in an open neighborhood of the origin, implying that all moments of $A(t)$ and $B(t)$ exist. Let $\varphi_n$ and $\psi_n$ be non-negative sequences such that $\varphi_n \psi_n = n$ and $\varphi_n \to \infty$ as $n \to \infty$. Our objective is to find the exact asymptotics of

$$\xi_n(u) := \mathbb{P}(A(\psi_n B(\varphi_n)) \geq un),$$

i.e., we wish to find a sequence $f_n$ such that $\xi_n(u)/f_n \to 1$ as $n \to \infty$. We assume $u > ab$, with $a := \mathbb{E}A(1) > 0$ and $b := \mathbb{E}B(1) > 0$, such that the event under consideration becomes increasingly rare as $n \to \infty$. One special case has been studied in detail: the choice $\varphi_n = n$ and $\psi_n = 1$ reduces $\xi_n(u)$ to $\mathbb{P}(A(B(n)) \geq un)$, whose exact asymptotics follow from [2] (using that $A(B(\cdot))$ is a Lévy process). To the best of our knowledge, other cases have not been analyzed in the literature.

Effect of multiple timescales. To get some intuition for the behavior of $\xi_n(u)$, we write it as

$$\xi_n(u) = \mathbb{P} \left( A \left( n \frac{B(\varphi_n)}{\varphi_n} \right) \geq un \right).$$

We proceed by explaining that there are two timescale regimes. (i) If $\varphi_n$ is superlinear, it is anticipated that $B(\varphi_n)/\varphi_n$ is close to $b$, such that $\xi_n(u)$ resembles $\mathbb{P}(A(bn) \geq un)$. We refer to this setting as the ‘fast regime’, as the fluctuations of the process $B(\cdot)$ are so fast that it can be replaced by its mean value. (ii) If on the contrary $\varphi_n$ is sublinear (which we will refer to as the ‘slow regime’ for obvious reasons), then one may expect that the event of interest roughly looks like $anB(\varphi_n)/\varphi_n \geq un$, and therefore $\xi_n(u)$ essentially behaves as $\mathbb{P}(aB(\varphi_n) \geq u\varphi_n)$. The objective of this paper is to make these claims precise. Our main contribution is that we succeed in identifying the exact asymptotics of $\xi_n(u)$ as $n \to \infty$. These turn out to have a non-standard form, in the sense that the exponent, in addition to the linear term that also appears in the classical asymptotics [10], may also contain sublinear terms.
To further investigate the two timescale regimes identified above, it is instructive to calculate the variance of \( C_n := A(\psi_n B(\varphi_n)) \). To this end, we first express the log-moment generating function (l-mgf) \( \gamma_n(\cdot) \) of \( C_n \) in terms of \( \alpha(\cdot) \) and \( \beta(\cdot) \). It requires a direct computation to verify that

\[
\gamma_n(\theta) = \varphi_n \beta(\alpha(\theta) \psi_n).
\]

Then it is direct that \( \mathbb{E} C_n = n\alpha'(0)\beta'(0) = nab \) and

\[
\operatorname{Var} C_n = \gamma''_n(0) = n\psi_n(\alpha'(0))^2 \beta''(0) + n\alpha''(0)\beta'(0) = n\psi_n\sigma^2_2 + n\sigma^2_1,
\]

with \( \sigma^2_1 := \alpha(0)^2 \beta''(0) \) and \( \sigma^2_2 := \alpha''(0)\beta \). In this decomposition of the variance, the aforementioned regime dichotomy is nicely reflected, as can be seen as follows. If \( \varphi_n \) is superlinear (and hence \( \psi_n \) vanishes), then the first term in the right-hand side of (1) is small relative to the second term, and \( \operatorname{Var} C_n \) essentially behaves as \( n\sigma^2_1 \) (which is also the variance of \( A(bn) \)). On the other hand, if \( \varphi_n \) grows sublinearly then so does \( \psi_n \), so that in this case the first term of (1) will dominate; as a consequence, \( \operatorname{Var} C_n \) behaves as \( n\psi_n\sigma^2_2 \) (which equals the variance of \( anB(\varphi_n)/\varphi_n \)).

The above intuition can be translated in terms of a central limit theorem by following a classical approach. In the fast regime (in which \( \varphi_n \) is superlinear), direct computations reveal that the characteristic function of

\[
D_n := \frac{C_n - (ab)n}{\sqrt{n\sigma_+}}
\]

converges to that of a standard Normal random variable. Hence, as a direct application of Lévy’s convergence theorem, \( D_n \) converges in distribution to a standard Normal random variable. Likewise, in the slow regime (in which \( \varphi_n \) is sublinear)

\[
E_n := \frac{C_n - (ab)n}{\sqrt{n\psi_n\sigma_-}}
\]

converges in distribution to a standard Normal random variable. We conclude that the dichotomy described above manifests itself in a central limit context.

**Large deviations, contributions.** In this paper we assess to what extent the above dichotomy carries over to a large-deviations setting. Based on the above reasoning, it is tempting to believe the following conjecture: in the fast regime \( \xi_n(u)/\mathbb{P}(A(bn) \geq un) \to 1 \) as \( n \to \infty \), and in slow regime \( \xi_n(u)/\mathbb{P}(aB(\varphi_n) \geq u\varphi_n) \to 1 \) as \( n \to \infty \). Our study, however, shows that this conjecture does not hold in general. In more detail, denoting by \( k_n \sim \ell_n \) that \( k_n/\ell_n \to 1 \) as \( n \to \infty \), the main result is that we find explicit sequences \( \lambda_{+,n} \) and \( \lambda_{-,n} \) such that in the fast regime, as \( n \to \infty \)

\[
\xi_n(u) \sim \mathbb{P}(A(bn) \geq un)\lambda_{+,n},
\]

and in the slow regime, as \( n \to \infty \),

\[
\xi_n(u) \sim \mathbb{P}(aB(\varphi_n) \geq u\varphi_n)\lambda_{-,n}.
\]

Here, \( \lambda_{+,n} \) and \( \lambda_{-,n} \) do not necessarily equal 1, thus refuting the above conjecture. Note that hereby the exact asymptotics of \( \xi_n(u) \) are fully identified, as sequences \( \kappa_{+,n} \) and \( \kappa_{-,n} \) such that \( \mathbb{P}(A(bn) \geq un)\kappa_{+,n} \to 1 \) and \( \mathbb{P}(aB(\varphi_n) \geq u\varphi_n)\kappa_{-,n} \to 1 \) are given in [2]. (As an aside, we mention that in specific situations \( \lambda_{+,n} \) and \( \lambda_{-,n} \) do equal 1, so that in those situations the conjecture applies; we return to this below.)

The resulting exact asymptotics of \( \xi_n(u) \) can be written as the product of a polynomial and an exponential part. In the fast regime, the polynomial part is inversely proportional to \( \sqrt{n} \), as was found before in various related settings (such as the one studied in [2]). The exponential part, however, has a rather unusual shape: not only does the exponent contain a commonly observed term that is linear in \( n \), in addition it consists of finitely many sublinear terms (the number of which depends on the specific form of \( \varphi_n \) and \( \psi_n \)). In the slow regime similar results apply, but with the role of \( n \) taken over by \( \varphi_n \), meaning that the polynomial part is inversely proportional to...
$\sqrt{\varphi_n}$ and the exponent is the sum of a term that is linear in $\varphi_n$ and finitely many terms that are $o(\varphi_n)$.

An immediate consequence of our result is that the conjecture we stated above is true if the timescales of both Lévy processes are sufficiently separated: then there are no sublinear terms in the exponent (where ‘sublinear’ means $o(n)$ in the fast regime and $o(\varphi_n)$ in the slow regime). More specifically, the conjecture holds in the fast regime if $n\psi_n \to 0$ (i.e., then $\lambda_{+n} = 1$), and in the slow regime if $\varphi_n / \psi_n \to 0$ (i.e., then $\lambda_{-n} = 1$). It is further remarked that on a logarithmic scale the conjecture always holds, albeit in the following (weaker) sense: we show that in the fast regime

$$\lim_{n \to \infty} \frac{1}{n} \log \xi_n(u) = \lim_{n \to \infty} \frac{1}{n} \log P(A(bn) \geq un),$$

whereas in the slow regime

$$\lim_{n \to \infty} \frac{1}{\varphi_n} \log \xi_n(u) = \lim_{n \to \infty} \frac{1}{\varphi_n} \log P(ab(\varphi_n) \geq u\varphi_n).$$

**Example 1.** The leading example, which we will visit several times in this paper, is that of $\varphi_n \sim n^f$ and $\psi_n \sim n^{1-f}$ for some $f > 0$. The fast regime corresponds to $f > 1$ and the slow regime to $f \in (0, 1)$. The above criterion yields that $\lambda_{+n} = 1$ if $f > 2$ and $\lambda_{-n} = 1$ if $f \in (0, \frac{1}{2})$.

**Remark 1.** We mentioned before that the case $\varphi_n = n$ can be dealt with in a straightforward way. For the sake of completeness, we include this argumentation here. In this case $\gamma_n(\vartheta) = n\beta(\alpha(\vartheta))$, which effectively means that $C_n\varphi_n$ can be written as the sum of $n$ i.i.d. random variables, bringing us directly in the setting of [2]. With $\vartheta^*$ solving $\beta'(\alpha(\vartheta))\alpha'(\vartheta) = 0$, one obtains

$$\xi_n(u) \sim \frac{1}{\vartheta^* \sigma_0 \sqrt{2\pi n}} \exp ((\beta(\alpha(\vartheta^*)) - \vartheta^* u)n),$$

where $\sigma_0 := \beta''(\vartheta^*)(\alpha'(\vartheta^*))^2 + \beta'(\alpha(\vartheta^*))\alpha''(\vartheta^*)$.

We proceed with a few words on the approach followed. In the large-deviations analysis a crucial role is played by the ‘twisting factor’, i.e., the solution $\vartheta^* = \vartheta_n$ of the equation $un = \gamma_n'(\vartheta)$, or

$$u = \beta'(\alpha(\vartheta) \psi_n) \alpha'(\vartheta).$$

As $u > ab$, it follows that $\vartheta_n$ is positive. In the fast regime, $\vartheta_n$ is close to the solution $\vartheta^*$ of $u = \beta a'(\vartheta^*)$; in the slow regime $\vartheta_n$ resembles $\tau^* / \psi_n$ where $\tau^*$ solves $u = \beta'(\alpha(\vartheta)^*) a$. Our proofs rely on a change of measure via an exponential twist that is based on the solution of (2). By and large, the proof presented in [10] underlying the Bahadur-Rao [2] result is followed: first we capture the exponential part of $\xi_n(u)$ by applying a change of measure, after which the polynomial part of $\xi_n(u)$ is identified using delicate Berry-Esséen-based calculations (that are considerably more involved than the ones needed in the ‘classical’ Bahadur-Rao case). In line with the proof in [10], the case where the underlying Lévy processes are lattice has to be dealt with separately.

**Literature and motivation.** Our work fits in the tradition of large-deviations asymptotics of sample-mean related quantities in the light-tailed setting. Without attempting to provide an exhaustive overview, we give a number of key references.

In the classical framework, $S_n$ is defined as the sum of i.i.d. random variables $X_1, \ldots, X_n$, where the $X_i$ are assumed to have a finite moment generating function in a neighborhood of the origin. The exceedance probability $\chi_n(u) := P(S_n \geq un)$ is the object of study, with $u > \mathbb{E}X_1$. Cramér [9] characterized a function $I(u)$ such that $\frac{1}{n} \log \chi_n(u) \to -I(u)$ as $n \to \infty$; this type of results is commonly referred to as logarithmic asymptotics. The proof technique used relied on change-of-measure argumentation that has been applied extensively since then.

Cramér’s seminal work was extended in several directions. In [8] a uniform upper bound on $\chi_n(u)$, generally known as the Chernoff bound, was derived. We refer to [3] for a generalization
of Cramér’s logarithmic asymptotics to infinite-dimensional topological vector spaces. Cramér’s theorem was also extended to sums of dependent random variables [20] and vectors [11, 13].

Whereas the above results focus on logarithmic asymptotics, another strand of research addresses exact asymptotics. In this respect we mention the pioneering paper [2] that shows that, under the conditions of Cramér’s result, \( \chi_n(u) e^{n I(u)} \sqrt{n} \) converges to a positive constant as \( n \to \infty \); in [6] a similar result had already been derived for the case that the \( X_i \) are lattice. We refer to e.g. [7] for exact asymptotics in the vector-valued case, and to [17] for a uniform framework covering both the CLT regime and large deviations.

