

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/102350>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# SINGULARITY ANALYSIS FOR HEAVY-TAILED RANDOM VARIABLES

NICHOLAS M. ERCOLANI, SABINE JANSEN, DANIEL UELTSCHI

**ABSTRACT.** We propose a novel complex-analytic method for sums of i.i.d. random variables that are heavy-tailed and integer-valued. The method combines singularity analysis, Lindelöf integrals, and bivariate saddle points. As an application, we prove three theorems on precise large and moderate deviations which provide a local variant of a result by S. V. Nagaev (1973). The theorems generalize five theorems by A. V. Nagaev (1968) on stretched exponential laws  $p(k) = c \exp(-k^\alpha)$  and apply to logarithmic hazard functions  $c \exp(-(\log k)^\beta)$ ,  $\beta > 2$ ; they cover the big jump domain as well as the small steps domain. The analytic proof is complemented by clear probabilistic heuristics. Critical sequences are determined with a non-convex variational problem.

## 1. INTRODUCTION

The motivation of the present article is two-fold. First, we present a new analytic method for the investigation of large powers of generating functions of sequences that satisfy some analyticity and log-convexity conditions. The method is explained and developed for probability generating functions but it has potentially broader applications and is motivated by techniques commonly used in analytic combinatorics [11]. Specifically, we show that methods akin to singularity analysis can be pushed beyond the realm of functions amenable to singularity analysis in the sense of [11, Chapter VI.1].

Second, we explore consequences for probabilistic limit laws and prove three theorems on precise large and moderate deviations for sums of independent identically distributed (i.i.d.) random variables that are heavy-tailed [7] and integer-valued. The theorems generalize results on stretched exponential laws by A. V. Nagaev [15] which have recently attracted interest in the context of the zero-range process [2]. They are close in spirit to results by S. V. Nagaev [17], however with more concrete conditions on the domain of validity of the theorems, and provide deviations results “on the whole axis” [20]. Our assumptions are more restrictive than one may wish from a probabilistic perspective; in return, they allow for sharp results and may

---

*Date:* 21 April 2017.

*1991 Mathematics Subject Classification.* 05A15; 30E20; 44A15; 60F05; 60F10.

*Key words and phrases.* local limit laws; large deviations; heavy-tailed random variables; asymptotic analysis; Lindelöf integral; singularity analysis; bivariate steepest descent.

Department of Mathematics, The University of Arizona, Tucson, AZ 85721–0089, (ercolani@math.arizona.edu). Supported by NSF grant DMS-1212167.

Department of Mathematics, University of Sussex, Falmer Campus, Brighton BN1 9QH, United Kingdom (s.c.jansen@ac.uk).

Department of Mathematics, University of Warwick, Coventry, CV4 7AL, United Kingdom (daniel@ueltschi.org).

provide a helpful class of explicit reference examples. For example, we prove that one of the bounds of the (local) big-jump domain for logarithmic hazard functions derived in [4] is sharp.

The analytic proof of the theorems is complemented by clear probabilistic heuristics. Our results cover different regimes: a small steps or moderate deviations regime, where a classical variant of a local central limit theorem with corrections expressed with the Cramér series holds [12], and a big-jump regime where the large deviation is realized by making one out of the  $n$  variables large. In the language of statistical mechanics and the zero-range process, they correspond to supersaturated gas and a condensed phase [2]. The critical scales that distinguish between regimes are defined with the help of a non-convex variational problem which encodes competing probabilistic effects. This complements the approach taken by Denisov, Dieker and Shneer [4].

The study of combinatorial generating functions shares much in common with the study of probability generating functions; in fact in many instances they coincide or run parallel as is the case for more recent investigations in the area of random combinatorial structures [1]. From the viewpoint of complex function theory the key here involves relating asymptotic questions to questions about the nature of the singularities of generating functions viewed as more global analytic objects. In most of the successful applications of this *singularity analysis* to coefficient asymptotics a bridge is provided by the realization that the series in question satisfies some global algebraic or differential equation. Generating functions for which this is the case are referred to as *holonomic*. Pushing beyond this class in a systematic way requires new ideas and one of the most promising of these is the use of Lindelöf integrals. Lindelöf introduced these classically [13] as a means to constructively carry out analytic continuations of function elements (series) in a fairly general setting. In more recent times his construction has begun to be used to study non-holonomic combinatorial generating functions [10]. The generating functions for heavy-tailed distributions studied in this paper are of non-holonomic type and our methods of studying them show a new application of Lindelöf's construction that has novel connections to other areas of analytic asymptotic analysis such as bivariate steepest descent. In future work we hope to build on the present article in a way that broadens the application of harmonic analysis and complex function theory to problems of asymptotic analysis in both probability and combinatorics, such as applying the theory of Hardy spaces and Riemann-Hilbert analysis and extensions of Tauberian theorems as originally envisioned by Paley and Wiener [9].

Our proof shares some features with [17], where cumulative distribution functions are approximate Laplace transforms and approximating moment generating function admit analytic extensions (see Section 7). Contour integrals that appear in inversion formulas are deformed and analyzed by Gaussian approximation—our proof details in Section 5.4 follow [17]. There are, however, key differences: we need not deal with approximation errors because of stronger analyticity assumptions, and our detailed analysis of the underlying variational problem allows us to formulate more concrete conditions for our theorems.

The remainder of this article is organized as follows. In Section 2 we formulate our main results and discuss applications to stretched exponential weights  $c \exp(-k^\alpha)$  and weights  $c \exp(-(\log x)^\beta)$  with logarithmic hazard functions. In Section 3 we explain the proof strategy in five steps, which are treated in detail

in the remaining sections. Steps 1 and 2 concern analytic extensions and notably use the Lindelöf and Bromwich integrals (Section 4). Steps 4 and 5 analyze the critical points of a bivariate function and deal with the Gaussian approximation to a double integral (Section 5). The pivotal Step 3 connects the contour integral and the bivariate double integral; it leads to the full proof of our theorems that can be found in Section 6. In Section 7 we sketch how the method developed in this paper may be extended to more general settings, in particular when the analyticity required by our proof methods is not a priori given.

## 2. RESULTS

We use the notation  $a_n \sim b_n \Leftrightarrow a_n = (1 + o(1))b_n$  and  $a_n \ll b_n \Leftrightarrow a_n = o(b_n)$ .

**2.1. Preliminaries.** In order to formulate the results, we need to introduce critical sequences deduced from a variational problem and the Cramér series. Let  $X, X_1, X_2, \dots$  be independent, identically distributed random variables with values in  $\mathbb{N}$  and law

$$\mathbb{P}(X = k) = p(k) = \exp(-q(k)) \quad (k \in \mathbb{N}) \quad (2.1)$$

for some sequence  $(q(k))_{k \in \mathbb{N}}$ . We assume that  $X$  is heavy-tailed and has moments of all orders, i.e., the generating function

$$G(z) = \sum_{k=1}^{\infty} p(k)z^k \quad (|z| \leq 1) \quad (2.2)$$

has radius of convergence 1 and  $\mathbb{E}[X^m] = \sum_{k=1}^{\infty} k^m p(k) < \infty$  for all  $m \in \mathbb{N}$ . Let  $\mu$  and  $\sigma^2$  be the expectation and variance of  $X$ . Set  $S_n = X_1 + \dots + X_n$ . We are interested in the asymptotic behavior of  $\mathbb{P}(S_n = \mu n + N_n)$  when  $n, N_n \rightarrow \infty$  with  $N_n \gg \sqrt{n}$ . The following assumption is similar to conditions considered by S. V. Nagaev [17].

**Assumption 2.1.** For some  $a > 0$ , the sequence  $(q(k))_{k \in \mathbb{N} \cap (a, \infty)}$  extends to a smooth function  $q : (a, \infty) \rightarrow \mathbb{R}$  which has the following properties:

- (i)  $q' > 0$ ,  $q'' < 0$ , and  $q^{(3)} > 0$ .
- (ii)  $\lim_{x \rightarrow \infty} xq'(x)/(\log x) = \infty$ .
- (iii)  $c_1 \frac{q'(x)}{x} \leq |q''(x)| \leq c_2 \frac{q'(x)}{x}$  for some constants  $c_1, c_2 > 0$ .
- (iv)  $c_3 \frac{|q''(x)|}{x} \leq q^{(3)}(x) \leq c_4 \frac{|q''(x)|}{x}$  for some constants  $c_3, c_4 > 0$ .
- (v)  $q'(x) \leq \alpha \frac{q(x)}{x}$  for some  $\alpha \in (0, 1)$ .

Assumption 2.1 allows for an easy analysis of an auxiliary variational problem, which is essential to the formulation of our main results. Let us collect a few elementary consequences. Under Assumption 2.1,  $q$  is concave on  $(a, \infty)$  and  $p = \exp(-q)$  is log-convex. Moreover,  $\lim_{x \rightarrow \infty} x^2 q''(x)/\log x = -\infty$  and for  $y > x > a$ , using

$$\frac{q'(y)}{q'(x)} = \exp\left(-\int_x^y \frac{|q''(u)|}{q'(u)} du\right), \quad (2.3)$$

we estimate

$$\left(\frac{y}{x}\right)^{-c_2} \leq \frac{q'(y)}{q'(x)} \leq \left(\frac{y}{x}\right)^{-c_1} \leq 1 \quad (2.4)$$

Similarly, for  $y > x > a$ ,

$$\left(\frac{y}{x}\right)^{c_3} \leq \frac{q''(y)}{q''(x)} \leq \left(\frac{y}{x}\right)^{c_4}. \quad (2.5)$$

Since  $G(z) = \sum_k z^k \exp(-q(k))$  has radius of convergence 1, we also know that

$$\lim_{x \rightarrow \infty} q'(x) = \lim_{x \rightarrow \infty} q''(x) = \lim_{x \rightarrow \infty} q^{(3)}(x) = 0. \quad (2.6)$$

Indeed by Assumption 2.1,  $q'$  is eventually decreasing and the limit  $\ell := \lim_{x \rightarrow \infty} q'(x)$  exists in  $\mathbb{R} \cup \{-\infty\}$ . Then  $\ell = \lim_{x \rightarrow \infty} q(x)/x$  and  $G(z)$  has radius of convergence  $\exp(\ell) = 1$ , whence  $\ell = 0$ . Assumption 2.1(iii) and (iv) leads to the statements on higher order derivatives. Assumption 2.1(v) implies  $q(x) = O(x^\alpha)$  as  $x \rightarrow \infty$ .