Our investigations were motivated by models recently suggested in order to incorporate overdispersion. The reason for developing such models was the observation that in various types of service systems [5, 21] the customer arrival process is intrinsically more variable than the traditionally used Poisson process. An approach to overcome the lack of overdispersion was proposed in [16]: extra variability is produced by periodically resampling the Poisson arrival rate. As a consequence, when resampling every unit of time, the number of arrivals in \([0, m]\) denoted by \( C(m) \) (for \( m \in \mathbb{N} \)) is Poisson distributed with (random) parameter \( \sum_{i=1}^{m} \Lambda_i \), where the \( \Lambda_i \) are i.i.d. non-negative random variables. Assuming for simplicity that the \( \Lambda_i \) are integer-valued as well, this means that

\[
C(m) \overset{d}{=} \sum_{j=1}^{\sum_{i=1}^{m} \Lambda_i} X_j,
\]

(3)

with \( X_j \) a sequence of i.i.d. Poisson random variables with parameter 1. We thus have introduced overdispersion, as desired; more specifically, we have

\[
\text{Var} \ C(m) = m \ E \Lambda \text{Var} \ X + m \ (E X)^2 \text{Var} \Lambda \geq m \ E \Lambda \text{Var} \ X = m \ E \Lambda = E \ C(m).
\]

making the process more variable than the Poisson process. Now observe that the ‘two-timescale random walk’ (3) we introduced can be seen as the discrete counterpart of the two-timescale Lévy processes considered in the present paper. In [16] the above two-timescale random walk model is studied under certain scalings (comparable to the scalings with \( \varphi_n \) and \( \psi_n \) that are imposed in the present paper). The results in [16] include logarithmic asymptotics, which are for special cases refined to exact asymptotics in [15].

**Organization.** The fast regime (in which \( \varphi_n \) grows superlinearly) is covered by Section 2, followed by the slow regime (in which \( \varphi_n \) grows sublinearly) in Section 3. Examples are presented in Section 4. Section 5 describes how are findings can be applied when modeling overdispersed customers streams.

## 2. Fast regime

In this section we analyze the case that \( \varphi_n \) is superlinear, entailing that \( \psi_n \to 0 \) as \( n \to \infty \); this corresponds to \( f > 1 \) in the context of Example 1. The approach followed echoes the proof of e.g. [10, Thm. 3.7.4], and comprises the following elements:

- We first identify in Section 2.1 the ‘twisting factor’ \( \vartheta_n \) (i.e., the solution of (2)). It means that twisting \( C_n \) by \( \vartheta_n \) leads to a random variable with mean \( un \). More concretely, we define a measure \( \mathbb{Q}_n \) through (in self-evident notation)

\[
\mathbb{Q}_n(C_n \in dx) = \mathbb{P}(C_n \in dx) e^{\vartheta_n x} \log \gamma_n(\vartheta_n),
\]

where it holds that (in self-evident notation)

\[
\mathbb{E}_{\mathbb{Q}_n} C_n = \int_{-\infty}^{\infty} x \mathbb{Q}_n(C_n \in dx) = un.
\]


Then we rewrite in Section 2.2 the probability \( \xi_n(u) \) using the \( \vartheta_n \)-twisted version of \( C_n \). As in [1, Ch. XIII], we can express \( \xi_n(u) \) into a mean under \( \mathbb{Q}_n \):

\[
\xi_n(u) = \mathbb{P}(C_n \geq un) = \mathbb{E}_{\mathbb{Q}_n}(L(C_n)1\{C_n \geq un\}),
\]

for an appropriately chosen \( L(\cdot) \) (which can be interpreted as a likelihood ratio).

- The right-hand side of (4) turns out to consist of an exponential part and a polynomial part (in \( n \), that is). Section 2.3 analyzes the exponential part. As opposed to the standard Bahadur-Rao result [10, Thm. 3.7.4], the exponent potentially also contains sublinear terms.

- The polynomial part in the right-hand side of (4) is technically the most demanding element of the proof, and relies on Berry-Esséen-based arguments; see Section 2.4.

- In Section 2.5 all elements are put together, and our result is stated.

The formal assumption we impose on \( \psi_n \) is the following.

**Assumption 1.** The sequence \( \psi_n \) satisfies

\[
\limsup_{n \to \infty} \frac{\log \psi_n}{\log n} < 0.
\]

This assumption means that there is an \( \varepsilon > 0 \) such that \( \psi_n < n^{-\varepsilon} \), and hence \( \varphi_n > n^{1+\varepsilon} \). We observe that \( \varphi_n \) is superlinear.

In the main theorem of this section, i.e., Thm. 1, we assume that \( A(\cdot) \) is non-lattice. To establish the result for the case where \( A(\cdot) \) is lattice (covering e.g. Poisson processes), a minor adaptation needs to be made. More specifically, the steps of Sections 2.1 up to 2.3 apply to the lattice as well as non-lattice case, whereas in Section 2.4 in the lattice case slightly different bounds have to be used (fully analogously to the treatment of the lattice and non-lattice cases in the proof of [10, Thm. 3.7.4]). The result for the lattice case is discussed separately in Remark 2. We refer to Sections 4 and 5 for illustrative examples covering both the lattice and non-lattice case.

**2.1. Analysis of twisting factor.** As \( \psi_n \to 0 \) as \( n \to \infty \), the twisting factor \( \vartheta_n \) converges to the solution \( \vartheta^* \) of \( b\alpha'(\vartheta) = u \), where \( b = \mathbb{EB}(1) = \beta'(0) \). To establish the exact asymptotics of \( \xi_n(u) \), it turns out that we need an expansion of \( \vartheta_n \); we start by arguing that \( \vartheta_n \) has the form

\[
\vartheta_n = \vartheta^* + \sum_{k=1}^{\infty} v_k \psi_n^k,
\]

and then we point out how to determine the coefficients \( v_k \).

- The main idea is that one can (implicitly) define a function \( \vartheta_+(x) \) as the solution of the equation

\[
\beta'(\alpha(\vartheta)x)\alpha'(\vartheta) = u;
\]

such a solution is unique due to the fact that the left-hand side of the previous display is the derivative of a convex function (hence increasing). Now observe that \( \vartheta_n \) (solving \( \gamma_n'(\vartheta_n) = u \)) equals \( \vartheta_+(\psi_n) \); recall that \( \psi_n \to 0 \) as \( n \to \infty \).

Observe that \( \vartheta_+(0) = \vartheta^* \) (with \( \vartheta^* \), as defined above, solving \( b\alpha'(\vartheta^*) = u \)). Now write down the Taylor expansion (at \( x = 0 \)) of the implicitly defined function \( \vartheta_+(x) \):

\[
\vartheta_+(x) = \sum_{k=0}^{\infty} v_k x^k,
\]

with \( v_0 = \vartheta^* \). The \( v_k \) can be determined in terms of the derivatives of \( \alpha(\cdot) \) and \( \beta(\cdot) \) (which are well-defined). Below we provide a constructive procedure that identifies all these coefficients. Upon combining the above elements, we thus obtain the expansion (5):

\[
\vartheta_n = \vartheta_+(\psi_n) = \sum_{k=0}^{\infty} v_k \psi_n^k.
\]
We now present a constructive procedure that yields the coefficients $v_k$. To this end, first observe that, by evaluating $\alpha(\cdot)$ as a Taylor series around $\vartheta^*$,
\[
\beta' \left( \sum_{\ell=0}^{\infty} \frac{\alpha^{(\ell)}(\vartheta^*)}{\ell!} \left( \sum_{k=1}^{\infty} v_k \psi_n^k \right) \psi_n^{\ell} \right) \cdot \sum_{\ell=0}^{\infty} \frac{\alpha^{(\ell+1)}(\vartheta^*)}{\ell!} \left( \sum_{k=1}^{\infty} v_k \psi_n^k \right)^{\ell} = u,
\]
where, by evaluating $\beta(\cdot)$ as a Taylor series around 0,
\[
\beta'(\vartheta) = \sum_{m=0}^{\infty} \frac{\beta^{(m+1)}(0)}{m!} \vartheta^m.
\]
This equation effectively determines all $v_k$, by grouping together the appropriate terms (i.e., terms with the same power).

We present this procedure for $v_1$. Up to terms that are $o(\psi_n)$,
\[
(\beta'(0) + \beta''(0) \alpha(\vartheta^*) \psi_n) \left( \alpha'(\vartheta^*) + \alpha''(\vartheta^*) v_1 \psi_n \right) = u.
\]
Now $v_1$ can be determined: using that $\beta'(0) \alpha'(\vartheta^*) = b \alpha'(\vartheta^*) = u$ we obtain
\[
v_1 = -\frac{\alpha'(\vartheta^*)}{\alpha''(\vartheta^*)} \frac{\alpha'(\vartheta^*) \beta''(0)}{\beta'(0)}.
\]
The $v_i$ for $i \in \{2, 3 \ldots\}$ can be computed in the exact same way, but this does not lead to clean expressions; for $k = 2$ we obtain
\[
v_2 = -\frac{1}{2} \frac{\beta^{(1)}(0) \alpha^{(3)}(\vartheta^*) v_1^2 + \beta^{(2)}(0) (\alpha(\vartheta^*) \alpha(\vartheta^*) + (\alpha'(\vartheta^*))^2)}{\beta^{(1)}(0) \alpha^{(2)}(\vartheta^*)} v_1.
\]

2.2. Change of measure. The next step is to rewrite the probability of interest, relying on the usual change-of-measure procedure. This effectively means that we let $Q_n$ correspond to twisting the distribution of $C_n$ by the solution $\vartheta_n$ of (2). The l-mgf of $C_n$ under $Q_n$ can thus be expressed in terms of the l-mgf of $C_n$ under the original measure:
\[
\gamma_n^{Q_n}(\vartheta) := \gamma_n(\vartheta + \vartheta_n) - \gamma_n(\vartheta_n).
\]
Later on we need the mean and variance of $C_n$ under $Q_n$. From the definition of $\vartheta_n$, it follows that $E_{Q_n} C_n = un$. Differentiating once more, we obtain
\[
\text{Var}_{Q_n} C_n = n \psi_n \beta'(\alpha(\vartheta_n) \psi_n) (\alpha'(\vartheta_n))^2 + n \beta'(\alpha(\vartheta_n) \psi_n) \alpha''(\vartheta_n).
\]
The next step is to use the change of measure to rewrite $\xi_n(u)$. By applying the definition of $Q_n$, we find the identity
\[
\xi_n(u) = E_{Q_n} \left( e^{\gamma_n(\vartheta_n) - \vartheta_n C_n} 1\{C_n \geq un\} \right),
\]
realizing that $e^{\gamma_n(\vartheta_n) - \vartheta_n C_n}$ plays the role of the likelihood ratio $dP/dQ_n$.

On the event $C_n \geq un$, $C_n$ will typically be relatively close to $un$ under the new measure $Q_n$ (recall $E_{Q_n} C_n = un$). To exploit this, we define, with $\sigma^Q_{\vartheta_n} := \sqrt{b \alpha''(\vartheta^*)}$,
\[
\overline{D}_n := \frac{C_n - un}{\sqrt{n \sigma^Q_{\vartheta_n}}},
\]
which is a random variable that has, by construction, mean 0 under $Q_n$. (Due to the $\sqrt{n}$ in the denominator, it is anticipated that, under $Q_n$, $\overline{D}_n$ can be approximated by a standard Normal random variable; we come back to this idea below.) We thus obtain that, for all $n$,
\[
\xi_n(u) = e^{\gamma_n(\vartheta_n) - \vartheta_n un} \Delta_n, \text{ with } \Delta_n := E_{Q_n} \left( e^{-\vartheta_n \sigma^Q_{\vartheta_n} \sqrt{n} \overline{D}_n} 1\{\overline{D}_n \geq 0\} \right).
\]
Consequently, we are left with analyzing the exponential factor $\delta_n := \exp(\gamma_n(\vartheta_n) - \vartheta_n un)$ and the expectation $\Delta_n$, as $n$ grows large.
2.3. Analysis of $\delta_n$ as $n \to \infty$. We proceed by analyzing the exponent in the expansion (6), i.e., $\gamma_n(\vartheta_n) - \vartheta_n un$. Define

$$m_+ := \sup \{ k \in \mathbb{N} : \liminf_{n \to \infty} \varphi_n \psi_n^k > 0 \};$$

note that $m_+ \geq 1$. Our claim is the following.

Lemma 1. As $n \to \infty$, for constants $\overline{\nu}_k$,

$$\gamma_n(\vartheta_n) - \vartheta_n un = \left( b\alpha(\vartheta^*) - \vartheta^* u \right) n + \sum_{k=2}^{m_+} \overline{\nu}_k \varphi_n \psi_n^k + o(1),$$

where the empty sum is defined as 0.