Our method of proof requires two more analyticity assumptions, though in Section 7 we discuss how, following [17], one might be able to relax these assumptions.

**Assumption 2.2.** There exists  $b \geq 0$  such that  $(p(n))_{n \in \mathbb{N} \cap [b, \infty)}$  extends to a function  $p(\xi)$  that is continuous on a closed half-plane  $\operatorname{Re} \xi \geq b$ , analytic on the open half-plane  $\operatorname{Re} \xi > b$ , and in addition satisfies

- (i) For every  $\varepsilon \in (0, \pi)$ , some  $C_\varepsilon > 0$ , and all  $\xi$ , we have  $|p(\xi)| \leq C_\varepsilon \exp(\varepsilon|\xi|)$ .
- (ii)  $\int_{-\infty}^{\infty} |(b + is)^k p(b + is)| ds < \infty$  for all  $k \in \mathbb{N}$ .

Moreover  $p(x) = \exp(-q(x))$  for all  $x \geq \max(a, b)$  with  $a, q(x)$  as in Assumption 2.1.

**Assumption 2.3.** Let  $p(\xi) = \exp(-q(\xi))$  be the analytic extension from Assumption 2.2, defined in  $\operatorname{Re} \xi \geq b$ . Then  $q(\xi) = -\operatorname{Log} \xi$ , defined with the principal branch of the logarithm is analytic as well, and the following holds:

- (i) For  $r > 0$  large, let  $z_r = b + i\sqrt{r^2 - b^2}$ . Then as  $r \rightarrow \infty$ ,

$$\left| \int_{\operatorname{Re} \xi = b, |\xi| \geq r} \exp(-\operatorname{Re} q(\xi)) d\xi \right| \leq \exp(-\operatorname{Re} q(z_r) + O(\log r)).$$

- (ii)  $\operatorname{Im}(\xi q'(\xi)) \leq \operatorname{Im}(\xi q'(r))$  for all large  $r$  and all  $\xi$  with  $\operatorname{Im} \xi \geq 0$  and  $|\xi| = r$ .
- (iii)  $|q^{(3)}(\xi)| \leq C|q''(\xi)/\xi|$  for some  $C > 0$  and all  $\xi$ .

Assumption 2.3 enters the proof of Theorem 4.4 only.

*Variational problem and critical scale.* Assumption 2.1 is tailored to the analysis of an auxiliary variational problem, motivated by the following heuristics. For subexponential random variables, the typical large deviations behavior is realized by making one out of the  $n$  variables large,

$$\mathbb{P}(S_n = \mu n + N_n) \approx n \mathbb{P}(X_n = N_n - k_n) \mathbb{P}(S_{n-1} = \mu n + k_n) \quad (2.7)$$

with a yet to be determined optimal  $k_n$ . Assuming that a normal approximation for the second factor is justified, we get

$$\mathbb{P}(S_n = \mu n + N_n) \approx \exp\left(-q(N_n - k_n) - \frac{k_n^2}{2n\sigma^2}\right) \quad (2.8)$$

where we have neglected prefactors  $n$  and  $1/\sqrt{2\pi n\sigma^2}$  (see Eq. (2.11) below for a more refined heuristics). The optimal  $k_n$  is then determined by minimizing the term in the exponential. Thus we are led to the minimization of

$$f_n(x) = q(x) + \frac{(N_n - x)^2}{2n\sigma^2}. \quad (2.9)$$

As illustrated in Figure 1, the nature of the variational problem changes with  $N_n$ . Define  $x_n^* > 0$  and  $N_n^*$  by

$$q''(x_n^*) = -\frac{1}{n\sigma^2}, \quad N_n^* = x_n^* + n\sigma^2 q'(x_n^*). \quad (2.10)$$

For sufficiently large  $n$ , the inflection point  $x_n^*$  is uniquely defined because of the monotonicity from Assumption 2.1 and Eq. (2.6), moreover  $x_n^* \rightarrow \infty$ . The quantity  $N_n^*$  is defined in such a way that the tangent to the curve  $y = q'(x)$  at  $x = x_n^*$  has equation  $y = (N_n^* - x)/(n\sigma^2)$ , see Fig. 3.

The next two lemmas characterize the minimization of  $f_n$ ; they are proven in Section 5.1. The first lemma relates the critical points of  $f_n$  to the location of  $N_n$  compared to  $N_n^*$ .

**Lemma 2.4.** *For sufficiently large  $n$ , the following holds true:*

- (a) *If  $N_n < N_n^*$ , then  $f_n' > 0$  on  $(a, \infty)$ .*
- (b) *If  $N_n = N_n^*$ , then  $f_n'$  has the unique zero  $x_n^*$ , moreover  $f_n'(x) \geq 0$  with equality if and only if  $x = x_n^*$ .*
- (c) *If  $N_n > N_n^*$  and  $\limsup_{n \rightarrow \infty} N_n/(n\sigma^2) < \lim_{x \searrow a} q'(x)$ , then  $f_n$  has exactly two critical points  $x_n$  and  $x_n'$ , which satisfy  $x_n' < x_n^* < x_n < N_n$  and*

$$f_n(x_n') = \max_{(a, x_n^*)} f_n, \quad f_n(x_n) = \min_{(x_n^*, \infty)} f_n.$$

- (d) *If  $N_n > N_n^*$  and  $\liminf_{n \rightarrow \infty} N_n/(n\sigma^2) > \lim_{x \searrow a} q'(x)$ , then  $f_n$  has a unique critical point  $x_n$ . It satisfies  $x_n \in (x_n^*, N_n)$  and is a global minimizer.*

For  $N_n > N_n^*$ , the function  $f_n$  may have two local minimizers:  $a$  and  $x_n$ , and we may wonder which one is the global minimizer. The answer depends on the location of  $N_n$  compared to a new critical sequence  $N_n^{**}$ . Concrete examples are given in Sections 2.3 and 2.4.

**Lemma 2.5.** *For  $n$  sufficiently large, there is a uniquely defined  $N_n^{**} > N_n^*$  such that:*

- (a) *If  $N_n^* < N_n < N_n^{**}$ , then  $f_n(a) < f_n(x_n)$ .*
- (b) *If  $N_n = N_n^{**}$ , then  $f_n(a) = f_n(x_n)$ .*
- (c) *If  $N_n > N_n^{**}$ , then  $f_n(x_n) < f_n(a)$ .*

In general it may not be straightforward to determine  $N_n^{**}$  exactly, but it is simple to find a lower bound: if  $a = 0$  and  $q(N_n) < N_n^2/(2n\sigma^2)$ , then  $N_n > N_n^{**}$ . This lower bound corresponds, roughly, to the sequence  $\Lambda(n)$  in [17].

The sequences introduced up to now are ordered as follows.

**Lemma 2.6.** *As  $n \rightarrow \infty$ , we have*

$$\sqrt{n} \ll x_n^* < N_n^* < N_n^{**} = O(n^{1/(2-\alpha)}) \ll n,$$

*and for some constants  $C, \delta > 0$*

$$(1 + \delta)x_n^* \leq N_n^* \leq Cx_n^*.$$

The lemma is proven in Section 5.1. Lemmas 2.4 and 2.5, together with the heuristics described above, suggest that for  $N_n > N_n^{**}$  the unlikely event  $S_n = n\mu + N_n$  is realized by making one component of the order of  $x_n$ . One may wonder how far  $x_n$  is from swallowing all of the overshoot  $N_n$ .

**Lemma 2.7.** *Suppose Assumption 2.1(i) holds true. Let  $N_n > N_n^*$ . Then*

$$N_n - N_n^* \leq x_n \leq N_n, \quad n\sigma^2 f_n''(x_n) = 1 - n\sigma^2 |q''(x_n)| = 1 + O\left(\frac{N_n^*}{N_n}\right).$$

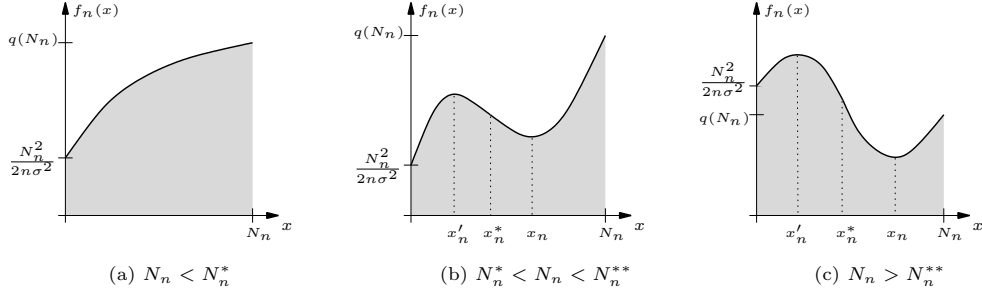


FIGURE 1. Minimization of  $f_n(x) = q(x) + (N_n - x)^2/(2n\sigma^2)$  and illustration of Lemmas 2.4 and 2.5 for weights  $q : (0, \infty) \rightarrow \mathbb{R}$  with  $q(0) = 0$ . For  $N_n > N_n^*$ ,  $f_n(x)$  has two critical points  $x'_n$  and  $x_n$  separated by an inflection point  $x_n^*$ . The global minimum is reached either at  $x = x_n$  or at  $x = 0$ .

In particular, for  $N_n \gg N_n^*$ , we have  $x_n \sim N_n$  and  $n\sigma^2 f''(x_n) \rightarrow 1$ . The lemma is proven in Section 5.1. The information on the second derivative enters a refined heuristics: we make the ansatz that conditional on the unlikely event  $S_n = n\mu + N_n$ , there is one large component of size  $x_n$ , but the size is not deterministic. Instead, there are fluctuations around  $x_n$ . This yields

$$\begin{aligned}
\mathbb{P}(S_n = n\mu + N_n) &\approx n \sum_{\ell} \mathbb{P}(X_1 = x_n + \ell) \mathbb{P}(S_{n-1} = N_n - x_n - \ell) \\
&\approx n \sum_{\ell} \frac{\exp(-f_n(x_n + \ell))}{\sqrt{2\pi n\sigma^2}} \\
&\approx n e^{-f_n(x_n)} \sum_{\ell} \frac{\exp(-f_n''(x_n)\ell^2/2)}{\sqrt{2\pi n\sigma^2}} \\
&\approx \frac{n \exp(-f_n(x_n))}{\sqrt{1 - n\sigma^2|q''(x_n)|}}. \tag{2.11}
\end{aligned}$$

Theorem 2.11 below confirms the heuristics for large  $N_n$ , up to correction terms both in the prefactor and in the exponential.