The validity of this claim (as well as the procedure to determine the coefficients $\overline{\nu}_k$) can be demonstrated as follows. Relying on Taylor expansions, we obtain

$$\gamma_n(\vartheta_n) - \vartheta_n un = \varphi_n \beta \left( \sum_{\ell=0}^{\infty} \frac{\alpha^{(\ell)}(\vartheta^*)}{\ell!} \left( \sum_{k=1}^{\infty} v_k \psi_n^k \right)^\ell \psi_n^0 - \left( \vartheta^* + \sum_{k=1}^{\infty} v_k \psi_n^k \right) un. \right)$$

The claim for $m_+ = 1$ can be directly verified. Now consider the case $m_+ = 2$; any higher value of $m$ can be dealt with fully analogously. For $m_+ = 2$, the sequence $\varphi_n \psi_n^k$ converges to 0 when $k > 2$, whereas $\varphi_n \psi_n^2 = n \psi_n$ stays away from 0; think of e.g. $\psi_n = n^{-2/3}$. Hence we obtain that, as $n \to \infty$,

$$\gamma_n(\vartheta_n) - \vartheta_n un = (b\alpha(\vartheta^*) - \vartheta^* u) n + v_1 \left( \frac{\varphi''(0)}{2} \alpha(\vartheta^*) + b\alpha'(\vartheta^*) - u \right) \varphi_n \psi_n^2 + o(1).$$

It is clear that for $m_+ = 3$, an additional term needs to be included. In general, using this procedure any value of $m_+$ can be dealt with.

2.4. Analysis of $\Delta_n$ as $n \to \infty$. We are left with analyzing $\Delta_n$ for $n$ large. Our objective is to prove that $\sqrt{n}\Delta_n$ converges to a positive constant as $n \to \infty$. Mimicking the line of reasoning of [10, Eqn. (3.7.7)], we apply integration by parts:

$$\Delta_n = \int_0^\infty e^{-\vartheta_n \sigma_+^2 \sqrt{n} x} Q_n(\overline{D}_n \leq x) \, dx = \sqrt{n} \vartheta_n \sigma_+^2 \int_0^\infty e^{-\vartheta_n \sigma_+^2 \sqrt{n} x} \left( Q_n(\overline{D}_n \leq x) - Q_n(\overline{D}_n \leq 0) \right) \, dx \right) \, dx.$$

As we will intensively rely on this representation of $\Delta_n$, an important role in our argumentation is played by the probability distribution

$$Q_n(\overline{D}_n \leq x) = Q_n \left( \frac{A(\psi_n B(\varphi_n)) - un}{\sqrt{n} \sigma_+^2} \leq x \right),$$

where it is noted that $E_{Q_n} A(\psi_n B(\varphi_n)) = un$. Estimates for this distribution can be established, essentially as variants of the classical Edgeworth expansion. As such expansions follow by applying well-established techniques, we restrict ourselves to providing the main steps of the derivation in the appendix. An excellent introduction on the Edgeworth expansion is provided in [14, Ch. II]; see also [4, Ch. IV].

- In the case where $\lim_{n \to \infty} \psi_n \sqrt{n} = 0$ (which corresponds to $f > \frac{3}{2}$ in the context of Example 1), we have, as pointed out in the appendix, as $n \to \infty$,

$$\sqrt{n} \sup_x \left( Q_n(\overline{D}_n \leq x) - \Phi(x) + \phi(x) H_2(x) \frac{\kappa_+}{\sqrt{n}} \right) \to 0, \quad \kappa_+ := \frac{1}{6} \left( \frac{\varphi''(0)}{\sigma_+^4} \right)^2;$$

cf. Eqn. (30), where $H_2(x) = x^2 - 1$, $\phi(\cdot)$ denotes the probability density function of a standard Normal random variable, and $\Phi(\cdot)$ the corresponding cumulative distribution function. We are now
in a position to prove that $\sqrt{n}\Delta_n$ converges to a constant; we provide the upper bound, but the lower bound follows fully analogously. By virtue of (8), for any $\varepsilon > 0$ and $n$ sufficiently large,

$$\sqrt{n} \Delta_n \leq \sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \left( \Phi \left( \frac{x}{\sqrt{n}} \right) - \Phi(0) \right) dx -$$

$$\sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \phi \left( \frac{x}{\sqrt{n}} \right) \frac{x^2 \kappa_+}{\sqrt{n}} dx +$$

$$\varepsilon \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \left( \Phi \left( \frac{x}{\sqrt{n}} \right) - \Phi(0) \right) dx + \varepsilon \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} dx.$$

Let us start by evaluating the first term on the right-hand side. It can be rewritten as

$$\sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} dx dy$$

$$= \sqrt{n} \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} dx dy$$

$$= \sqrt{n} \exp \left( \frac{1}{2} \left( \vartheta_n \sigma^Q_+ \right)^2 \right) \left( 1 - \Phi(\vartheta_n \sigma^Q_+ / \sqrt{n}) \right).$$

Using the known limit $x(1 - \Phi(x))/\phi(x) \rightarrow 1$ as $x \rightarrow \infty$, and $\vartheta_n \rightarrow \vartheta^*$, we have that the expression in the previous display converges to

$$(\vartheta^* \sigma^Q_+ / \sqrt{2\pi})^{-1}. \quad (9)$$

Using that $H_2(x) = x^2 - 1$ the second term can be written as the sum of

$$t_{+1}^{(1)} := \sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \left( \phi(0) - \phi \left( \frac{x}{\sqrt{n}} \right) \right) \frac{\kappa_+}{\sqrt{n}} dx,$$

$$t_{+1}^{(2)} := \sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \phi \left( \frac{x}{\sqrt{n}} \right) \frac{x^2 \kappa_+}{\sqrt{n}} dx.$$

From (i) $\sqrt{n} (\phi(0) - \phi(x/\sqrt{n})) \rightarrow x \phi'(0) = 0$, (ii) $\vartheta_n \rightarrow \vartheta^*$ as $n \rightarrow \infty$, and (iii) the fact that $\phi'(\cdot)$ is bounded, applying the dominated convergence theorem yields that $t_{+1}^{(1)} \rightarrow 0$. Due to (i) $\phi(x/\sqrt{n}) \leq 1/\sqrt{2\pi}$, (ii) $\vartheta_n \rightarrow \vartheta^*$ as $n \rightarrow \infty$, and (iii)

$$\lim \sup_{n \rightarrow \infty} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x^2} dx = \lim \sup_{n \rightarrow \infty} \frac{2}{(\vartheta_n \sigma^Q_+)^2} < \infty,$$

we also have $t_{+1}^{(2)} \rightarrow 0$.

The third term equals $\varepsilon$, which can be made arbitrarily small. As mentioned before, by the same token it can be verified that the corresponding lower bound applies as well. We thus conclude that $\sqrt{n} \Delta_n$ converges to the constant in (9).

o We now proceed with the case $\lim \inf_{n \rightarrow \infty} \psi_n \sqrt{n} > 0$ (which simplifies to $f \in (1, \frac{3}{2})$ in the context of Example 1). From Eqn. (31) in the appendix we have

$$\sqrt{n} \sup_{x} \left( Q_n (D_n \leq x) - \Phi(x) + \phi(x) \left( H_1(x) \sum_{k=1}^{m} c_k \psi_n^k + H_2(x) \frac{\kappa_+}{\sqrt{n}} \right) \right) \rightarrow 0. \quad (10)$$

By (10), for any $\varepsilon > 0$ and $n$ sufficiently large, using that $H_1(x) = x$ and hence $H_1(0) = 0$,

$$\sqrt{n} \Delta_n \leq \sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \left( \Phi \left( \frac{x}{\sqrt{n}} \right) - \Phi(0) \right) dx -$$

$$\sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \phi \left( \frac{x}{\sqrt{n}} \right) \frac{x}{\sqrt{n}} \left( \sum_{k=1}^{m} c_k \psi_n^k \right) dx -$$

$$\sqrt{n} \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} \left( \Phi \left( \frac{x}{\sqrt{n}} \right) - \Phi(0) \right) H_2 \left( \frac{x}{\sqrt{n}} \right) dx + \varepsilon \vartheta_n \sum_{k=1}^{m} \chi_{k} \int_{0}^{\infty} e^{-\vartheta_n \sigma^Q_+ x} dx$$
The first, third, and fourth term can be dealt with as in the case \( \psi_n \sqrt{n} \to 0 \). Due to (i) \( \phi(x) \leq 1/\sqrt{2\pi} \) for all \( x \), (ii) \( \vartheta_n \to \vartheta^* \) as \( n \to \infty \), (iii) \( \psi_n^k \to 0 \) as \( n \to \infty \) for any \( k \in \{1, \ldots, m_+ \} \), and (iv)
\[
\limsup_{n \to \infty} \int_0^\infty e^{-\vartheta_n \sigma_n^2 x} x \, dx = \limsup_{n \to \infty} \frac{1}{(\vartheta_n \sigma_n^2)^2} < \infty,
\]
the second term vanishes. It follows that again \( \sqrt{n} \Delta_n \) converges to the constant in (9). We have thus established the following lemma.

**Lemma 2.** As \( n \to \infty \),
\[
\sqrt{n} \Delta_n \to (\vartheta^* \sigma_+^2 \sqrt{2\pi})^{-1}.
\]

**2.5. Result.** Upon combining Lemmas 1 and 2, as presented in the previous subsections, we have arrived at the following result.

**Theorem 1.** As \( n \to \infty \), under Assumption 1, for non-lattice \( A(\cdot) \),
\[
\frac{1}{\vartheta^* \sigma_+^2 \sqrt{2\pi n}} \exp \left( \left( b\alpha(\vartheta^*) - \vartheta^* u \right) n + \sum_{k=2}^{m_+} \bar{v}_k \varphi_n \psi_n^k \right).
\]

An immediate consequence of this theorem is that \( \xi_n(u) \) behaves as \( \mathbb{P}(A(bn) \geq un) \) when \( \varphi_n \psi_n^2 = n \psi_n \to 0 \). This is for instance the case when \( \psi_n = n^\zeta \) for \( \zeta < -1 \), or, in the setting of Example 1, \( f > 2 \). It reflects that the timescale of \( B(\cdot) \) is so much faster than that of \( A(\cdot) \), that it can be replaced by its mean. In addition, it implies that the rough (logarithmic) asymptotics are not affected by the choice of \( \psi_n \) (as long as Assumption 1 is fulfilled). These findings are summarized in the following corollary.

**Corollary 1.** If \( \varphi_n \psi_n^2 = n \psi_n \to 0 \) as \( n \to \infty \), then, under Assumption 1, for non-lattice \( A(\cdot) \),
\[
\frac{1}{\vartheta^* \sigma_+^2 \sqrt{2\pi n}} \exp \left( \left( b\alpha(\vartheta^*) - \vartheta^* u \right) n \right).
\]

As \( n \to \infty \), under Assumption 1, for non-lattice \( A(\cdot) \),
\[
\frac{1}{n} \log \xi_n(u) \to b\alpha(\vartheta^*) - \vartheta^* u.
\]

**Remark 2.** So far we throughout assumed that the Lévy processes \( A(\cdot) \) be non-lattice, thus ruling out e.g. Poisson processes. With a minor adaptation, however, the lattice case can be dealt with as well. Let, for some \( x_0 \) and \( d \) and any \( t \geq 0 \), the random variable \( d^{-1} (A(t) - x_0) \) be integer almost surely, and let \( d \) be the largest number with this property; e.g. in the Poisson process case \( d = 1 \). Then, following the proof in e.g. [10, Thm. II.7.4.b], under Assumption 1, we find the following counterpart of Thm. 1:
\[
\frac{d}{1 - e^{-\sigma_d}} \frac{1}{\sigma_+^2 \sqrt{2\pi n}} \exp \left( \left( b\alpha(\vartheta^*) - \vartheta^* u \right) n + \sum_{k=2}^{m_+} \bar{v}_k \varphi_n \psi_n^k \right),
\]
as \( n \to \infty \). This behavior is consistent with that of Thm. 1 when taking \( d \downarrow 0 \).

3. Slow regime

In this section we analyze the case that \( \varphi_n \) grows sublinearly, implying that \( \psi_n \to \infty \) as \( n \to \infty \) (also sublinearly). As before, we first characterize the solution \( \vartheta_n \) of (2), then we rewrite \( \xi_n(u) \) using the \( \vartheta_n \)-twisted version of \( C_n \), and finally we determine the corresponding exact asymptotics. The formal assumption we impose on \( \psi_n \) in this section is the following.

**Assumption 2.** The sequence \( \psi_n \) satisfies
\[
0 < \liminf_{n \to \infty} \frac{\log \psi_n}{\log n} \leq \limsup_{n \to \infty} \frac{\log \psi_n}{\log n} < 1.
\]
The first inequality of this assumption ensures that there is an \( \varepsilon \in (0,1) \) such that \( \psi_n > n^\varepsilon \), and hence \( \varphi_n < n^{1-\varepsilon} \), so that that \( \varphi_n \) is sublinear. In addition, the second inequality entails that \( \psi_n \) is sublinear, too. For the moment we assume that \( B(\cdot) \) be non-lattice; see Remark 3 for the corresponding result in the lattice case.

The procedure we follow to prove the exact asymptotics in the slow regime is in line with the one we developed for the fast regime: we perform a change of measure, take out the exponential factor \( \delta_n \), and analyze the remainder term \( \Delta_n \). As this procedure echoes the one developed in the previous section, we only include the main steps.

The change of measure is again based on the solution \( \vartheta_n \) that solves equation (2). In this case, as we argue below, \( \vartheta_n \) obeys the expansion
\[
\vartheta_n = \sum_{k=1}^{\infty} w_k \psi_n^{-k};
\]
the coefficients \( w_k \) can be recursively determined. The reasoning is analogous to the argumentation followed in the fast regime.

- To show the validity of the expansion (11), we write \( \vartheta_- (x) \) as the solution of the equation
  \[
  \beta' \left( \frac{\alpha(\vartheta)}{x} \right) \alpha'(\vartheta) = u,
  \]
so that \( \vartheta_n = \vartheta_-(1/\psi_n) \). Recall that in this regime \( \psi_n \to \infty \) as \( n \to \infty \).

- Observe that \( \vartheta_-(x)/x \to \tau^* \) as \( x \to 0 \), where \( \tau^* \) solves \( a \beta'(a \tau^*) = u \). We now write \( \vartheta_- (\cdot) \) as a Taylor expansion around \( x = 0 \):
  \[
  \vartheta_-(x) = \sum_{k=1}^{\infty} w_k x^k,
  \]
with \( w_1 = \tau^* \). As in the fast regime, the coefficients allow expressions in terms of the derivatives of \( \alpha(\cdot) \) and \( \beta(\cdot) \). Combining the above, we thus obtain the expansion (11):
  \[
  \vartheta_n = \vartheta_-(1/\psi_n) = \sum_{k=1}^{\infty} w_k \psi_n^{-k}.
  \]

- The procedure to identify the coefficients \( w_k \) in this slow regime is analogous to the procedure to identify the coefficients \( v_k \) in this fast regime. In particular, \( w_1 \) solves \( a \beta'(aw_1) = u \); in the sequel, we refer to \( w_1 \) by \( \tau^* \).

We define \( \sigma^Q_n := a \sqrt{\beta''(a \tau^*)} \); cf. the decomposition (1). For this regime we define
\[
\overline{E}_n := \frac{C_n - u n}{\psi_n \sqrt{\varphi_n \Delta_n}},
\]
which is a random variable that has, by construction, mean 0 and variance converging to 1 under \( Q_n \); comparing \( \overline{E}_n \) with \( D_n \), observe that there is now \( \psi_n \sqrt{\varphi_n} \) rather than \( \sqrt{n} \) in the denominator. Again we obtain, for all \( n \), the factorization
\[
\xi_n(u) = e^{\gamma_n(\vartheta_n) - \vartheta_n u n} \Delta_n, \quad \text{with } \Delta_n := E_{Q_n} \left( e^{-\vartheta_n \sigma^Q_n \psi_n \sqrt{\varphi_n \overline{E}_n} 1 \{ \overline{E}_n \geq 0 } \right).
\]
As before, it remains to analyze the exponential factor \( \delta_n := \exp(\gamma_n(\vartheta_n) - \vartheta_n u n) \) and the expectation \( \Delta_n \), in the regime of \( n \) growing large.

The analysis of \( \delta_n \) precisely follows the corresponding step in the fast regime. Define
\[
m_- := \sup \left\{ k \in \mathbb{N} : \liminf_{n \to \infty} \varphi_n \psi_n^{-k} > 0 \right\}.
\]
It thus follows that as \( n \to \infty \), for constants \( \overline{w}_k \), with the empty sum being defined as 0,
\[
\delta_n = \gamma_n(\vartheta_n) - \vartheta_n u n = (\beta(a \tau^*) - \tau^* u) \varphi_n + \sum_{k=1}^{m_-} \overline{w}_k \varphi_n \psi_n^{-k} + o(1).
\]
We continue with the analysis of $\Delta_n$. Analogously to our analysis in the fast regime, applying integration by parts, we write $\Delta_n$ as

\[
\Delta_n = \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} d\varphi_n Q_n(E_n \in \{C\})
\]

\[
= \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} (Q_n(E_n \leq x) - Q_n(E_n \leq 0)) dx
\]

\[
= \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} (Q_n(E_n \leq x/\sqrt{\varphi_n}) - Q_n(E_n \leq 0)) dx.
\]

We again proceed by using the Edgeworth expansion presented in the appendix.

- In the case $\lim_{n \to \infty} 3\varphi_n^2/n = 0$ (which corresponds to $f < \frac{2}{3}$ in the context of Example 1), we have, as pointed out in Eqn. (34) in the appendix, as $n \to \infty$,

\[
\sqrt{\varphi_n} \sup_x \left( Q_n(E_n \leq x) - \Phi(x) + \phi(x)H_2(x)\frac{\kappa_n}{\sqrt{\varphi_n}} \right) \to 0,
\]

\[
\kappa_n := \frac{1}{6}\frac{\beta''(a\tau^*)}{(\sigma^2)^3}.
\]

Our objective is to prove that $\sqrt{\varphi_n} \Delta_n$ converges to a constant; as in the fast regime, we provide the upper bound, but the lower bound follows fully analogously. By (13) and (14), for any $\varepsilon > 0$ and $n$ sufficiently large,

\[
\sqrt{\varphi_n} \Delta_n \leq \sqrt{\varphi_n} \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} \left( \Phi\left(\frac{x}{\sqrt{\varphi_n}}\right) - \Phi(0) \right) dx
\]

\[
- \sqrt{\varphi_n} \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} \left( \phi\left(\frac{x}{\sqrt{\varphi_n}}\right)H_2\left(\frac{x}{\sqrt{\varphi_n}}\right) - \phi(0)H_2(0) \right) \frac{\kappa_n}{\sqrt{\varphi_n}} dx + \varepsilon.
\]

The first term on the right-hand side equals

\[
\sqrt{\varphi_n} \exp \left( \frac{1}{2} (\psi_n^2)^2 \right) \left( 1 - \Phi(\psi_n^2) \right) ,
\]

which, using $x(1 - \Phi(x))/\phi(x) \to 1$ and $\psi_n \varphi_n \to \tau^*$, converges to

\[
(\tau^* \sigma^2 \sqrt{2\pi})^{-1}.
\]

As before, we split the second term into (recalling that $H_2(x) = x^2 - 1$)

\[
t_{(1),n}^{(1)} := \sqrt{\varphi_n} \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} \left( \phi(0) - \phi\left(\frac{x}{\sqrt{\varphi_n}}\right) \right) \frac{\kappa_n}{\sqrt{\varphi_n}} dx,
\]

\[
t_{(2),n}^{(2)} := \sqrt{\varphi_n} \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} \phi\left(\frac{x}{\sqrt{\varphi_n}}\right) \frac{x^2}{\sqrt{\varphi_n}} \frac{\kappa_n}{\sqrt{\varphi_n}} dx.
\]

Mimicking the reasoning used in the fast regime, it can be shown that $t_{(1),n}^{(1)} \to 0$ and $t_{(2),n}^{(2)} \to 0$ as $n \to \infty$. The third term equals $\varepsilon$, which can be made arbitrarily small. Combining this with the corresponding lower bound applies, we find that $\sqrt{\varphi_n} \Delta_n$ converges to (15).

- Considering $\liminf_{n \to \infty} 3\varphi_n^2/n > 0$ (which simplifies to $f \in [\frac{2}{3}, 1]$ in the context of Example 1), we obtain from Eqn. (35) in the appendix,

\[
\sqrt{\varphi_n} \sup_x \left( Q_n(E_n \leq x) - \Phi(x) + \phi(x)H_1(\sum_{k=1}^{k} c_k \psi_n^{-k} + H_2(\frac{\kappa_n}{\sqrt{\varphi_n}})) \right) \to 0.
\]

By (16), for any $\varepsilon > 0$ and $n$ sufficiently large, using that $H_1(x) = x$ and hence $H_1(0) = 0$,

\[
\sqrt{\varphi_n} \Delta_n \leq \sqrt{\varphi_n} \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} \left( \Phi\left(\frac{x}{\sqrt{\varphi_n}}\right) - \Phi(0) \right) dx
\]

\[
- \sqrt{\varphi_n} \psi_n \varphi_n^2 \int_0^\infty e^{-\psi_n^2 x \sqrt{\varphi_n}} \phi\left(\frac{x}{\sqrt{\varphi_n}}\right) \frac{x}{\sqrt{\varphi_n}} \left( \sum_{k=1}^{k} c_k \psi_n^{-k} \right) dx.
\]
The first, third, and fourth term can be dealt with as in the case $\varphi_n^{3/2}/n \to 0$. Analogously to the reasoning used in the fast regime, the second term vanishes. The corresponding lower bound is established in the same way. Conclude that also in this case $\sqrt{n} \Delta_n$ converges to (15). We have thus proven the following results.

**Theorem 2.** As $n \to \infty$, under Assumption 2, for non-lattice $B(\cdot)$,  
\[
\xi_n(u) \sim \frac{1}{\tau \sigma_n^2 \sqrt{2\pi \varphi_n}} \exp \left( (\beta(a\tau^*) - \tau^*u)\varphi_n + \sum_{k=1}^{m} w_k \varphi_n^{1/2} k \right).
\]

**Corollary 2.** If $\varphi_n^{1/2} = n/\varphi_n^2 \to 0$ as $n \to \infty$, then, under Assumption 2, for non-lattice $B(\cdot)$,  
\[
\xi_n(u) \sim \mathbb{P}(aB(\varphi_n) \geq u\varphi_n) \sim \frac{1}{\tau \sigma_n^2 \sqrt{2\pi \varphi_n}} \exp \left( (\beta(a\tau^*) - \tau^*u)\varphi_n \right).
\]

As $n \to \infty$, under Assumption 2, for non-lattice $B(\cdot)$,  
\[
\frac{1}{\varphi_n} \log \xi_n(u) \to \beta(a\tau^*) - \tau^*u.
\]

**Remark 3.** In the case $B(\cdot)$ is lattice, the result of Thm. 2 has to be slightly adjusted; cf. Remark 2. Let, for some $x_0$ and $d$ and any $t \geq 0$, the random variable $d^{-1}(B(t) - x_0)$ be integer almost surely, and let $d$ be the largest number with this property. Then, under Assumption 2, we obtain  
\[
\xi_n(u) \sim \frac{d}{1 - e^{-a\tau^* a}} \frac{1}{\sigma_n^2/a \sqrt{2\pi \varphi_n}} \exp \left( (\beta(a\tau^*) - \tau^*u)\varphi_n + \sum_{k=1}^{m} w_k \varphi_n^{1/2} k \right),
\]
as $n \to \infty$. Analogously to what we observed in the fast regime, this behavior is consistent with that of Thm. 2 when taking $d \downarrow 0$.  

**Remark 4.** In our analysis we have assumed that $a > 0$; one may wonder whether our findings extend to general $a$. When picking $a \leq 0$, however, in the slow regime complications arise. The constant $\tau^*$, that plays a crucial role in this regime, should solve the equation $u = \beta'(a\tau^*)a$, but observe that $\beta'(\cdot)$ is positive (as it the characteristic exponent of a subordinator); as a consequence, $u = \beta'(a\tau^*)a$ has no solution for $a \leq 0$. This fact implies that the case $a \leq 0$ should be dealt with an entirely different technique, in the sense that the change-of-measure technique (as was used above) cannot work.  

4. **Examples**

In this section we include two examples that demonstrate how the asymptotic expansion can be evaluated.

4.1. **$A(\cdot)$ is a Poisson process and $B(\cdot)$ a Gamma process.** Let $A(\cdot)$ be a Poisson process; we assume its rate is $\lambda > 0$, so that $\alpha(\vartheta) = \lambda(e^{\vartheta} - 1)$ for $\vartheta \in \mathbb{R}$. Let $B(\cdot)$ be a Gamma process; we call the parameters $r > 0$ (shape) and $\mu > 0$ (rate), so that $\beta(\vartheta) = r \log \mu - r \log (\mu - \vartheta)$, where we require $\vartheta < \mu$. Observe that $A(t)$ has a Poisson distribution with parameter $\lambda t$, and that $B(t)$ has a Gamma distribution with parameters $rt$ and $\mu$. In particular, in the terminology of our paper, $a = \lambda$ and $b = r/\mu$. To make the event of interest rare, we assume $u > ab$; writing $\vartheta := \lambda r/\mu$ this translates to assuming $\vartheta < 1$. Note that we have that $A(\cdot)$ is lattice but $B(\cdot)$ is not, so that we should apply Remark 2 in the fast regime, and Thm. 2 in the slow regime.