*Cumulants and Cramér series.* The heuristics together with Lemma 2.7 suggest that the optimal  $k_n = N_n - x_n$  in Eq. (2.8) is of order up to  $N_n^* \gg \sqrt{n}$ . At this scale the normal approximation fails and requires correction terms. The latter are usually expressed with the Cramér series [12], whose definition we briefly recall. Let  $\varphi(t)$  be the cumulant generating function of  $X$ ,

$$\varphi(t) = \log \mathbb{E}[e^{tX}] = \log G(e^t) \quad (\operatorname{Re} t \leq 0). \tag{2.12}$$

Note  $\varphi(0) = 0$ . As  $t \rightarrow 0$ ,  $\varphi(t)$  can be approximated to arbitrary order

$$\varphi(t) = \sum_{j=1}^r \kappa_j \frac{t^j}{j!} + O(t^{r+1}) \tag{2.13}$$

with finite and real expansion coefficients  $\kappa_j \in \mathbb{R}$ , the cumulants, see Section 4. Notice that  $\kappa_1 = \mu$  is the expectation and  $\kappa_2 = \sigma^2 > 0$  is the variance of  $X$ .

**Definition 2.8.** Let  $t(\tau) = \frac{1}{\sigma^2}\tau + \sum_{j \geq 2} a_j \tau^j$  be the formal power series obtained by inverting

$$\sigma^2 t(\tau) + \sum_{j \geq 2} \kappa_{j+1} \frac{t(\tau)^j}{j!} = \tau.$$

The Cramér series  $\sum_{j \geq 0} \lambda_j \tau^j$  is the formal power series defined by composing the expansion of  $t(\tau)$  with the expansion of  $(\mu + \tau)t - \varphi(t)$ ,

$$(\mu + \tau)t(\tau) - \sum_{j \geq 1} \kappa_j \frac{t(\tau)^j}{j!} = -\frac{\tau^2}{2\sigma^2} + \tau^3 \sum_{j \geq 0} \lambda_j \tau^j.$$

Equivalently, the Cramér series is the left-sided Taylor expansion of the Legendre transform  $\varphi^*$  at  $\mu$ : let  $\varphi^*(x) := \sup_{t \leq 0} (tx - \varphi(t))$ . Then as  $\tau \nearrow 0$ ,

$$\varphi^*(\mu + \tau) = -\frac{\tau^2}{2\sigma^2} + \tau^3 \sum_{j=0}^r \lambda_j \tau^j + O(\tau^{r+4}) \quad (2.14)$$

to arbitrarily high order  $r$ .

*Remark 1.* For  $t > 0$ ,  $\log[\sum_{k \geq 1} p(k) \exp(kt)]$  is infinite and the standard convention is to set  $\varphi(t) = \infty$ ; then  $\varphi^*(\mu + \tau) \equiv 0$  for  $\tau \geq 0$  and Eq. (2.14) no longer applies. We adopt a different convention, however, for which  $\varphi(t)$  is smooth in a neighborhood of 0 (see Theorem 4.2), though it becomes complex-valued, and Eq. (2.14) applies to  $(\operatorname{Re} \varphi)^*$  for positive  $\tau$  as well.

**2.2. Main theorems.** Set  $f_{n0} = f_n$  and for  $r \geq 1$ ,

$$f_{nr}(x) = q(x) + \frac{(N_n - x)^2}{2n\sigma^2} - \frac{(N_n - x)^3}{n^2} \sum_{j=0}^{r-1} \lambda_j \left(\frac{N_n - x}{n}\right)^j. \quad (2.15)$$

Remember the minimization of  $f_n(x)$  and the critical scales  $\sqrt{n} \ll N_n^* < N_n^{**} = O(n^{1/(2-\alpha)}) \ll n$  illustrated in Fig. 1. In Proposition 2.12 below we check that the properties of  $f_n$  carry over to  $f_{nr}$ . The following theorems provide a local variant of a large deviations theorem by S. V. Nagaev [17], see also [18, Theorem 2.1]. The principal difference, apart from the local character of the theorems, is that our detailed investigation of the variational problem and notably Lemma 2.7 allows us to formulate conditions directly in terms of  $N_n$ , whereas S. V. Nagaev's criteria included an indirect condition on the sign of some second derivative.

**Theorem 2.9.** Let  $N_n \rightarrow \infty$  with  $\sqrt{n} \ll N_n \leq (1 + o(1))N_n^*$ . Pick  $r$  large enough so that  $n(N_n/n)^r \rightarrow 0$ . Then

$$\mathbb{P}(S_n = \mu n + N_n) \sim \frac{1}{\sqrt{2\pi\sigma^2 n}} \exp\left(-\frac{N_n^2}{2n\sigma^2} + \frac{N_n^3}{n^2} \sum_{j=0}^{r-1} \lambda_j \left(\frac{N_n}{n}\right)^j\right).$$

**Theorem 2.10.** Let  $N_n \rightarrow \infty$  with  $\liminf N_n/N_n^* > 1$  and  $N_n = O(n^{1/[2-\alpha]})$ . Pick  $r$  large enough so that  $n(N_n^*/n)^r \rightarrow 0$ . Then

$$\begin{aligned} \mathbb{P}(S_n = \mu n + N_n) &= (1 + o(1)) \frac{1}{\sqrt{2\pi\sigma^2 n}} \exp\left(-\frac{N_n^2}{2n\sigma^2} + \frac{N_n^3}{n^2} \sum_{j=0}^{r-1} \lambda_j \left(\frac{N_n}{n}\right)^j\right) \\ &\quad + (1 + o(1)) \frac{n}{\sqrt{1 - n\sigma^2 |q''(x_{nr})|}} \exp\left(-f_{nr}(x_{nr})\right) \end{aligned}$$



with  $x_{nr} = N_n + O(N_n^*)$  the largest solution of  $f'_{nr}(x_{nr}) = 0$ .

Lemma 2.5 suggests that for  $N_n \ll N_n^{**}$ , the first contribution dominates and for  $N_n \gg N_n^{**}$  the second contribution wins, but one has to be careful because of the factors  $n$  and  $1/\sqrt{2\pi n\sigma^2}$  as well as the Cramér corrections; a detailed evaluation is best left to concrete examples (see, however, Corollary 2.13 below).

**Theorem 2.11.** *Let  $N_n \rightarrow \infty$  with  $N_n \gg n^{1/(2-\alpha)}$ . Pick  $r$  large enough so that  $n(N_n^*/n)^r \rightarrow 0$ . Then*

$$\mathbb{P}(S_n = \mu n + N_n) \sim n \exp\left(-f_{nr}(x_{nr})\right)$$

with  $x_{nr} = N_n + O(N_n^*)$  the largest solution of  $f'_{nr}(x_{nr}) = 0$ .

In practice one may prefer not to deal with the Cramér corrections or the variational problem, and the following proposition is helpful.

**Proposition 2.12.** *Suppose that  $\liminf_{n \rightarrow \infty} N_n/N_n^* > 1$ . Fix  $r \in \mathbb{N}_0$ . Then, for sufficiently large  $n$ ,  $x_{nr}$  is the maximizer of  $f_{nr}$  restricted to  $(x_n^*, N_n)$  and the unique zero of  $f'_{nr}$  in that interval. Moreover  $1 - n\sigma^2|q''(x_{nr})| = 1 + O(N_n^*/N_n)$  stays bounded away from zero and*

$$\begin{aligned} x_{nr} &= N_n - (1 + o(1))n\sigma^2 q'(x_{nr}) = N_n + O(N_n^*), \\ f_{nr}(x_{nr}) &= q(x_{nr}) + \frac{1}{2}(1 + o(1))n\sigma^2 q'(x_{nr}) = q(N_n) \left(1 + O\left(\frac{N_n^*}{N_n}\right)\right). \end{aligned}$$

The proposition is proven in Section 5.1. For  $N_n \gg N_n^*$ , we obtain  $x_{nr} \sim N_n$ ,  $q'(x_{nr}) \sim q'(N_n)$ , and  $q(x_{nr}) = q(N_n) - (1 + o(1))n\sigma^2 q'(N_n)$ , hence

$$f_{nr}(x_{nr}) = q(N_n) - \frac{1}{2}(1 + o(1))n\sigma^2 q'(N_n)^2. \quad (2.16)$$

Now suppose in addition that  $\liminf \frac{N_n^2}{2n\sigma^2}/q(N_n) > 1$ . Then in Theorem 2.10, the first summand is of order  $\exp(-(1 + o(1))N_n^2/(2n\sigma^2))$ , the second of order  $\exp(-(1 + o(1))q(N_n))$ , so the first contribution is negligible and the validity of Theorem 2.11 extends accordingly, since  $1 - n\sigma^2|q''(x_{nr})| \rightarrow 1$  for  $N_n \gg N_n^*$ . Eq. (2.16) now yields the following corollary.

**Corollary 2.13.** *Take  $N_n \rightarrow \infty$  with  $N_n \gg N_n^*$  and  $\liminf \frac{N_n^2}{2n\sigma^2}/q(N_n) > 1$ . Then*

$$\mathbb{P}(S_n = n\mu + N_n) \sim n e^{-q(N_n)} = n\mathbb{P}(X = N_n)$$

if and only if  $\sqrt{n\sigma^2}q'(N_n) \rightarrow 0$ .

In concrete examples, Theorem 2.9 should allow us to extend the domain of validity of the corollary to  $N_n \gg N_n^{**}$ . The condition  $\sqrt{n\sigma^2}q'(N_n) \rightarrow 0$  is closely related to the *insensitivity* scale discussed by Denisov, Dieker and Shneer [4], as

$$\frac{p(N_n \pm \sqrt{n\sigma^2})}{p(N_n)} \rightarrow 1 \Leftrightarrow \sqrt{n\sigma^2}q'(N_n) \rightarrow 0. \quad (2.17)$$

*Remark 2* (Big-jump vs small steps). The domain where  $\mathbb{P}(S_n = n\mu + N_n) \sim n\mathbb{P}(X = N_n)$  is sometimes called *big-jump domain*. Think of  $S_n$  as the position of a random walker with step size distribution  $p(k)$ . In the situation of Corollary 2.13, the unlikely event that the walker has travelled a distance  $\mu n + N_n$  much larger than the expected distance  $\mu n$  is realized by one big jump of size  $N_n$ . Finding the boundary of the big-jump domain is an active field of research [4].

The interpretation of Theorem 2.9, in contrast, is that the moderate overshoot  $N_n$  is achieved by a collective effort: all steps tend to stay small, though each stretches a little beyond its expected value  $\mu$ . For stretched exponential variables, this interpretation is made rigorous in [15] and [2].

*Remark 3* (Condensation in the zero-range process). In the zero-range process, the random variables  $X_1, \dots, X_n$  model the number of particles at lattice sites  $j = 1, \dots, n$ , with  $S_n$  the total number of particles, and  $\mu$  is a critical density. Theorem 2.9 corresponds to supersaturated gas. In Corollary 2.13, the particle excess  $N_n$  is absorbed by a condensate, i.e., one large occupation number. In Theorem 2.11, the particle excess is shared by a condensate of size  $x_{nr} < N_n$  and supersaturated gas. See [2] and [8, Section 7].

We conclude with an equivalent but more intrinsic formulation of Theorem 2.11. In Section 4 we shall see that  $G(z)$  extends to a function that is analytic in the slit plane  $\mathbb{C} \setminus [1, \infty)$ , and in addition the limit  $G(e^t) = \lim_{\varepsilon \searrow 0} G(e^t + i\varepsilon)$  exists for all  $t \geq 0$ . So the cumulant generating function  $\varphi(t) = \log G(e^t)$  extends to a function that is well-defined and smooth in a neighborhood of the origin in  $\mathbb{R}$ , and Eq. (2.13) stays valid for small positive  $t$ . We define

$$\begin{aligned} \Phi_n(t, \xi) &= -q(\xi) + n \operatorname{Re} \varphi(t) - (\mu n + N_n)t + t\xi \\ &= -q(\xi) + n \sum_{j=2}^{r-1} \kappa_j \frac{t^j}{j!} - t(N_n - \xi) + O(t^r). \end{aligned} \quad (2.18)$$

The asymptotic expansion holds for every order  $r$ . For  $N_n \gg N_n^*$  and sufficiently large  $n$ , the bivariate function  $\Phi_n(t, \xi)$  has exactly two critical points in  $(0, \frac{N_n}{n\sigma^2}) \times (a, \infty)$ . We label them as  $(t_n, \xi_n)$  and  $(t'_n, \xi'_n)$  with  $t_n < t'_n$ . Then, in the situation of Theorem 2.11, we have

$$\mathbb{P}(S_n = n\mu + N_n) \sim n e^{\Phi_n(t_n, \xi_n)}. \quad (2.19)$$

It is in this form that we prove the theorem. Let us explain how to recover the expression in terms of  $f_{nr}$ . We may solve for  $\nabla \Phi_n(t, \xi) = 0$  in two steps: (1) use  $\partial_t \Phi_n(t, \xi) = 0$  to express  $t = t(\xi)$  as a function of  $\xi$ , (2) plug the expression into  $\partial_\xi \Phi_n(t, \xi)$  to obtain an equation for  $\xi$ . This latter step breaks into the following two stages:

(2a) substitute the expression of  $t = t(\xi)$  into the expression for  $\Phi_n(t, \xi)$  so as to obtain a function  $\Phi_n(t(\xi), \xi)$  of  $\xi$  only;

(2b) set the derivative of  $\Phi_n(t(\xi), \xi)$  with respect to  $\xi$  to zero,

which is valid since

$$\frac{d}{d\xi} \Phi_n(t(\xi), \xi) = \frac{\partial \Phi_n}{\partial \xi}(t(\xi), \xi) + \frac{\partial \Phi_n}{\partial t}(t(\xi), \xi) \frac{dt}{d\xi}(t) = \frac{\partial \Phi_n}{\partial \xi}(t(\xi), \xi). \quad (2.20)$$

By the definition of the Cramér series, step (2a) gives

$$\Phi_n(t(\xi), \xi) = -q(\xi) - \frac{(N_n - \xi)^2}{2n\sigma^2} + \frac{(N_n - \xi)^3}{n^2} \sum_{j \geq 0} \lambda_j \left( \frac{N_n - \xi}{n} \right)^j \quad (2.21)$$

Truncation of the asymptotic expansion on the right-hand side gives precisely the function  $-f_{nr}(\xi)$ . For  $r$  large enough, step (2b) shows that we may approximate

$$\Phi_n(t_n, \xi_n) = -f_{nr}(x_{nr}) + o(1), \quad n\sigma^2 q''(\xi_n) = n\sigma^2 q''(x_{nr}) + o(1) \quad (2.22)$$

hence the equivalence of Eq. (2.19) with the expression from Theorem 2.11.

**2.3. Application to stretched exponential laws.** Here we explain how to recover five theorems by A. V. Nagaev [15] for stretched exponential variables. Let  $\alpha \in (0, 1)$ ,  $c > 0$ , and

$$p(k) = c \exp(-k^\alpha), \quad q(k) = k^\alpha - \log c \quad (k \in \mathbb{N}). \quad (2.23)$$

We need not check Assumption 2.3 since Theorem 4.4 for stretched exponential weights has already been proven in [12, Theorem 2.4.6].

**Lemma 2.14.** *The probability weights (2.23) satisfy Assumptions 2.1 and 2.2, and we have*

$$x_n^* = [\alpha(1-\alpha)n\sigma^2]^{1/(2-\alpha)}, \quad N_n^* = \frac{2-\alpha}{1-\alpha}x_n^*, \quad N_n^{**} = C_\alpha(n\sigma^2)^{1/(2-\alpha)}$$

with  $C_\alpha = (2-\alpha)(2-2\alpha)^{-(1-\alpha)/(2-\alpha)}$ . Moreover  $\sqrt{n\sigma^2}q'(N_n) \rightarrow 0$  if and only if  $N_n \gg n^{-1/(2-2\alpha)}$ .

The proof of the lemma is sketched in Appendix A. The critical scale  $n^{-1/(2-\alpha)}$  is explained by a simple scaling relation: for  $N_n = kn^{1/(2-\alpha)}$ , we have

$$f_n(yn^{1/(2-\alpha)}) = n^{\alpha/(2-\alpha)} \left( y^\alpha + \frac{(k-y)^2}{2\sigma^2} \right) - \log c. \quad (2.24)$$

A careful examination of the expressions in Theorem 2.10 shows that the first summand dominates if  $N_n \leq (1-\delta)N_n^{**}$  while the second dominates if  $N_n \geq (1+\delta)N_n^{**}$  for some  $\delta > 0$ . In particular, Theorem 2.9 extends to  $N_n \leq (1-\delta)N_n^{**}$ , which corresponds to Theorem 1 in [15]. Theorem 2.10 for  $N_n \sim N_n^{**}$  is Theorem 4 in [15]. For  $N_n \geq (1+\delta)N_n^{**}$ , we have

$$\mathbb{P}(S_n = n\mu + N_n) \sim \frac{1}{\sqrt{1 - n\sigma^2 q''(x_{nr})}} e^{-f_{nr}(x_{nr})}. \quad (2.25)$$

This regime can be divided into three cases:

- (a)  $N_n \ll n^{-1/(2-2\alpha)}$  corresponds to Theorem 2 in [15].
- (b) When  $N_n$  is of the order of  $n^{-1/(2-2\alpha)}$ , the corrections from the Cramér series are irrelevant and

$$\mathbb{P}(S_n = \mu n + N_n) \sim n \exp(-f_n(x_n)), \quad (2.26)$$

i.e., we may choose  $r = 0$ . This corresponds to Theorem 6 from the erratum [16], replacing Theorem 5 in the original article [15]. The statement actually extends to  $N_n \gg n^{-1/(3-3\alpha)}$ . Indeed  $q'(x_{nr}) \sim (N_n - x_{nr})/(n\sigma^2) = o(N_n/n)$  yields  $N_n - x_{nr} = (1 + o(1))n\sigma^2\alpha N_n^{\alpha-1}$  and

$$f_{nr}(x_{nr}) = f_n(x_{nr}) + O(nN_n^{-(3-3\alpha)}), \quad (2.27)$$

and one can check that  $f_n(x_{nr}) = f_n(x_n) + o(1)$  if  $N_n \gg n^{-1/(3-3\alpha)}$ .

- (c)  $N_n \gg n^{-1/(2-2\alpha)}$  corresponds to Theorem 3 in [15] and our Corollary 2.13.

**2.4. Application to logarithmic hazard functions.** Here we specialize to

$$p(k) = c \exp(-(\log k)^\beta), \quad q(k) = -\log c + (\log k)^\beta \quad (2.28)$$

with  $\beta > 2$ .<sup>1</sup>

<sup>1</sup>For  $\beta \in (1, 2]$ ,  $xq'(x) = \beta(\log x)^{\beta-1} \rightarrow \infty$  but the stronger condition  $xq'(x)/\log x \rightarrow \infty$  from Assumption 2.1(ii) fails. We suspect that this restriction is technical and could be lifted with more detailed estimates, but a proof or disproof is beyond this article's scope.

**Lemma 2.15.** *The weights (2.28) satisfy Assumptions 2.1– 2.3. Moreover*

$$N_n^* \sim 2x_n^* \sim 2\sqrt{2^{1-\beta}\beta n\sigma^2(\log n)^{\beta-1}}, \quad N_n^{**} \sim \sqrt{2n\sigma^2(\log n)^\beta},$$

and  $\sqrt{n\sigma^2}q'(N_n) \rightarrow 0$  if and only if  $N_n \gg \sqrt{n}(\log n)^{\beta-1}$ .

The lemma is proven Appendix A. Notice that unlike the stretched exponential case (Lemma 2.14),  $N_n^{**}$  is much larger than  $N_n^*$ . The scaling relation (2.24) is modified as follows: for  $N_n = k\sqrt{n\sigma^2(\log n)^{\beta-1}} \sim kx_n^*$ , we have

$$f_n(yx_n^*) = (\log x_n^*)^\beta + \beta(\log n)^{\beta-1} \left( \log y + \frac{(k-y)^2}{2} \right) + o((\log n)^{\beta-1}) \quad (2.29)$$

and  $(\log x_n^*)^\beta \sim (\log n)^\beta \sim (\log N_n)^\beta$ .

Our results may now be applied to obtain a sharp boundary for the big-jump domain.

**Theorem 2.16.** *Let  $p(k)$  be as in Eq. (2.28) with  $\beta > 2$  and  $N_n \gg \sqrt{n}$ . Then  $\mathbb{P}(S_n = \mu n + N_n) \sim n\mathbb{P}(X = N_n)$  if and only if  $N_n \gg \sqrt{n}(\log n)^{\beta-1}$ .*

The “if” part of the theorem is actually a special case of [4, Theorem 8.2] and as such not new. The “only if” part shows that the boundary derived in [4] is in fact sharp.