We start our computations by providing the l-mgf of $A(\psi_n B(\varphi_n))$ for this special case:  
\[
\gamma_n(\vartheta) = \varphi_n \beta(\alpha(\vartheta) \psi_n) = r \varphi_n \log \left( \frac{\mu}{\mu - \lambda(e^{\vartheta} - 1)\varphi_n} \right).
\]
As \( \vartheta_n \) satisfies the first-order condition \( \gamma_n(\vartheta_n) = un \), we are to solve

\[
\beta'(\alpha(\vartheta_n)\psi_n)\alpha'(\vartheta_n) = \frac{r \lambda e^{\vartheta_n}}{\mu - \lambda(e^{\vartheta_n} - 1)\psi_n} = u. \tag{18}
\]

We thus find

\[
\vartheta_n = \log \left( \frac{\mu u + \lambda \psi_n u}{\lambda r + \lambda \psi_n u} \right) = \log \left( 1 + \frac{\lambda}{\mu} \psi_n \right) - \log \left( \varrho (1 + \frac{u}{r} \psi_n) \right).
\]

We now distinguish between the fast regime and the slow regime. In the fast regime, in which \( \psi_n \to 0 \) as \( n \to \infty \), applying the Taylor expansion of the logarithm yields an expression for the coefficients \( v_k \). With \( \zeta_1 := \lambda/\mu \) and \( \zeta_2 := u/r \),

\[
\vartheta_n = \vartheta^* + \sum_{k=1}^{\infty} v_k \psi_n^k, \quad \vartheta^* = \log \frac{1}{\varrho}, \quad v_k := \frac{(-1)^{k+1}}{\varrho} \left( \zeta_1^k - \zeta_2^k \right).
\]

We now compute the coefficients \( \overline{v}_k \) featuring in Lemma 1. To this end, note that by inserting the first-order condition (18) into (17),

\[
\gamma_n(\vartheta_n) = -r \varphi_n \log(r e^{\vartheta_n}) = -r \varphi_n (\vartheta_n - \vartheta^*) = -r \sum_{k=1}^{\infty} v_k \varphi_n \psi_n^k,
\]

so that

\[
\delta_n = \gamma_n(\vartheta_n) - \vartheta_n un = -r \varphi_n \sum_{k=1}^{\infty} v_k \psi_n^k - \vartheta^* un - un \sum_{k=1}^{\infty} v_k \psi_n^k.
\]

It now follows that \( -r \vartheta_1 = (1 - \varrho) u = b \alpha' (\vartheta^*) \). We thus observe that \( \delta_n \) indeed has the form that was established in Lemma 1, with, for \( k \in \{2, 3, \ldots\} \), \( \overline{v}_k = -(r v_k + u v_{k-1}) \). More explicitly,

\[
\overline{v}_k = (-1)^k \left( \frac{r (\zeta_1^k - \zeta_2^k)}{k} - \frac{u (\zeta_1^{k-1} - \zeta_2^{k-1})}{k - 1} \right).
\]

Noting that \( (\sigma^Q_\tau)^2 = b \alpha''(\vartheta^*) = u \), and bearing in mind Remark 2, we conclude that

\[
\xi_n(u) \sim \frac{1}{1 - \varrho \sqrt{2\pi} un} \exp \left( (1 - \varrho + \log \varrho) un + \sum_{k=2}^{m_+} \overline{v}_k \varphi_n \psi_n^k \right), \tag{19}
\]

as \( n \to \infty \).

The computations pertaining to the slow regime work very similarly. The crucial step is that we now rewrite \( \vartheta_n \) by ‘Tayloring’ with respect to \( \psi_n^{-1} \) (rather than to \( \psi_n \)): with \( \zeta_1 := \mu/\lambda \) and \( \zeta_2 := r/u \),

\[
\vartheta_n = \log \left( 1 + \frac{\mu}{\lambda} \psi_n^{-1} \right) - \log \left( 1 + \frac{r}{u} \psi_n^{-1} \right) = \sum_{k=1}^{\infty} w_k \psi_n^{-k}, \quad w_k = \frac{(-1)^{k+1}}{k} \left( \zeta_1^k - \zeta_2^k \right).
\]

Then \( \tau^* = w_1 = \overline{\zeta}_1 - 1 = \overline{\zeta}_1 (1 - \varrho) = \frac{\lambda}{\mu} \varrho - 1 \) and \( \beta(a \tau^*) = r \log \frac{1}{\varrho} \). This gives

\[
\gamma_n(\vartheta_n) = -r \varphi_n \log(r e^{\vartheta_n}) = \beta(a \tau^*) \varphi_n - r \varphi_n \sum_{k=1}^{\infty} w_k \psi_n^{-k},
\]

so that

\[
\delta_n = \gamma_n(\vartheta_n) - \vartheta_n un = (\beta(a \tau^*) - \tau^* u) \varphi_n - \sum_{k=1}^{\infty} (r w_k + u w_{k+1}) \varphi_n \psi_n^{-k}.
\]

Defining \( (\sigma^O_\tau)^2 = a^2 \beta''(a \tau^*) = u^2/r \), application of Thm. 2 leads after minor calculations to

\[
\xi_n(u) \sim \frac{1}{\varrho - 1 \sqrt{2\pi} r \varphi_n} \exp \left( (1 - \frac{1}{\varrho} \log \frac{1}{\varrho}) \right) r \varphi_n + \sum_{k=1}^{m_-} \overline{w}_k \varphi_n \psi_n^{-k}, \tag{20}
\]

as \( n \to \infty \), where \( \overline{w}_k = -(r w_k + u w_{k+1}) \) for \( k \in \{1, 2, \ldots\} \); observe the similarity with (19).
Remark 5. It is noted that the expressions encountered in this example align with those featuring in [15, Section 3] that deal with a case in which $\xi_n(u)$ could be evaluated explicitly. It can be checked that $C_n$ has a negative binomial distribution, as follows. Observe that

$$
\gamma_n(\vartheta) = r\varphi_n \log \left( \frac{\mu}{\mu - \lambda (e^\vartheta - 1)\psi_n} \right).
$$

In case a random variable has a negative binomial distribution with parameters $k$ and $p$, then its l-mgf is of the form

$$
k \log \left( \frac{p}{1 - (1-p)e^\vartheta} \right).
$$

It thus follows that in the setting considered in this subsection, we have that $C_n$ is negative binomial, with success probability equal to $p = \mu/(\mu + \lambda \psi_n)$ and the allowed number of failures $k = r\varphi_n$. Further intuition behind the appearance of the negative binomial distribution has been provided in [15, Remark 6].

4.2. $A(\cdot)$ is a Gamma process and $B(\cdot)$ a Poisson process. In our second example the Gamma process and the Poisson process swap roles. In other words, we have $\alpha(\vartheta) = r \log \mu - r \log(\mu - \vartheta)$ for $\vartheta < \mu$, and $\beta(\vartheta) = \lambda (e^\vartheta - 1)$ for $\vartheta \in \mathbb{R}$. Again, to make the event of interest rare, we assume that $u > ab$, or alternatively, $\vartheta := \lambda r/(\mu u) < 1$ (where $a = r/\mu$ and $b = \lambda$). In this case $A(\cdot)$ is non-lattice, whereas $B(\cdot)$ is, so that we have to use Thm. 1 in the fast regime, and Remark 3 in the slow regime.

Remark 6. Note that the random variable $A(\psi_n B(\varphi_n))$ is compound Poisson in this case, with Gamma distributed jumps. More precisely, the jumps are generated according to a Poisson process with rate $\varphi_n \lambda$, where the jumps are Gamma, with parameters $r\psi_n$ and $\mu$.

In this case the l-mgf is given by

$$
\gamma_n(\vartheta) = \varphi_n \lambda \left( \left( \frac{\mu}{\mu - \vartheta} \right)^{r\psi_n} - 1 \right).
$$

(21)

Again we have to distinguish between the fast regime and the slow regime. As before, we start with the fast regime. From (21) it follows, by solving the first-order condition, that

$$
\vartheta_n = \mu (1 - \vartheta_n^m), \quad \eta_n := \frac{1}{1 + r\psi_n},
$$

so that

$$
\gamma_n(\vartheta_n) = \varphi_n \lambda \left( \frac{1}{\vartheta} \cdot \vartheta_n^m - 1 \right).
$$

To find the $v_k$, we write $\vartheta_n$ as a Taylor series in $\psi_n$. It is directly seen that $\vartheta^* = \mu(1 - \vartheta)$; some elementary calculus thus yields for the first coefficients $v_1 = -\mu r \vartheta \log \frac{1}{\vartheta}$, and $v_2 = \mu r^2 \vartheta (\log \frac{1}{\vartheta}) (1 - \frac{1}{2} \log \frac{1}{\vartheta})$.

higher coefficients can be found in the same way.

Also the coefficients $\pi_k$ can be found:

$$
\delta_n = \gamma_n(\vartheta_n) - \vartheta_n^m u n = \varphi_n \lambda \left( \frac{1}{\vartheta} \cdot \vartheta_n^m - 1 \right) - (1 - \vartheta_n^m) u n
$$

$$
= \frac{\lambda}{\vartheta} (\vartheta_n^m - 1 + 1 - \vartheta) \varphi_n - \mu(1 - \vartheta_n^m) u n = \frac{\lambda}{\vartheta} (1 - \vartheta) \varphi_n - \mu u \left( \frac{1}{r} \varphi_n + n \right) (1 - \vartheta_n^m)
$$

$$
= \frac{\lambda}{\vartheta} (1 - \vartheta) \varphi_n - u \left( \frac{1}{r} \varphi_n + n \right) \left( \vartheta^* + \sum_{k=1}^{\infty} v_k \psi_n^k \right)
$$

$$
= \left( \frac{\lambda}{\vartheta} - \frac{\mu u}{r} \right) (1 - \vartheta) \varphi_n - (1 - \vartheta) u m n - \frac{u}{r} v_1 n - u \sum_{k=2}^{\infty} \left( \frac{1}{r} \psi_k + v_{k-1} \varphi_n \psi_n^k \right)
\[
\xi_n(u) \sim \frac{1}{\sqrt{2\pi \lambda r n}} \exp \left( \left( 1 - \frac{1}{\theta} + \log \frac{1}{\theta} \right) \lambda r n + \sum_{k=2}^{\infty} \left( \frac{1}{r} v_k + v_{k-1} \right) \varphi_n \psi_n^{-k} \right).
\]

so that \( \tau_k = -u \left( \frac{1}{r} v_k + v_{k-1} \right) \). Note that the first term equals \((b_\theta(\vartheta^*) - \vartheta^*u)n\), in accordance with the result in Lemma 1. Thm. 1, with \((\sigma_\theta^2)^2 = u^2/\lambda r\), yields

\[
\xi_n(u) \sim \frac{1}{\sqrt{2\pi \lambda r n}} \exp \left( \left( 1 - \frac{1}{\theta} + \log \frac{1}{\theta} \right) \lambda r n + \sum_{k=2}^{m_n} \vartheta_k \varphi_n \psi_n^{-k} \right).
\]

We now proceed with the slow case. The \( w_k \) follow by expanding \( \vartheta_n^{-1} \) (recalling that \( \psi_n \to \infty \):

\[
\vartheta_n = \mu \left( 1 - \varrho \cdot \varrho_n \right), \quad \vartheta_n := -\frac{r}{r + 1/\psi_n}.
\]

Routine calculations show that \( \tau^* = w_1 = (\mu/r) \log \frac{1}{\theta} \), and

\[
w_2 = -\frac{\mu}{r} \left( \log \frac{1}{\theta} \right) (1 + \frac{1}{2} \log \frac{1}{\theta}),
\]

where higher coefficients follow along the same lines. The \( \vartheta_k \) can be determined as before; leaving out a few intermediate steps,

\[
\delta_n = \gamma_n(\vartheta_n) - \vartheta_n un = \varphi_n \lambda (\varrho_n - 1) - \mu (1 - \varrho \cdot \varrho_n) un
\]

\[
= (1 - \varrho + \log \varrho) \frac{\lambda}{\vartheta_n} \varphi_n - u \sum_{k=1}^{\infty} \left( \frac{1}{r} w_k + w_{k+1} \right) \varphi_n \psi_n^{-k}.
\]

By applying Remark 3, with \((\sigma_\vartheta^2)^2 = (r/\mu)^2 \cdot \lambda/\varrho = ru/\mu \) and \( \vartheta_k = -u \left( \frac{1}{r} w_k + w_{k+1} \right) \),

\[
\xi_n(u) \sim \frac{1}{\sqrt{2\pi \lambda r n}} \exp \left( \left( 1 - \varrho + \log \varrho \right) \frac{\lambda}{\vartheta_n} \varphi_n + \sum_{k=1}^{m_n} \bar{w}_k \varphi_n \psi_n^{-k} \right).
\]

5. Applications in modeling overdispersed customer streams

In operations research, the performance evaluation and design of service systems is a key topic of interest. Traditionally, the dominant assumption when modeling customer arrival processes is that of Poisson arrivals. As we mentioned in the introduction, however, measurements indicate that the level of variation observed in practice may significantly exceed that predicted by the Poisson process [5, 18, 19, 21]. More concretely, where for Poisson arrivals the mean and variance of the number of arrivals in a given time interval coincide, in practice one typically observes that the variance is larger than the mean; the phenomenon is typically referred to as overdispersion.

In [16] a mechanism is proposed that produces overdispersion. The main idea is that the arrival rate is random rather than deterministic; this is achieved by resampling the arrival rate periodically. The main underlying idea behind the resampling is that there is a (perhaps unobservable) environmental process that influences all potential customers in the same way; think for instance of the weather conditions, or the occurrence of specific events. As we argued in the introduction, our two-timescale framework can be used to represent this ‘resampled Poisson process’.

Renormalizing time such that the resampling intervals have length 1, considering \( K \) such intervals, and letting the samples be i.i.d. non-negative random variable \( \Lambda_1, \ldots, \Lambda_K \), then the number of arrivals in there \( K \) intervals is distributed as

\[
C(K) = \text{Pois} \left( \sum_{i=1}^{K} \Lambda_i \right).
\]

When designing service centers (for instance when making staffing decisions), one would like to control the probability of excessive waiting times. For that reason, it could be informational to
have a handle on the tail distribution corresponding to the number of arrivals in a certain time window.

The remainder of this section has two objectives. In the first place, we quantify the error made by neglecting overdispersion (i.e., the situation in which the number of arrivals in the $K$ slots has a Poisson distribution with parameter $K \Lambda$). In the second place, we compare our refined asymptotics with the crude asymptotics that were found in [16], just involving the dominant term in the exponent (as featuring in the second statements of Corollaries 1 and 2).

5.1. Setup. In our experiments, the following example is studied. Let $\Lambda_1, \ldots, \Lambda_K$ be i.i.d. samples from an exponential distribution with mean $1/\mu$. We are interested in using our expansions to approximate the probability

$$\Pi(K, \overline{\mu}) := \mathbb{P}(C(K) \geq u),$$

for some $\overline{\mu} > K/\mu$.

- Importantly, as we have seen in Remark 5, this example allows an explicit solution: here $C(K)$ has a negative binomial distribution with parameters $K$ and success probability $\overline{\mu}/(1 + \overline{\mu})$. We thus have, in self-evident notation,

$$\Pi(K, \overline{\mu}, \mu) = \mathbb{P}(\text{NegBin}(K, \frac{\mu}{1 + \mu}) \geq u).$$

We chose the example of exponentially distributed $\Lambda_i$, as the explicit expressions allow us to compare the various approximations with the corresponding exact values.

- When neglecting the overdispersion, one does not work with sampled values $\Lambda_1, \ldots, \Lambda_K$, but rather with a deterministic arrival rate $\mathbb{E}\Lambda$ in any interval. In other words, when one would ignore the overdispersion, the probability $\Pi(K, \overline{\mu}, \mu)$ would be approximated by

$$\Pi^{(\text{Pois})}(K, \overline{\mu}) := \mathbb{P}(\text{Pois}(K \mathbb{E}\Lambda) \geq \overline{\mu}).$$

This approximation, corresponding with the fast regime, is expected to work well if $K$ is large relative to $\overline{\mu}$, and the contributions of the individual $\Lambda_i$ are relatively small (i.e., $\overline{\mu}$ is relatively large).

- Along the same lines, one could consider the regime that $K$ is substantial, but not in the order of $\overline{\mu}$, and the contributions of the individual $\Lambda_i$ are relatively large (i.e., $\overline{\mu}$ is relatively small). Now the slow regime is expected to apply, suggesting to ignore the Poisson sampling. Bearing in mind that the sum of exponentially distributed random variables (with equal parameter) has an Erlang distribution, we end up with (in self-evident notation)

$$\Pi^{(\text{Gamma})}(K, \overline{\mu}) := \mathbb{P}(\text{Erl}(K, \mu) \geq \overline{\mu}).$$

The crude approximations $\Pi^{(\text{Pois})}(K, \overline{\mu}, \mu)$ and $\Pi^{(\text{Gamma})}(K, \overline{\mu})$ we wish to compare with approximations based on our two-timescale model. To this end, we convert the asymptotics that we found in Section 4 into approximations. We fix a number $n > 1$; this scaling parameter can be considered as ‘artificial’ in the sense that the choice of $n$ will not affect the approximation (as will turn out below). Then we pick $f$ such that $nf = K$ (i.e., $f = \log K/\log n > 0$), and we set $u := \overline{\mu}/n$ and $\mu := \overline{\mu}n^{1-f}$. We thus obtain

$$\sum_{i=1}^{K} \Lambda_i \overset{d}{=} n^{1-f} \sum_{i=1}^{nf} \Lambda_i^o,$$

where $\Lambda_i^o$ is exponentially distributed with parameter $\mu$. We have thus rewritten the probability (22) as

$$\mathbb{P}\left(\text{Pois}\left(n^{1-f} \sum_{i=1}^{nf} \Lambda_i^o\right) \geq un\right).$$
Now notice that this representation falls in the framework of Section 4.1, with $\varphi_n = n^f$. More specifically, we are in the situation that $A(\cdot)$ is a Poisson process with rate $\lambda = 1$, and $B(\cdot)$ a Gamma process with shape parameter $r = 1$ and rate parameter $\mu$.

The next step is to translate the asymptotics that we identified for $\xi_n(u)$, which are in terms of the parameters of the scaled model (i.e., $\mu$, $f$, and $u$), into approximations for $\Pi(K, \overline{\mu}, \overline{\mu})$, which are in terms of the original parameters (i.e., $\overline{\mu}$, $K$, and $\overline{\mu}$).

5.2. Fast regime. We start by considering the regime in which $K$ is large relative to $\overline{\mu}$, so that it makes sense to work with the asymptotics pertaining to the case $f > 1$ (i.e., the fast regime). First observe that

$$\varphi_n = \frac{n}{\mu u} = \frac{K}{\overline{\mu} \overline{u}},$$

which we assumed to be smaller than 1. In addition,

$$\zeta_1 = \frac{1}{\overline{\mu} n^{1-f}} = \frac{1}{\overline{\mu} \psi_n}, \quad \zeta_2 = \frac{\overline{\mu}}{n} = \frac{\overline{\mu}}{K \psi_n},$$

so that (after some elementary computations)

$$\overline{v}_k \varphi_n \psi_n^k = (-1)^k \left( \frac{K t_k^+}{k} - \frac{\overline{\mu} t_k^+}{k-1} \right), \quad t_k^+ := \left( \frac{1}{\overline{\mu}} \right)^k - \left( \frac{\overline{\mu}}{K} \right)^k.$$ (24)

Also,

$$\varphi^* = \log \frac{\overline{\mu} \overline{u}}{K}.$$ Based on (19), we eventually obtain the following approximation for the probability (22):

$$\Pi^{(\text{fast})}(K, \overline{\mu}, \overline{\mu}) := \frac{1}{1 - \varphi} \frac{1}{\sqrt{2\pi \overline{u}}} \exp \left( (1 - \varphi + \log \varphi) \overline{u} + \sum_{k=2}^{M_+} (-1)^k \left( \frac{K t_k^+}{k} - \frac{\overline{\mu} t_k^+}{k-1} \right) \right),$$ (25)

with $\varphi$ given by (23) and $t_k^+$ by (24). In practice one should choose the threshold $M_+$ such that adding more terms in the sum does not affect the outcome; realize that the object $m_+$ is not well-defined in the pre-limit context. Observe that, as announced, the scaling parameter $n$ does not appear in the approximation.

In [16] logarithmic asymptotics of $\xi_n(u)$ were found. These lead to (crude) approximations that only cover the dominant term in the exponent (as in the second statement of Corollary 1). More concretely, one would approximate the probability of our interest by

$$\hat{\Pi}^{(\text{fast})}(K, \overline{\mu}, \overline{\mu}) = \exp ((1 - \varphi + \log \varphi) \overline{u}).$$

5.3. Slow regime. Now consider the regime in which $K$ is small relative to $\overline{\mu}$, in which we opt for relying on the asymptotics corresponding to $f < 1$. Then

$$\overline{v}_1 = \overline{\mu} n^{1-f} = \overline{\mu} \psi_n, \quad \overline{v}_2 = \frac{n}{\overline{\mu}} = \frac{K \psi_n}{\overline{\mu}},$$

implying that

$$\overline{v}_k \varphi_n \psi_n^{-k} = (-1)^k \left( \frac{K t_k^-}{k} - \frac{\overline{\mu} t_k^-}{k+1} \right), \quad t_k^- := \overline{\mu}^k - \left( \frac{\overline{\mu}}{K} \right)^k.$$ (26)

In addition,

$$\tau^* = K^{1/f} \left( \frac{\overline{\mu}}{K} - \frac{1}{\overline{\mu}} \right).$$ Based on (20), we thus obtain the following approximation for (22):

$$\Pi^{(\text{slow})}(K, \overline{\mu}, \overline{\mu}) := \frac{1}{1 - \varphi} \frac{1}{\sqrt{2\pi K}} \exp \left( (1 - \frac{1}{\varphi} + \log \frac{1}{\varphi}) K + \sum_{k=1}^{M_-} (-1)^k \left( \frac{K t_k^-}{k} - \frac{\overline{\mu} t_k^-}{k+1} \right) \right),$$ (27)
with \( \varrho \) given by (23) and \( t_k^- \) by (26). As before, the truncation level \( M_- \) is chosen such that including additional terms would not have any impact on the approximation. Again, we observe that the scaling parameter \( n \) does not appear in the approximation.

We also compare with the approximation that only covers the dominant term in the exponent (as in the second statement of Corollary 2, in line with the decay rate identified in [16]). This approximation reads

\[
\hat{\Pi}^{(\text{slow})}(K, \overline{\varpi}, \overline{\mu}) = \exp \left( \left( 1 - \frac{1}{\varrho} + \log \frac{1}{\varrho} \right) K \right).
\]

5.4. Numerical experiments. We conclude this section by presenting a set of examples that illustrate the use of the developed approximations. We subsequently consider an example that corresponds to the fast regime, and one that corresponds to the slow regime. They provide indications of the accuracy that can be achieved by the expressions (25) and (27) that were based on our refined asymptotics, relative to more crude approximations.

We start with an example corresponding to the fast regime. Recalling (25), notice that our approximation \( \Pi^{(\text{fast})}(K, \overline{\varpi}, \overline{\mu}) \) is parameterized by the threshold \( M_+ \). In Table 1 below, the column with superscript 0 corresponds to neglecting the sum in the exponent of (25) altogether, which could be seen as choosing \( M_+ := 1 \), whereas the column with superscript 1 corresponds to including one term in the sum in the exponent of (25), i.e., \( M_+ := 2 \). Throughout this example \( \varrho = \frac{1}{3} \) and \( \overline{\varpi} = 150 \) are held fixed, explaining that the values in the third, fourth and fifth column are constant (as can be seen from the expressions for the approximations \( \Pi^{(\text{Pois})}, \hat{\Pi}^{(\text{fast})}, \Pi^{(\text{fast},1)} \)).

| \((K, \overline{\mu}, \overline{\mu})\) | \(\Pi\) | \(\Pi^{(\text{Pois})}\) | \(\Pi^{(\text{fast})}\) | \(\Pi^{(\text{fast},0)}\) | \(\Pi^{(\text{fast},1)}\) |
|-----------------|--------|--------|--------|--------|--------|
| \((100 000, 1000, 150)\) | 1.90 \(\cdot\) 10\(^{-6}\) | 1.88 \(\cdot\) 10\(^{-6}\) | 2.00 \(\cdot\) 10\(^{-5}\) | 1.95 \(\cdot\) 10\(^{-6}\) | 1.97 \(\cdot\) 10\(^{-6}\) |
| \((50 000, 500, 150)\) | 1.93 \(\cdot\) 10\(^{-6}\) | 1.88 \(\cdot\) 10\(^{-6}\) | 2.00 \(\cdot\) 10\(^{-5}\) | 1.95 \(\cdot\) 10\(^{-6}\) | 2.00 \(\cdot\) 10\(^{-6}\) |
| \((10 000, 100, 150)\) | 2.13 \(\cdot\) 10\(^{-6}\) | 1.88 \(\cdot\) 10\(^{-6}\) | 2.00 \(\cdot\) 10\(^{-5}\) | 1.95 \(\cdot\) 10\(^{-6}\) | 2.21 \(\cdot\) 10\(^{-6}\) |
| \((5 000, 50, 150)\) | 2.41 \(\cdot\) 10\(^{-6}\) | 1.88 \(\cdot\) 10\(^{-6}\) | 2.00 \(\cdot\) 10\(^{-5}\) | 1.95 \(\cdot\) 10\(^{-6}\) | 2.51 \(\cdot\) 10\(^{-6}\) |
| \((1 000, 10, 150)\) | 5.89 \(\cdot\) 10\(^{-6}\) | 1.88 \(\cdot\) 10\(^{-6}\) | 2.00 \(\cdot\) 10\(^{-5}\) | 1.95 \(\cdot\) 10\(^{-6}\) | 6.82 \(\cdot\) 10\(^{-6}\) |

**Table 1.** Approximations for fast regime.

The following observations can be made from Table 1. (i) In the two top rows there is so much timescale separation that \( \Pi^{(\text{fast},0)} \) provides highly accurate results. In the other rows the timescale separation becomes less pronounced, which leads to \( \Pi^{(\text{fast},1)} \) becoming the preferred approximation. Here it is noted that including 2 or more terms in the exponent hardly improves the approximation in the third and fourth row. In the last row the timescales are so poorly separated that \( \Pi^{(\text{fast},1)} \) is still relatively far off; we mention that the performance is significantly improved by using \( M_+ = 3 \), which leads to the highly accurate approximation \( 5.99 \cdot 10^{-6} \). (ii) As anticipated, the Poisson approximation (neglecting the overdispersion) performs well if there is a substantial degree of timescale separation, but leads to substantial underestimation if the timescales are relatively close together.