*Proof Theorem 2.16.* Let  $I_n := \sqrt{n\sigma^2}(\log n)^{\beta-1}$  and notice  $I_n \gg N_n^{**}$  for  $\beta > 2$ . Suppose that  $N_n \gg I_n$ . Then we have, in particular,  $N_n \gg N_n^{**}$ . Write  $N_n = \alpha_n N_n^{**}$  with  $\alpha_n \rightarrow \infty$ . Then

$$\frac{N_n^2/(2n\sigma^2)}{(\log N_n)^\beta} = \alpha_n^2 \left( 1 + o(1) \frac{\log \alpha_n}{\log n} \right)^{-\beta} \geq \frac{\alpha_n^2}{(\log \alpha_n)^\beta} \rightarrow \infty \quad (2.30)$$

and it follows from Corollary 2.13 that  $\mathbb{P}(S_n = n\mu + N_n) \sim n\mathbb{P}(X = n)$ , hence  $N_n \gg N_n^{**}$  is indeed a sufficient condition. In order to check that it is necessary, we treat the case  $N_n = O(I_n)$  with Theorems 2.9 and 2.10.

*Case 1:*  $N_n^* \ll N_n = O(I_n)$ . In Theorem 2.10 we obtain a lower bound by neglecting the first contribution and estimating  $1 - n\sigma^2|q''(x_{nr})| \leq 1$ . Combining with Proposition 2.12, we find

$$\begin{aligned} \mathbb{P}(S_n = n\mu + N_n) &\geq n \exp(-f_{nr}(x_{nr}) + o(1)) \\ &\geq n \exp\left(-q(N_n) + (1 + o(1))n\sigma^2 q'(N_n)^2 + o(1)\right) \end{aligned} \quad (2.31)$$

hence in view of  $N_n = O(I_n)$  and Lemma 2.15

$$\liminf_{n \rightarrow \infty} \frac{\mathbb{P}(S_n = n\mu + N_n)}{n\mathbb{P}(X = N_n)} \geq \liminf_{n \rightarrow \infty} \exp((1 + o(1))n\sigma^2 q'(N_n)^2) > 1. \quad (2.32)$$

*Case 2:*  $N_n = O(N_n^*)$ . Then we have in particular  $N_n = o(N_n^{**})$ . Write  $N_n = \alpha_n N_n^{**}$  with  $\alpha_n \rightarrow 0$ , then

$$\frac{N_n^2}{2n\sigma^2(\log N_n)^\beta} \leq \frac{\alpha_n^2(\log n)^\beta}{(\log \sqrt{n})^\beta} \rightarrow 0 \quad (2.33)$$

and Theorems 2.9 and 2.10 show

$$\log \frac{\mathbb{P}(S_n = n\mu + N_n)}{n\mathbb{P}(X = N_n)} \geq -\frac{N_n^2}{2n\sigma^2} + (\log N_n)^\beta + O(\log n) \rightarrow \infty. \quad (2.34)$$

□

### 3. PROOF STRATEGY

Here we explain the strategy for the proof of Theorems 2.9–2.11. We focus on the case  $N_n = o(n)$  and Theorem 2.10. Set  $m = \mu n + N_n$ . We start from the observation that  $\mathbb{P}(S_n = m)$  is equal to  $[z^m]G(z)^m$ , the coefficient of  $z^m$  in the expansion of  $G(z)^m$ , which in turn is given by contour integrals

$$[z^m]G(z)^n = \frac{1}{2\pi i} \oint \frac{G(z)^n}{z^m} \frac{dz}{z} = \frac{1}{2\pi i} \int_0^{2\pi i} e^{n\varphi(t)-mt} dt. \quad (3.1)$$

The contour integral can be taken over any circle centered at the origin with radius  $r \leq 1$ . A steepest descent ansatz would look for a point  $z_n$  such that  $z_n G'(z_n) = m/n = \mu + N_n/n$ , or  $\eta_n$  with  $\varphi'(\eta_n) = \mu + N_n/n$ , and then integrate over  $|z| = |z_n|$  (or  $\operatorname{Re} t = \operatorname{Re} \eta_n$ ). However in the regime  $m/n > G'(1) = \mu$  that we investigate there is no such point, and instead we follow an approach that is in the spirit of singularity analysis [11] but with several novel ingredients. Crucially, the generating function  $G(z)$  *does not* fall into the class of functions which Flajolet and Sedgewick call “amenable to singularity analysis” [11][Chapter VI].

**Step 1: Analytic extension to slit plane.** Observe that  $G(z)$  has an analytic extension to the slit plane  $\mathbb{C} \setminus [1, \infty)$ . This is proven with the help of the Lindelöf integral [13, 10], see Proposition 4.1. The key ingredient here is that  $p(\xi)$  is analytic in a complex half-plane containing the integers  $k \in \mathbb{N}$  and growth slower than  $\exp((\pi - \varepsilon)|\xi|)$  as  $\xi \rightarrow \infty$ .

**Step 2: Behavior near the dominant singularity and along the slit.** The Lindelöf integral actually shows that the analytic extension  $G(z)$  has well-defined limits as  $z$  approaches the slit  $[1, \infty)$  from above or below, i.e., the limits  $\lim_{\varepsilon \searrow 0} G(e^t + i\varepsilon)$  and  $\lim_{\varepsilon \searrow 0} G(e^t - i\varepsilon)$  exist for all  $t \in \mathbb{R}$ . Moreover the imaginary part along the slit is given by a Bromwich integral,

$$\lim_{\varepsilon \searrow 0} \operatorname{Im} G(e^t + i\varepsilon) = \frac{1}{2i} \int_{1/2-i\infty}^{1/2+i\infty} e^{t\xi} p(\xi) d\xi \quad (t \geq 0). \quad (3.2)$$

The line of integration  $\operatorname{Re} \xi = 1/2$  can be replaced by any other line  $\operatorname{Re} \xi = x > 0$ . As  $t \searrow 0$ , the imaginary part vanishes faster than any power of  $t$ , whereas the real part can be approximated to arbitrarily high order by a Taylor polynomial.

**Step 3: Contour integrals.** We may now deform the contour of integration: in the  $z$ -plane, we replace the circle of radius 1 by a Hankel-type contour consisting of a circle of radius  $e^\varepsilon$  and a piece hugging the segment  $[1, \varepsilon)$ , see Figure 2. In the  $t$ -plane, we replace the vertical segment joining 0 and  $2\pi i$  by the three other sides of the rectangle with corners 0,  $\varepsilon$ ,  $\varepsilon + 2\pi i$ ,  $2\pi i$ . This yields

$$[z^m]G(z)^n = \frac{1}{2\pi i} \left( \int_0^\varepsilon e^{n\varphi(t)-mt} dt + \int_0^{2\pi} e^{n\varphi(\varepsilon+i\theta)-m(\varepsilon+i\theta)} i d\theta - \int_0^\varepsilon e^{n\varphi(t+2\pi i)-m(t+2\pi i)} dt \right). \quad (3.3)$$

We focus on  $N_n = o(n)$  and choose  $\varepsilon = \eta_n$  as the solution of

$$\operatorname{Re} \varphi'(\eta_n) = \frac{m}{n} = \mu + \frac{N_n}{n}. \quad (3.4)$$

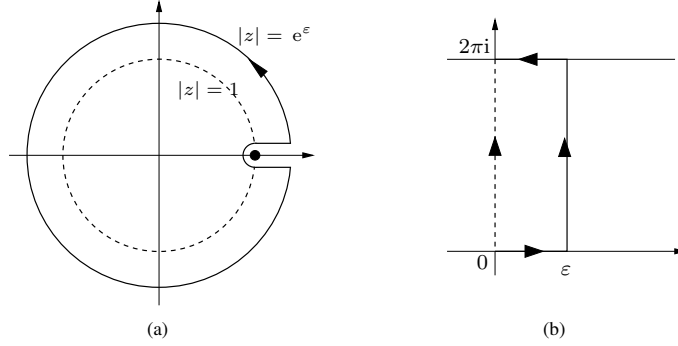


FIGURE 2. Contour integrals in the  $z$ -plane (a) and in the  $t$ -plane (b). Dotted lines: Eq. (3.1). Solid lines: deformed contours in Eq. (3.3). Recall that  $z = e^t$ . Later  $\varepsilon = \eta_n$  will be chosen in a judicious way. For  $N_n \ll N_n^{**}$ , the dominant contribution should come from the horizontal pieces of the deformed contour in the  $t$ -plane. For  $N_n \gg N_n^{**}$ , the dominant contribution should instead be from the vertical line.

Notice  $\eta_n \sim N_n/(n\sigma^2)$ . Using the identities

$$G(\bar{z}) = \overline{G(z)}, \quad \forall t \geq 0: \varphi(t + 2\pi i) = \overline{\varphi(t)}, \quad (3.5)$$

Eq. (3.3) becomes

$$[z^m]G(z)^n = H_n + V_n \quad (3.6)$$

with

$$\begin{aligned} H_n &= \frac{1}{\pi} \int_0^{\eta_n} e^{n\operatorname{Re} \varphi(t) - mt} \sin(n\operatorname{Im} \varphi(t)) dt \\ V_n &= \frac{1}{\pi} \int_0^\pi e^{n\operatorname{Re} \varphi(\eta_n + i\theta) - m\eta_n} \cos(n\operatorname{Im} \varphi(\eta_n + i\theta)) d\theta. \end{aligned} \quad (3.7)$$

Standard arguments show that the dominant contribution to  $V_n$  come from small  $\theta$ . Since  $\operatorname{Im} G(e^t) \rightarrow 0$  faster than any power of  $t$  as  $t \searrow 0$  and

$$\operatorname{Im} G(e^t) = \operatorname{Im} e^{\varphi(t)} = e^{\operatorname{Re} \varphi(t)} \operatorname{Im} \varphi(t) \sim \operatorname{Im} \varphi(t), \quad (3.8)$$

we may drop the trigonometric functions from Eq. (3.7) and find

$$\begin{aligned} H_n &\sim \frac{n}{\pi} \int_0^{\eta_n} e^{n\operatorname{Re} \varphi(t) - mt} \operatorname{Im} G(e^t) dt, \\ V_n &\sim \frac{1}{\pi} \int_0^\pi e^{n\operatorname{Re} \varphi(\eta_n + i\theta) - m\eta_n} d\theta. \end{aligned} \quad (3.9)$$

The vertical contribution is evaluated with the help of a Gaussian approximation around  $\theta = 0$ , which yields

$$V_n \sim \frac{1}{\sqrt{2\pi n \operatorname{Re} \varphi''(\eta_n)}} e^{n\operatorname{Re} \varphi(\eta_n) - m\eta_n}. \quad (3.10)$$

With Definition 2.8, we recognize in Eq. (3.10) the asymptotic expression from Theorem 2.9 and obtain

$$V_n \sim \frac{1}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{N_n^2}{2n\sigma^2} + \frac{N_n^3}{n^2} \sum_{j=0}^{r-1} \lambda_j \left(\frac{N_n}{n}\right)^j\right). \quad (3.11)$$

The evaluation of  $H_n$  is more involved. As a preliminary step, we express  $\text{Im } G(e^t)$  through the Bromwich integral and find

$$H_n \sim \frac{n}{2\pi i} \int_0^{\eta_n} \left( \int_{1/2-i\infty}^{1/2+i\infty} e^{\Phi_n(t,\xi)} d\xi \right) dt \quad (3.12)$$

with  $\Phi_n$  as in Eq. (2.18).

**Step 4: Critical points of  $\Phi_n(t, \xi)$ .** In order to apply a Gaussian approximation to the bivariate integral (3.12), we look for a critical points  $(t_n, \xi_n)$  of  $\Phi_n$  with  $t_n \in (0, \eta_n)$  and  $\xi_n \in (0, \infty)$ . The gradient  $\nabla \Phi_n(t_n, \xi_n)$  vanishes if and only if

$$\begin{cases} t_n = q'(\xi_n), \\ \text{Re}(\varphi'(t_n) - \varphi'(0)) = \frac{N_n - \xi_n}{n}. \end{cases} \quad (3.13)$$

Since  $\text{Re } \varphi'(t) = \mu + \sigma^2 t + O(t^2)$  as  $t \searrow 0$ , Eq. (3.13) implies

$$q'(\xi_n) \sim \frac{N_n - \xi_n}{n\sigma^2}. \quad (3.14)$$

We recognize the equation for the critical points of  $f_n$ . Lemma 2.4 suggests the following: for  $N_n \ll N_n^*$ , there should be no critical point, for  $N_n^* \ll N_n \ll n$ , there should be two. Let us focus on the latter case and label the critical points as  $(t_n, \xi_n)$  and  $(t'_n, \xi'_n)$  with  $t_n < t'_n$ . In view of Lemmas 2.4 and 2.7, we expect  $\xi_n \approx x_n$  and  $\xi'_n \approx x'_n$ , hence

$$t_n \sim q'(N_n), \quad \xi_n = N_n + O(N_n^*) \quad (3.15)$$

and  $\xi'_n < x_n^* < \xi_n$ . The Hessian of  $\Phi_n$  is

$$\text{Hess } \Phi_n(t, \xi) = \begin{pmatrix} n \text{Re } \varphi''(t) & 1 \\ 1 & -q''(\xi) \end{pmatrix}. \quad (3.16)$$

Using again Lemma 2.7 and  $\xi_n \approx x_n$ , we expect

$$\det \text{Hess } \Phi_n(t_n, \xi_n) = -1 - (1 + o(1))n\sigma^2 q''(\xi_n) = -1 + O\left(\frac{N_n^*}{N_n}\right) < 0 \quad (3.17)$$

thus  $(t_n, \xi_n)$  is a saddle point. (More precisely, it is a saddle point of  $\text{Re } \Phi_n$ , but the abuse of terminology is natural and not problematic in our context.) On the other hand  $\xi'_n \approx x'_n < x_n^*$  with  $1 + n\sigma^2 q'(x_n^*) = 0$  by definition of  $x_n^*$ , so we expect

$$\det \text{Hess } \Phi_n(t'_n, \xi'_n) = -1 - (1 + o(1))n\sigma^2 q''(\xi'_n) > 0. \quad (3.18)$$

**Step 5: Gaussian approximation for  $H_n$ .** In order to evaluate the double integral in Eq. (3.12), we use a good change of variables and a Gaussian approximation. Let  $\xi(t)$  be the solution of  $q'(\xi) = t$ , so that  $\partial_\xi \Phi_n(t, \xi) = 0$  if and only if  $\xi = \xi(t)$ . It is convenient to deform the contour and integrate along  $\text{Re } \xi = \xi(t)$  instead of  $\text{Re } \xi = 1/2$ . The integral becomes

$$H_n \sim \frac{n}{2\pi} \int_0^{\eta_n} \left( \int_{-\infty}^{\infty} e^{\Phi_n(t, \xi(t)+is)} ds \right) dt. \quad (3.19)$$

A straightforward computation shows that  $F_n(t, s) = S_n(t, \xi(t) + is)$ , a function of two real variables  $t, s$ , has a critical point at  $(t_n, 0)$  with positive definite Hessian

$$\text{Hess } F_n(t_n, 0) = \begin{pmatrix} \beta_n & 0 \\ 0 & -\partial_\xi^2 \Phi_n(t_n, \xi_n) \end{pmatrix}, \quad \beta_n = \frac{\det \text{Hess } \Phi_n(t_n, \xi_n)}{\partial_\xi^2 \Phi_n(t_n, \xi_n)}, \quad (3.20)$$

see Lemma B.1.  $F_n : (0, \eta_n) \times \mathbb{R} \rightarrow \mathbb{C}$  has another critical point at  $(t'_n, 0)$ , with negative determinant of the Hessian; later we show that it does not contribute to the integral. The evaluation of  $H_n$  is concluded by replacing the double integral (3.19) by the integral of the Gaussian approximation around  $(t_n, 0)$  which yields

$$H_n \sim \frac{n}{2\pi} \sqrt{\frac{(2\pi)^2}{\det \text{Hess } F_n(t_n, 0)}} e^{F_n(t_n, 0)} = \frac{n}{\sqrt{|\det \text{Hess } \Phi_n(t_n, \xi_n)|}} e^{\Phi_n(t_n, \xi_n)}. \quad (3.21)$$

The argument leading leading to Eq. (2.22) and Eq. (3.17) show

$$H_n \sim \frac{n}{\sqrt{1 - n\sigma^2 q''(x_{nr})}} e^{-f_{nr}(x_{nr})}. \quad (3.22)$$

#### 4. ANALYTIC CONTINUATION. LINDELÖF AND BROMWICH INTEGRALS

Here we take care of steps 1 and 2, starting from Assumption 2.2. For concreteness' sake we write down the results for  $b = 1/2$ ; they apply for general  $b$  with straight-forward modifications. Define

$$\Lambda(w) = -\frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} p(\xi) w^\xi \frac{\pi}{\sin \pi \xi} d\xi \quad (w \in \mathbb{C} \setminus (-\infty, 0]), \quad (4.1)$$

the Lindelöf integral with symbol  $p(\xi)$ .

**Proposition 4.1** ([13]).

- (a)  $\Lambda(w)$  is analytic in the slit plane  $\mathbb{C} \setminus (-\infty, -1]$ .
- (b) In the unit disk  $|w| \leq 1$ ,

$$\Lambda(w) = \sum_{k=1}^{\infty} p(k) (-w)^k = G(-w).$$

A detailed proof and many additional properties of  $\Lambda(w)$  can be found in [10]. Proposition 4.1 shows right away that  $G(z) = \Lambda(-z)$  has an analytic continuation from the unit disk to the open slit plane  $\mathbb{C} \setminus [1, \infty)$ . We use the same letter  $G(z)$  for the analytic continuation, and set  $\varphi(t) = \text{Log } G(e^t)$  with  $\text{Log}$  the principal branch of the logarithm, and  $\text{Im } t \in [0, 2\pi)$ . We prove the following additional properties of  $G(z)$  and  $\varphi(t)$ .

**Theorem 4.2.**

- (a) The boundary value  $G(e^t) = \lim_{\varepsilon \searrow 0} G(e^t + i\varepsilon)$  exists for all  $t \in \mathbb{R}$  and is a smooth function of  $t \in \mathbb{R}$ .
- (b) The imaginary part  $\text{Im } G(e^t)$ ,  $t \geq 0$  is given by the Bromwich integral (3.2).
- (c)  $\varphi(t)$  is well-defined and smooth in a neighbourhood of the origin; the derivatives  $\kappa_j = \varphi^{(j)}(0)$  are real. As  $t \rightarrow 0$  in the strip  $\text{Im } t \in [0, 2\pi)$ , we have

$$\varphi(t) = \text{Log } G(e^t) = \sum_{j=1}^r \kappa_j \frac{t^j}{j!} + O(t^{r+1}).$$

to arbitrarily high order  $r$ .



In (c)  $z = e^t$  is allowed to approach the slit  $[1, \infty)$  as fast as we like; we may even take  $t$  real. Because the coefficients  $\kappa_j$  are real, we find in particular that  $\text{Im} G(e^t)$  vanishes faster than any power of  $t$  as  $t \rightarrow 0$ ,  $t \in \mathbb{R}$ .

*Proof of Theorem 4.2.* For  $u \in \mathbb{C}$  in the closed strip  $\text{Im } u \in [-\pi, \pi]$ , define

$$L(u) = -\frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} p(\xi) e^{\xi u} \frac{\pi}{\sin \pi \xi} d\xi. \quad (4.2)$$

When  $\text{Im } u$  is in the open strip  $\text{Im } u \in (-\pi, \pi)$ , we have  $w = e^u \in \mathbb{C} \setminus (-\infty, 0]$  and  $L(u) = \Lambda(e^u)$ . Along the vertical line  $\text{Re } \xi = 1/2$ , we have

$$\left| \frac{\exp(\xi u)}{\sin(\pi \xi)} \right| = 2e^{\text{Re } u/2} \frac{\exp(-s \text{Im } u)}{\exp(\pi s) + \exp(-\pi s)} \leq 2e^{\text{Re } u/2} \quad (\xi = \frac{1}{2} + is) \quad (4.3)$$

By Assumption 2.2(ii), since  $\xi^k p(\xi)$  is integrable along  $\text{Re } \xi = 1/2$ . Eq. (4.3) then shows that the integral defining  $L(u)$  is absolutely convergent, and it stays absolutely convergent if we replace the symbol  $p(\xi)$  by  $\xi^k p(\xi)$ . Standard arguments for parameter-dependent integrals then show that  $L(u)$  is continuous on the closed strip, differentiable in the open strip, and we may exchange differentiation, limits, and integration, which shows that the restriction of  $L$  to the boundaries  $\text{Im } u = \pm\pi$  yield smooth functions.