In the slow regime our approximation \( \Pi^{(\text{fast})}(K, \overline{\varpi}, \overline{\mu}) \), as given by (27), is parameterized by the threshold \( M_- \). Analogously to the notation used in Table 1, in Table 2 the column with superscript 0 corresponds to neglecting the sum in the exponent of (27), which one could identify with the choice \( M_- := 0 \), whereas the column with superscript 1 corresponds to including one term in the sum in the exponent of (27), i.e., \( M_- := 1 \). We chose scenarios in which \( \varrho = \frac{2}{3} \) and \( K = 100 \) are held fixed, implying that the values in the fourth and fifth column are constant (as can be seen from the expressions for \( \hat{\Pi}^{(\text{fast})} \) and \( \Pi^{(\text{fast},0)} \)); note that in this slow regime the values in the third column (i.e., \( \Pi^{(\text{Gamma})} \)) are not constant, in that they (very slowly) increase.
EXACT ASYMPTOTICS FOR A MULTI-TIMESCALE MODEL

The conclusions from Table 2 are as follows. (i) Again, in the top rows (with a strong timescale separation) \( \Pi^{(\text{fast,0})} \) performs well. This approximation, however, degrades in the lower rows, where the accuracy substantially improves if \( M_\infty = 1 \) (i.e., the approximation \( \Pi^{(\text{fast,1})} \)) is used. We mention that for the parameters in the last row, the accuracy is further improved by taking \( \Pi^{(\text{fast,0})} \) as an infinite-server queue with overdispersed input. Using the insights from \[ (100, 0.001, 150000) \]
\[ (100, 0.05, 30000) \]
\[ (100, 0.01, 15000) \]
\[ (100, 0.05, 3000) \]
\[ (100, 0.1, 1500) \]

| \( (K, \pi, \bar{\pi}) \) | \( \Pi \) | \( \Pi^{(\text{Gamma})} \) | \( \Pi^{(\text{fast})} \) | \( \Pi^{(\text{fast,0})} \) | \( \Pi^{(\text{fast,1})} \) |
|-----------------|-----|------------|------------|----------------|----------|
| \( (100, 0.001, 150000) \) | 5.98 \( \cdot 10^{-6} \) | 5.93 \( \cdot 10^{-6} \) | 7.84 \( \cdot 10^{-5} \) | 6.26 \( \cdot 10^{-6} \) | 6.31 \( \cdot 10^{-6} \) |
| \( (100, 0.05, 30000) \) | 6.20 \( \cdot 10^{-6} \) | 5.93 \( \cdot 10^{-6} \) | 7.84 \( \cdot 10^{-5} \) | 6.26 \( \cdot 10^{-6} \) | 6.52 \( \cdot 10^{-6} \) |
| \( (100, 0.01, 15000) \) | 6.48 \( \cdot 10^{-6} \) | 5.95 \( \cdot 10^{-6} \) | 7.84 \( \cdot 10^{-5} \) | 6.26 \( \cdot 10^{-6} \) | 6.80 \( \cdot 10^{-6} \) |
| \( (100, 0.05, 3000) \) | 9.11 \( \cdot 10^{-6} \) | 6.03 \( \cdot 10^{-6} \) | 7.84 \( \cdot 10^{-5} \) | 6.26 \( \cdot 10^{-6} \) | 9.49 \( \cdot 10^{-6} \) |
| \( (100, 0.1, 1500) \) | 1.35 \( \cdot 10^{-5} \) | 6.14 \( \cdot 10^{-6} \) | 7.84 \( \cdot 10^{-5} \) | 6.26 \( \cdot 10^{-6} \) | 1.44 \( \cdot 10^{-5} \) |

Table 2. Approximations for slow regime.

6. Discussion and concluding remarks

Motivated by recent developments in arrival process modeling, this paper presents tail asymptotics for a multi-timescale model. The focus has been on approximating \( \xi_n(u) = \mathbb{P}(A(\varphi_n)B(\varphi_n) > un) \), with \( A(\cdot) \) a Lévy process and \( B(\cdot) \) a Lévy subordinator. Our analysis shows that subtle analysis is required to identify the exact asymptotics. The structure found is considerably richer than in the model’s classical (single timescale) counterpart; in particular, the exponent includes additional terms.

Effect of leaving customers. In the case \( A(\cdot) \) is a Poisson process, we recover the mixed Poisson model proposed in [16], which can be thought of as a Poisson process with periodically resampled rate (so as to generate overdispersion). An interesting extension of the results developed in the presented paper could concern the model in which the arrived clients leave after a service time (so as to generate overdispersion). An interesting extension of the results developed in the presented paper could concern the model in which the arrived clients leave after a service time distributed as the random variable \( J \) with \( F_J(t) := \mathbb{P}(J \leq t) \). Such a setting could be considered as an infinite-server queue with overdispersed input. Using the insights from [16], one thus obtains the following expression for the number of customers \( \overline{C}_n \) present at time \( \varphi_n \), under the scaling considered in the present paper:

\[
\overline{C}_n = A\left( \psi_n \int_0^{\varphi_n} (1 - F_J(t)) dB(t) \right).
\]

Using calculation rules for Lévy processes, one obtains

\[
\log \mathbb{E} e^{\theta \overline{C}_n} = \int_0^{\varphi_n} \beta(\alpha(\varphi)\psi_n(1 - F_J(t))) dt;
\]

plugging in \( F_J(t) = 0 \) for all \( t \in [0, \varphi_n] \), we recover \( \varphi_n \beta(\alpha(\varphi)\psi_n) \) (which makes sense, as for this choice customers do not leave the system). As a topic for further research, one could pursue deriving the exact asymptotics of \( \mathbb{P}(\overline{C}_n \geq un) \) (which may require scaling the service times \( J \)). These asymptotics can be converted into approximations, to be used as the basis of refined staffing rules.

Follow-up research. Other directions for future work include:

- In the first place, multivariate extensions could be considered. One could for instance consider the probability that the vector

\[
(A_1(\psi_n B(\varphi_n)), A_2(\psi_n B(\varphi_n))),
\]

- In the second place, multivariate extensions could be considered. One could for instance consider the probability that the vector

\[
(A_1(\psi_n B(\varphi_n)), A_2(\psi_n B(\varphi_n))),
\]

- In the third place, multivariate extensions could be considered. One could for instance consider the probability that the vector

\[
(A_1(\psi_n B(\varphi_n)), A_2(\psi_n B(\varphi_n))),
\]
with \( \mathbf{A}(\cdot) = (A_1(\cdot), A_2(\cdot)) \) a bivariate Lévy processes, attains a value in the set \([u_1 n, \infty) \times [u_2 n, \infty)\). The components of \( \mathbf{A}(\cdot) \) could be assumed dependent, but observe that \( A_1(\psi_n B(\varphi_n)) \) and \( A_2(\psi_n B(\varphi_n)) \) are dependent anyway (as the same \( B(\varphi_n) \) is used). The multivariate single-timescale results of [7] show that one should expect that various cases arise: there will be cases in which one component exceeding its threshold (with high probability) implies that the other component exceeds its threshold as well, but also cases in which both constraints are ‘tight’.

- In Remark 4 we pointed out that in our setup it is required to assume \( a > 0 \), particularly in the slow regime. For \( a \leq 0 \), one may anticipate that \( \xi_n(u) \) decays exponentially in \( n \) (unlike in the case of \( a > 0 \), where \( \xi_n(u) \) decays exponentially in \( \varphi_n \)), based on the following reasoning. The starting point is the large-deviations heuristic

\[
\xi_n(u) \approx \max_{x > 0} \mathbb{P}(B(\varphi_n) \approx x \varphi_n) \mathbb{P}(A(xn) \approx un). \tag{28}
\]

The main insight is that, taking into account only the leading term of the asymptotics, the first probability on the right-hand side of (28) decays exponentially in \( \varphi_n \), but the second exponentially in \( n \), so that their product decays exponentially in \( n \). This should be contrasted with the case \( a > 0 \) studied in the present paper: there the maximum will be attained (roughly) at \( x = u/a > 0 \), so that \( \{A(xn) \approx un\} \) is no rare event, and hence \( \mathbb{P}(aB(\varphi_n) \approx w\varphi_n) \) dictates the tail behavior (thus leading to exponential decay in \( \varphi_n \)). A comparable situation has been considered in [15, Section 2.2]. We stress however that, relative to the \( a > 0 \) case, this \( a \leq 0 \) case has a considerably lower practical relevance.

References

[1] S. Asmussen (2003). *Applied Probability and Queues*, 2nd ed. Springer, New York.
[2] R. Bahadur and R. Ranga Rao (1960). On deviations of the sample mean. *Annals of Mathematical Statistics*, 31, pp. 1015-1027.
[3] R. Bahadur and S. Zabell (1979). Large deviations of the sample mean in general vector spaces. *Annals of Probability* 7, pp. 587-621.
[4] O. Barndorff-Nielsen and D. Cox (1989). *Asymptotic Techniques for Use in Statistics*. Chapman & Hall, London.
[5] A. Bassamboo, S. Ramandeep, and A. Zeevi (2010). Capacity sizing under parameter uncertainty: Safety staffing principles revisited. *Management Science* 56, pp. 1668-1686.
[6] D. Blackwell and J. Hodges (1959). The probability in the extreme tail of a convolution. *Annals of Mathematical Statistics* 30, pp. 1113-1120.
[7] N. Chaganty and J. Sethuraman (1996). Multidimensional strong large deviation theorems. *Journal of Statistical Planning and Inference* 55, pp. 265-280.
[8] H. Chernoff (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics* 23, pp. 493-507.
[9] H. Cramér (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualités Scientifiques et Industrielles* 736, pp. 5-23.
[10] A. Dembo and O. Zeitouni (1998). *Large Deviations Techniques and Applications*, 2nd ed. Springer, New York.
[11] R. Ellis (1984). Large deviations for a general class of random vectors. *Annals of Probability* 12, pp. 1-12.
[12] W. Feller (1971). *An Introduction to Probability Theory and its Applications*, Vol. II. 2nd ed. Wiley, New York.
[13] J. Gärtner (1977). On large deviations from the invariant measure. *Theory of Probability and Its Applications* 22, pp. 24-39.
[14] P. Hall (1992). *The Bootstrap and Edgeworth Expansion*. Springer, New York.
[15] M. Heemskerk, J. Kuhn, and M. Mandjes (2017). Rare-event analysis of mixed Poisson random variables, and applications in staffing. *Performance Evaluation* 114, pp. 56-77.
Appendix A. Edgeworth expansions

In this appendix we establish Edgeworth expansions for $C_n$. We successively address the fast and the slow regime.