When  $z \rightarrow e^t \in [1, \infty)$  along  $\text{Im } z > 0$ , we have  $w = -z \rightarrow -e^t$  along  $\text{Im } w < 0$ . Thus we may write  $w = e^u$  with  $\text{Re } u \rightarrow t$  and  $\text{Im } u = \arg w \searrow -\pi$ . Therefore

$$\lim_{\varepsilon \searrow 0} G(e^t + i\varepsilon) = L(t - i\pi) = i \int_{-\infty}^{\infty} p(\frac{1}{2} + is) e^{(1/2+is)t} \frac{\exp(\pi s)}{\exp(\pi s) + \exp(-\pi s)} ds. \quad (4.4)$$

This proves the existence of the limit and, in view of the above mentioned properties of  $L(u)$ , the smoothness as a function of  $t$ . The complex conjugate is

$$\begin{aligned} -i \int_{-\infty}^{\infty} p(\frac{1}{2} - is) e^{(1/2-is)t} \frac{\exp(\pi s)}{\exp(\pi s) + \exp(-\pi s)} ds \\ = i \int_{-\infty}^{\infty} p(\frac{1}{2} + is) e^{(1/2+is)t} \frac{\exp(-\pi s)}{\exp(\pi s) + \exp(-\pi s)} ds. \end{aligned} \quad (4.5)$$

Therefore

$$\text{Im } G(e^t) = \frac{1}{2} \int_{-\infty}^{\infty} p(\frac{1}{2} + is) e^{(1/2+is)t} ds = \frac{1}{2i} \int_{1/2-i\infty}^{1/2+i\infty} p(\xi) e^{t\xi} d\xi.$$

This proves (b). For (c), consider first real  $t \in \mathbb{R}$ . We have already checked (a) hence  $G(e^t)$  is in  $C^\infty(\mathbb{R})$ . It is real and strictly positive for  $t \leq 0$  (this follows from the series representation and  $p(k) > 0$ ), and non-zero though possibly complex-valued for sufficiently small  $t > 0$ . Therefore  $\varphi(t) = \log G(e^t)$  is well-defined and smooth in some interval  $(-\infty, \delta)$ ,  $\delta > 0$ , and real-valued for  $t \leq 0$ . In particular, the derivatives  $\kappa_j = \varphi^{(j)}(0)$  exist and are real, and  $\varphi(t)$  can be approximated to arbitrarily high order by Taylor polynomials. The extension to complex  $t$ ,  $\text{Im } t \in [0, 2\pi)$ , follows again from the smoothness of  $L(u)$  in the closed strip  $\text{Im } u \in [-\pi, \pi]$ .  $\square$

Theorem 4.2(b) has an interesting consequence. Eq. (3.2) is, up to a factor  $\pi$ , the formula for the inverse Laplace transform, therefore

$$p(\lambda) = \pi \int_0^\infty e^{-t\lambda} \operatorname{Im} G(e^t) dt \quad (\operatorname{Re} \lambda > 0). \quad (4.6)$$

In the special case of stretched exponential weights, we can draw on an extensive literature as  $\exp(-\lambda^\alpha)$  is known to be the Laplace transform of a probability density, an  $\alpha$ -stable law. For  $\alpha = 1/2$  [5]

$$\operatorname{Im} G(e^t) = \frac{c\sqrt{\pi}}{2t^{3/2}} e^{-1/(4t)} \quad (t \geq 0). \quad (4.7)$$

For general  $\alpha \in (0, 1)$ , we have instead [12, Theorem 2.4.6]

$$\operatorname{Im} G(e^t) \sim \frac{c}{2} \sqrt{\frac{2\pi}{(1-\alpha)\alpha^{-1/(1-\alpha)}}} \frac{\exp(-(1-\alpha)\left(\frac{\alpha}{t}\right)^{\alpha/(1-\alpha)})}{t^{(2-\alpha)/(2-2\alpha)}} \quad (4.8)$$

as  $t \searrow 0$ . This is proven in [12] by applying a steepest descent approach to the Bromwich integral. For general weights, Eq. (4.8) is generalized as follows.

Assume that  $t < \lim_{x \searrow a} |q''(x)|$ . By Assumption 2.1,  $q''$  is strictly increasing and negative on  $(a, \infty)$ . By Eq. (2.6), we have  $q''(x) \rightarrow 0$  as  $x \rightarrow \infty$ . Consequently there exists a uniquely defined  $\xi(t)$  that solves  $q'(\xi(t)) = t$ . We define

$$\psi(t) = t\xi(t) - q(\xi(t)) \quad (4.9)$$

and note the relations

$$\psi'(t) = \xi(t), \quad \psi''(t) = \frac{1}{q''(\xi(t))}, \quad (4.10)$$

so  $\psi(t)$  is monotone increasing and strictly concave. Since  $-\psi(-t)$  is the Legendre transform of the convex function  $-q(x)$ , it comes as no surprise that Assumption 2.1 on large  $x$  translates into information on small  $t$ .

**Lemma 4.3.** *The following holds:*

- (a)  $\lim_{t \searrow 0} t\psi'(t)/\log t = \infty$ .
- (b)  $-\psi''(t) \geq c \frac{\psi'(t)}{t}$  for some  $c > 0$  and all sufficiently small  $t > 0$ .
- (c)  $0 \leq \psi^{(3)}(t) \leq C \frac{|\psi''(t)|}{t}$  for some  $C > 0$  and all sufficiently small  $t > 0$ .

The lemma has been proven in [17, Lemma 2.2].

**Theorem 4.4.** *As  $t \searrow 0$ ,*

$$\operatorname{Im} G(e^t) \sim \frac{1}{2} \sqrt{2\pi |\psi''(t)|} e^{\psi(t)}.$$

*Proof.* By Theorem 4.2(b) we may start from the Bromwich representation of  $\operatorname{Im} G(e^t)$ . The analyticity of  $q(\xi)$  allows us to replace the contour  $\operatorname{Re} \xi = 1/2$  by  $\Gamma = \Gamma_1 \cup \Gamma_2$  where

$$\Gamma_1 = \{\xi \in \mathbb{C} \mid |\xi| = \xi(t), \operatorname{Re} \xi \geq 1/2\}, \quad \Gamma_2 = \{\xi \in \mathbb{C} \mid |\xi| > \xi(t), \operatorname{Re} \xi = 1/2\}. \quad (4.11)$$

To lighten notation set  $r = \xi(t)$  and suppress the  $t$ - and  $r$ -dependence from the notation. Let  $\theta_0 = \arcsin(1/2r)$  and notice  $\theta_0 \nearrow \pi/2$  as  $r \rightarrow \infty$  ( $t \searrow 0$ ). For small

$\theta$  we have

$$\begin{aligned} t r e^{i\theta} - q(r e^{i\theta}) &= \psi(t) - \frac{1}{2} r^2 q''(r) (e^{i\theta} - 1)^2 + O(r^3 q^{(3)}(r) (e^{i\theta} - 1)^2) \\ &= \psi(t) - \frac{1}{2} r^2 |q''(r)| (\theta^2 + O(\theta^3)) \end{aligned} \quad (4.12)$$

The estimate is uniform in  $r = \xi(t)$  by Assumption 2.3(iii). Let  $\varepsilon(r) \searrow 0$  with  $\varepsilon(r)^2 r^2 q''(r) / \log r \rightarrow \infty$  (this is possible by Assumption 2.1), then

$$\frac{1}{2i} \int_{-\varepsilon(r)}^{\varepsilon(r)} \exp(tr e^{i\theta} - q(r e^{i\theta})) i r e^{i\theta} d\theta \sim \frac{1}{2} \sqrt{2\pi |\psi''(t)|} e^{\psi(t)}. \quad (4.13)$$

As we veer away from  $r = \xi(t)$  along  $\Gamma_1$ , the real part of  $\xi t - q(\xi)$  decreases. Indeed for  $\theta \in (0, \pi/2)$

$$\begin{aligned} \frac{d}{d\theta} \operatorname{Re}(tr e^{i\theta} - q(r e^{i\theta})) &= \operatorname{Re}(i\xi(t - q'(\xi))) \Big|_{\xi=r e^{i\theta}} \\ &= -\operatorname{Im}(\xi q'(r) - \xi q'(\xi)) \Big|_{\xi=r e^{i\theta}} \leq 0. \end{aligned} \quad (4.14)$$

At the very end we have used Assumption 2.3(ii) It follows that

$$\begin{aligned} \left| \frac{1}{2i} \int_{\varepsilon(r)}^{\theta_0} \exp(tr e^{i\theta} - q(r e^{i\theta})) i r e^{i\theta} d\theta \right| \\ \leq \frac{r\pi}{4} \exp\left(\psi(t) - \frac{1}{2} r^2 |q''(r)| \varepsilon(r) (1 + o(1))\right) = o\left(\exp(\psi(t))\right) \end{aligned}$$

Taking complex conjugates, we obtain a similar estimate for the integral from  $-\theta_0$  to  $-\varepsilon(r)$ . Together with (4.13) we obtain

$$\frac{1}{2i} \int_{\Gamma_1} e^{t\xi - q(\xi)} d\xi \sim \frac{1}{2} \sqrt{2\pi |\psi''(t)|} e^{\psi(t)}. \quad (4.15)$$

It remains to estimate the contribution from  $\Gamma_2$ . Because of the monotonicity (4.14) we have

$$\operatorname{Re}(t\xi_0 - q(\xi_0)) \leq \psi(t) - \frac{1}{2} r^2 |q''(r)| \varepsilon(r)^2 (1 + o(1)) \quad (4.16)$$

Assumption 2.3(ii) ensures that

$$\left| \int_{\Gamma_2} e^{t\xi - q(\xi)} d\xi \right| \leq \exp\left(\operatorname{Re}[t\xi_0 - q(\xi_0)] + O(\log r)\right) = o\left(\exp(\psi(t))\right). \quad (4.17)$$

□

## 5. CRITICAL POINT AND GAUSSIAN APPROXIMATION

Here we prove Lemmas 2.4-2.7 and Proposition 2.12 and we address steps 4 and 5 of the proof strategy.

### 5.1. Variational analysis of $f_n(x)$ . Critical scales.

*Proof of Lemma 2.4.* We treat the case  $a = 0$ . Under Assumption 2.1(i),  $q'$  is strictly convex and decreasing. Therefore

$$\begin{aligned} f'_n(x) &= q'(x) - \frac{N_n - x}{n\sigma^2} \\ &\geq q'(x_n^*) + q''(x_n^*)(x - x_n^*) - \frac{N_n - x}{n\sigma^2} = \frac{N_n^* - N_n}{n\sigma^2} \end{aligned} \quad (5.1)$$

with equality if and only if  $x = x_n^*$ . If  $N_n < N_n^*$ , we obtain  $f_n'(x) > 0$  on  $(0, \infty)$  and parts (a) and (b) of the lemma follow right away.