A.1. Fast regime. The goal is to develop an expansion for $Q_n(\mathcal{D}_n < x)$. The proof is a variation of that for sums of i.i.d. random variables [12], and therefore we only provide the main steps. First observe that

$$\mathbb{E}_{Q_n} e^{\vartheta \mathcal{D}_n} = (e^{\Gamma_n(\vartheta/\sigma^Q_n \sqrt{n})})^{\varphi_n}, \quad (29)$$

with $\Gamma_n(\vartheta) := \beta(\vartheta + \psi_n) - \beta(\vartheta) - \vartheta \psi_n$. Obviously,

$$e^{\Gamma_n(\vartheta/\sigma^Q_n \sqrt{n})} = \sum_{k=0}^{\infty} \frac{1}{k!} \omega_n^{(k)} = \left( \frac{\vartheta}{\sigma^Q_n \sqrt{n}} \right)^k.$$

Due to the definition of $\vartheta_n$, $\Gamma_n(0) = 0$. It requires some elementary calculus to verify that, with $\gamma_n := \beta''(0)(\alpha'(\vartheta^*))^2 + \beta''(0)\alpha(\vartheta^*)\alpha''(\vartheta^*) + \beta'(0)\alpha'''(\vartheta^*)\psi_1$,

$$\Gamma_n''(0) = \beta''(\vartheta_n) + \gamma_n \psi_n^2 + o(\psi_n^2),$$

and $\Gamma_n''(0) = b \alpha'''(\vartheta^*)\psi_n + o(\psi_n)$. Upon combining the above results, we obtain

$$e^{\Gamma_n(\vartheta/\sigma^Q_n \sqrt{n})} = 1 + \frac{\vartheta^2}{2\varphi_n} + \frac{1}{2} \frac{\gamma_n}{\sigma^Q_n^2} \varphi_n^2 + \frac{1}{6} \frac{b \alpha'''(\vartheta^*)}{\sigma^Q_n^3} \varphi_n^3 + o \left( \max \left\{ \frac{\psi_n}{\varphi_n}, \frac{1}{\varphi_n \sqrt{n}} \right\} \right).$$

Note that we have to distinguish between two cases: $\lim_{n \to \infty} \psi_n \sqrt{n} = 0$ and $\lim_{n \to \infty} \psi_n \sqrt{n} > 0$.

In the case where $\lim_{n \to \infty} \psi_n \sqrt{n} = 0$, we have due to (29),

$$\mathbb{E}_{Q_n} e^{\vartheta \mathcal{D}_n} = \left( 1 + \frac{\vartheta^2}{2\varphi_n} + \frac{1}{6} \frac{b \alpha'''(\vartheta^*)}{\sigma^Q_n^3} \varphi_n \sqrt{n} + o \left( \frac{1}{\varphi_n \sqrt{n}} \right) \right)^{\varphi_n},$$

which can be rewritten as

$$\left( 1 + \frac{\vartheta^2}{2\varphi_n} \right)^{\varphi_n} + \left( 1 + \frac{\vartheta^2}{2\varphi_n} \right)^{\varphi_n-1} \frac{1}{6} \frac{b \alpha'''(\vartheta^*)}{\sigma^Q_n^3} \varphi_n \sqrt{n} + o \left( \frac{1}{\sqrt{n}} \right).$$

By a direct computation, we find

$$\mathbb{E}_{Q_n} e^{\vartheta \mathcal{D}_n} = \exp \left( \frac{1}{2} \vartheta^2 \right) \left( 1 + \frac{1}{6} \frac{b \alpha'''(\vartheta^*)}{\sigma^Q_n^3} \varphi_n \sqrt{n} + o \left( \frac{1}{\sqrt{n}} \right) \right).$$
Using the familiar inversion procedure for characteristic functions, we thus obtain, with \( \phi(\cdot) \) denoting the probability density function of a standard Normal random variable,

\[
\mathbb{Q}_n(D_n \in dx) = \phi(x) \left( 1 + H_3(x) \frac{1}{6} \frac{b \alpha''(\vartheta^*)}{\sigma_+^3} \frac{1}{\sqrt{n}} \right) + o \left( \frac{1}{\sqrt{n}} \right),
\]

with \( H_k(\cdot) \) the Hermite polynomial of degree \( k \). This leads to, with \( \Phi(\cdot) \) denoting the cumulative distribution function of a standard Normal random variable,

\[
\mathbb{Q}_n(D_n \leq x) = \Phi(x) - \phi(x) \left( H_2(x) \frac{1}{6} \frac{b \alpha''(\vartheta^*)}{\sigma_+^3} \frac{1}{\sqrt{n}} \right) + o \left( \frac{1}{\sqrt{n}} \right).
\]  \( \text{(30)} \)

With the same reasoning as in the standard Edgeworth expansion (i.e., that for sums of i.i.d. random variables), the error term (being small relative to \( 1/\sqrt{n} \)) is uniform in \( x \).

In the case where \( \lim \inf_{n \to \infty} \psi_n \sqrt{n} > 0 \), by inserting our expansion into \( (29) \),

\[
\mathbb{E} \mathbb{Q}_n e^{\mathbb{D}_n} = \left( 1 + \frac{2}{\varphi_n} \frac{\vartheta^2}{2} + \frac{1}{2} \frac{\gamma^2}{(\sigma_+^2)^2} \varphi_n + o \left( \frac{\varphi_n}{\varphi_n} \right) \right)^{\varphi_n},
\]

which can be rewritten as

\[
\left( 1 + \frac{2}{\varphi_n} \frac{\vartheta^2}{2} \right)^{\varphi_n} + \left( 1 + \frac{2}{\varphi_n} \frac{\vartheta^2}{2} \right)^{\varphi_n-1} \frac{1}{2} \frac{\gamma^2}{(\sigma_+^2)^2} \varphi_n + o \left( \varphi_n \right).
\]

Using the same procedure as above, this yields, uniformly in \( x \),

\[
\mathbb{Q}_n(D_n \leq x) = \Phi(x) - \phi(x) \left( H_1(x) \frac{1}{2} \frac{\gamma^2}{(\sigma_+^2)^2} \psi_n \right) + o \left( \psi_n \right).
\]

In our setting, however, we need an error that is small relative to \( 1/\sqrt{n} \), which can be achieved by expanding \( \Gamma''_n(0) \) further (as it can be verified that the contributions of the higher derivatives \( \Gamma''_n(k)(0) \) for \( k \geq 3 \) are \( o(1/\sqrt{n}) \)). Define

\[
k_+ := \sup \{ k \in \mathbb{N} : \lim \inf_{n \to \infty} \psi_n^k \sqrt{n} > 0 \};
\]

due to Assumption 1, this is a finite constant. Using the above type of reasoning, we conclude that there are constants \( c_1, \ldots, c_{k_+} \) so that

\[
\mathbb{Q}_n(D_n \leq x) = \Phi(x) - \phi(x) \left( H_1(x) \sum_{k=1}^{k_+} c_k \psi_n^k + H_2(x) \frac{1}{6} \frac{b \alpha''(\vartheta^*)}{\sigma_+^3} \frac{1}{\sqrt{n}} \right) + o \left( \frac{1}{\sqrt{n}} \right).
\]  \( \text{(31)} \)

In the boundary case where \( \psi_n \sqrt{n} \) converges to a constant we have \( k_+ = 1 \); the sum in \( (31) \) consists of only one term.

A.2. Slow regime. Now our objective is to find an expansion for \( \mathbb{Q}_n(F_n \leq x) \); the reasoning is analogous to that of the fast regime. The starting point is the mgf of \( F_n \) under the twisted distribution:

\[
\mathbb{E} \mathbb{Q}_n e^{\mathbb{F}_n} = \left( e^{\Gamma_n(\vartheta/\sigma^2 \sqrt{\psi_n})} \right)^{\varphi_n},
\]  \( \text{(32)} \)

with, as before, \( \Gamma_n(\vartheta) := \beta(\alpha(\vartheta + \vartheta_n)\psi_n) - \beta(\alpha(\vartheta_n)\psi_n) - \vartheta_n \psi_n \). Now,

\[
\frac{1}{2} \left( \frac{\vartheta}{\sigma^2 \sqrt{\psi_n}} \right)^2 \Gamma''_n(0) = \frac{1}{2} \frac{\vartheta^2}{(\sigma_+^2)^2} \left( \beta''(\alpha(\vartheta_n)\psi_n)(\vartheta')^2 \frac{1}{\varphi_n} + \beta'(\alpha(\vartheta_n)\psi_n) \alpha''(\vartheta_n) \frac{1}{n} \right).
\]  \( \text{(33)} \)

Using the expansion of \( \vartheta_n \), the right-hand side of the previous display reads

\[
\frac{1}{2} \frac{\vartheta^2}{(\sigma_+^2)^2} \left( \beta''(\alpha \vartheta^*) \alpha^2 \frac{1}{\varphi_n} + \gamma^2 \frac{1}{n} + o \left( \frac{1}{n} \right) \right) = \frac{\vartheta^2}{2\varphi_n} + \frac{1}{2} \frac{\gamma^2}{(\sigma_+^2)^2 n} \frac{\vartheta^2}{n} + o \left( \frac{1}{n} \right),
\]
where \( \gamma := \beta'(a\tau^*)a''(0) + 2a a''(0)\beta''(a\tau^*)\tau^* + \frac{1}{3} a^2 a''(0)\beta'''(a\tau^*) (\tau^*)^2 + a^3 \beta'''(a\tau^*)w_2 \). In addition,

\[
\frac{1}{6} \left( \frac{\vartheta}{\sigma^2 \sqrt{n} \psi_n} \right)^3 \Gamma_n(0) = \frac{1}{6} \frac{\beta''(a\tau^*)a^3}{(\sigma^2)^3} \vartheta^3 \varphi_n^{1/2} + o \left( \frac{1}{\varphi_n^{1/2}} \right).
\]

We distinguish between the cases \( \lim_{n \to \infty} \varphi_n^{3/2}/n = 0 \) and \( \lim_{n \to \infty} \varphi_n^{3/2}/n > 0 \). Mimicking the reasoning used for the fast regime, in the case where \( \lim_{n \to \infty} \varphi_n^{3/2}/n = 0 \),

\[
\mathbb{E}_{Q_n} e^{\vartheta E_n} = \left( 1 + \frac{\vartheta^2}{2\varphi_n} + \frac{1}{6} \frac{\beta''(a\tau^*)a^3}{(\sigma^2)^3} \varphi_n^{1/2} + o \left( \frac{1}{\varphi_n^{1/2}} \right) \right)^{\varphi_n},
\]

which can be rewritten as

\[
\left( 1 + \frac{\vartheta^2}{2\varphi_n} \right)^{\varphi_n} + \left( 1 + \frac{\vartheta^2}{2\varphi_n} \right)^{\varphi_n-1} \frac{1}{6} \frac{\beta''(a\tau^*)a^3}{(\sigma^2)^3} \frac{\vartheta^3}{\sqrt{\varphi_n}} + o \left( \frac{1}{\sqrt{\varphi_n}} \right).
\]

This leads to, inserting our expansion into (32),

\[
\mathbb{E}_{Q_n} e^{\vartheta E_n} = \exp \left( \frac{1}{2} \vartheta^2 \right) \left( 1 + \frac{1}{6} \frac{\beta''(a\tau^*)a^3}{(\sigma^2)^3} \frac{\vartheta^3}{\sqrt{\varphi_n}} \right) + o \left( \frac{1}{\sqrt{\varphi_n}} \right),
\]

and, after inversion, to the Edgeworth expansion

\[
Q_n(E_n \leq x) = \Phi(x) - \phi(x) \left( H_2(x) \frac{1}{6} \frac{\beta''(a\tau^*)a^3}{(\sigma^2)^3} \frac{1}{\sqrt{\varphi_n}} \right) + o \left( \frac{1}{\sqrt{\varphi_n}} \right). \tag{34}
\]

Finally, focus on \( \lim_{n \to \infty} \varphi_n^{3/2}/n > 0 \). Performing the same steps,

\[
\mathbb{E}_{Q_n} e^{\vartheta E_n} = \exp \left( \frac{1}{2} \vartheta^2 \right) \left( 1 + \frac{1}{6} \frac{\gamma}{(\sigma^2)^2} \frac{\vartheta^2}{\psi_n} + o \left( \frac{1}{\psi_n} \right) \right),
\]

leading to, uniformly in \( x \),

\[
Q_n(E_n \leq x) = \Phi(x) - \phi(x) \left( H_1(x) \frac{1}{2} \frac{\gamma}{(\sigma^2)^2} \frac{1}{\psi_n} \right) + o \left( \frac{1}{\psi_n} \right).
\]

Our objective, however, is to obtain an error that is \( o(1/\sqrt{\varphi_n}) \). This is achieved by expanding (33) further; again the contributions of the higher derivatives (i.e., derivatives of order 3 and higher) can be verified to be negligible. Following the line of reasoning of the fast regime, we define

\[
k_- := \sup \left\{ k \in \mathbb{N} : \lim_{n \to \infty} \frac{\varphi_n^{1/2}}{\psi_n^k} > 0 \right\};
\]

due to Assumption 2, this is a finite constant. Using the above type of reasoning, we conclude that there are constants \( c_1', \ldots, c_{k_-}' \) so that

\[
Q_n(E_n \leq x) = \Phi(x) - \phi(x) \left( H_1(x) \sum_{k=1}^{k_-} c_k' \frac{1}{\varphi_n^{1/2}} + H_2(x) \frac{1}{6} \frac{\beta''(a\tau^*)a^3}{(\sigma^2)^3} \frac{1}{\varphi_n^{1/2}} \right) + o \left( \frac{1}{\sqrt{\varphi_n}} \right). \tag{35}
\]

Again, in the boundary case where \( \varphi_n^{3/2}/n \) converges to a constant we have \( k_- = 1 \) and the sum in (35) consists of only one term.