If  $N_n > N_n^*$ , then  $f_n'(x_n^*) = (N_n^* - N_n)/(n\sigma^2) < 0$  and  $\lim_{x \rightarrow \infty} f_n'(x) = \infty$ , so by the intermediate value theorem  $f_n'$  has at least one zero in  $(x_n^*, \infty)$ . On the other hand

$$f_n''(x) = q''(x) + (n\sigma^2)^{-1} = q''(x) - q''(x_n^*) > 0 \text{ on } (x_n^*, \infty) \quad (5.2)$$

so  $f_n'$  is strictly increasing and  $f_n'$  has exactly one zero  $x_n$  in  $(x_n^*, \infty)$ , moreover  $f_n(x_n) = \min_{(x_n^*, \infty)} f_n$ . Since  $q'(x_n) > 0$  by Assumption 2.1 and  $q'(x_n) = (N_n - x_n)/(n\sigma^2)$ , we must have  $x_n < N_n$ . This proves the first part of (c).

If in addition to  $N_n > N_n^*$ , we have  $\limsup_{n \rightarrow \infty} N_n/(n\sigma^2) < \lim_{x \searrow a} q'(x)$ , then  $\lim_{x \searrow a} f_n'(x) > 0$ . We have already observed that  $f_n'(x_n^*) < 0$ . By the intermediate value theorem,  $f_n'$  has at least one zero  $x_n'$  in  $(0, x_n^*)$ . Since  $f_n''(x) = q''(x) - q''(x_n^*) < 0$  on  $(0, x_n^*)$ , the zero is unique and corresponds to maximizer. This completes the proof of (c). The proof of (d) is similar to (c) and therefore omitted.  $\square$

*Proof of Lemma 2.5.* We treat the case  $a = 0$ . Write  $f_n(x) = I_n(x, N_n)$  with  $I_n(x, y) = q(x) + [y - x]^2/[2n\sigma^2]$ . For  $y > N_n^*$ , let  $x_n(y) > x_n^* > x_n'(y)$  be the solutions of  $\partial_x I_n(x, y) = 0$ , with  $x_n'(y)$  well-defined for  $N_n \leq n\sigma^2 \sup q'$  only. Notice that  $x \mapsto I_n(x, y)$  is increasing in  $(0, x_n'(y))$ , decreasing in  $(x_n'(y), x_n(y))$ , and increasing in  $(x_n(y), \infty)$ . We have

$$\frac{d}{dy} [I_n(x_n(y), y) - I(0, y)] = \frac{y - x_n(y)}{n\sigma^2} - \frac{y}{n\sigma^2} = -\frac{x_n(y)}{n\sigma^2} < 0. \quad (5.3)$$

As  $y \searrow N_n^*$  at fixed  $n$ , a careful examination of the proof of Lemma 2.4 shows  $x_n(y) \searrow x_n^*$  and  $x_n'(y) \nearrow x_n^*(y)$ , hence  $I_n(x_n(y), y) \rightarrow I_n(x_n^*, N_n^*)$ . But  $x \mapsto I_n(x, N_n^*)$  is strictly increasing on  $(0, \infty)$  because for  $N_n = N_n^*$ ,  $\partial_x I_n(\cdot, N_n) = f_n'(x) \geq 0$  by Eq. (5.1), hence  $I_n(x_n^*, N_n^*) > I_n(0, N_n^*)$  and by continuity

$$\lim_{y \searrow N_n^*} [I_n(x_n(y), y) - I(0, y)] > 0. \quad (5.4)$$

Assumption 2.1 implies that  $q(y) = o(y)$  as  $y \rightarrow \infty$ . It follows that

$$\lim_{y \rightarrow \infty} [I_n(y, y) - I(0, y)] = \lim_{y \rightarrow \infty} [q(y) - \frac{y^2}{2n\sigma^2}] = -\infty. \quad (5.5)$$

Eqs. (5.3)–(5.5) guarantee the existence and uniqueness of a solution  $y = N_n^{**}$  to the equation  $I_n(x_n(y), y) = I(0, y)$ , and (a)–(c) follow with the observation  $f_n(x_n) - f_n(0) = [I_n(x_n(y), y) - I_n(0, y)]|_{y=N_n}$ .  $\square$

*Proof of Lemma 2.6.* As noted after Assumption 2.1, we have  $\lim_{x \rightarrow \infty} x^2 q''(x) = -\infty$ , moreover from the definition (2.10) of  $x_n^*$  and the observation  $x_n^* \rightarrow \infty$  we get

$$1 = \lim_{n \rightarrow \infty} n\sigma^2 |q''(x_n^*)| \gg \frac{n\sigma^2}{(x_n^*)^2} \quad (5.6)$$

hence  $x_n^* \gg \sqrt{n}$ . The inequality  $x_n^* < N_n^*$  follows from the definition 2.10 of  $N_n^*$  and the positivity of  $q'$ . The inequality  $N_n^* < N_n^{**}$  holds true by definition of  $N_n^{**}$ . By Assumption 2.1(v) there exists  $C > 0$  such that  $q(x) \leq Cx^\alpha$  for all sufficiently large  $x$ . Fix  $C' > C$  and  $N_n \geq (2C'n\sigma^2)^{1/(2-\alpha)}$ . Then for large  $n$ ,

$$\frac{q(N_n)}{N_n/(2n\sigma^2)} \leq 2n\sigma^2 C N_n^{\alpha-2} \leq \frac{C}{C'} < 1. \quad (5.7)$$

Write  $C/C' = 1 - \varepsilon$ . It follows that  $f_n(N_n) = q(N_n) < (1 - \varepsilon)\frac{N_n}{n\sigma^2} \leq (1 - \varepsilon)(1 + o(1))f_n(a)$ , and a fortiori  $\min f_n \leq f_n(N_n) < f_n(a)$ , which shows  $N_n \geq N_n^{**}$ . This proves  $N_n^{**} = O(n^{1/(2-\alpha)})$  and completes the proof of the first part of the lemma.

Next suppose by contradiction that  $N_n^*/x_n^* \rightarrow 1$ . Then by Eq. (2.10) we must have  $n\sigma^2 q'(x_n^*)/x_n^* \rightarrow 0$ . Since  $x_n^* \rightarrow \infty$ , Assumption 2.1 yields  $n\sigma^2 q''(x_n^*) \rightarrow 0$ , contradicting the definition of  $x_n^*$ . Similarly, the assumption  $N_n^*/x_n^* \rightarrow \infty$  leads to  $n\sigma^2 q''(x_n^*) \rightarrow \infty$ , contradicting again the definition of  $x_n^*$ . So  $N_n^*/x_n^*$  stays bounded away from 1 and from  $\infty$ .  $\square$

*Proof of Lemma 2.7.* Remember  $q'(x_n^*) = [N_n^* - x_n^*]/[n\sigma^2]$  by definition of  $N_n^*$ , and  $q'(x_n) = [N_n - x_n]/[n\sigma^2]$  by definition of  $x_n$ . Since  $q'$  is strictly decreasing we deduce

$$\frac{N_n - x_n}{n\sigma^2} = q'(x_n) < q'(x_n^*) = \frac{N_n^* - x_n^*}{n\sigma^2} < \frac{N_n^*}{2\sigma^2}, \quad (5.8)$$

so  $N_n - x_n \leq N_n^*$  (see Figure 3). Since  $q'$  is strictly convex, we have  $q'(x_n^*) > q'(x_n) + q''(x_n)(x_n^* - x_n)$  hence

$$q''(x_n) > \frac{q'(x_n) - q'(x_n^*)}{x_n - x_n^*} = \frac{(N_n - x_n) - (N_n^* - x_n^*)}{n\sigma^2(x_n - x_n^*)} = \frac{O(N_n^*)}{n\sigma^2(N_n + O(N_n^*))} \quad (5.9)$$

We also know that  $q''(x_n) < 0$ , so we obtain

$$f_n''(x_n) = q''(x_n) + \frac{1}{n\sigma^2} = \frac{1}{n\sigma^2} \left(1 + O\left(\frac{N_n^*}{N_n}\right)\right). \quad (5.10)$$

$\square$

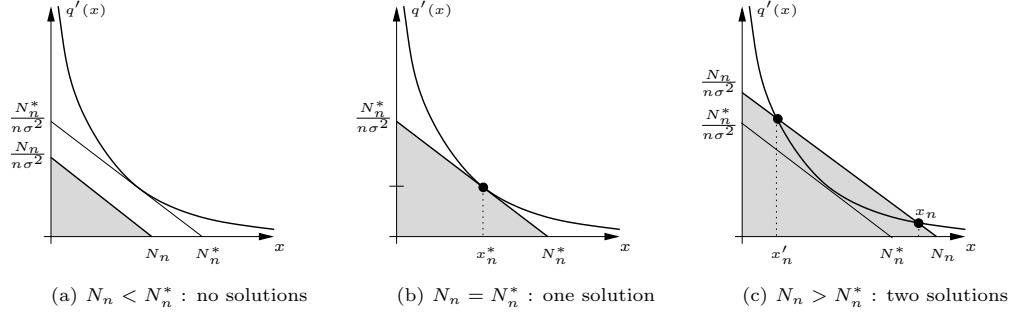


FIGURE 3. Solutions to  $q'(x) = (N_n - x)/(n\sigma^2)$  (=critical points of  $f_n$ ) as in Lemma 2.7.

For the proof of Theorem 2.9 in the case  $N_n \sim N_n^*$  we need the following.

**Lemma 5.1.** *Let  $N_n = N_n^*$ . We have*

$$f_n(x_n^*) - \frac{(N_n^*)^2}{2n\sigma^2} \geq \varepsilon \frac{(N_n^*)^2}{2n\sigma^2}. \quad (5.11)$$

for some  $\varepsilon > 0$  and all sufficiently large  $n$ ,

*Proof.* Assume  $a = 0$  and  $q(0) = 0$ . By Lemma 2.5 we already know  $f_n(x_n^*) - f_n(0) > 0$ , we prove  $f_n(x_n^*) - f_n(0) > \varepsilon f_n(0)$ . Since  $f_n'(x_n^*) = 0$  for  $N_n = N_n^*$ ,

$$f_n'(x) = \int_{x_n^*}^x f_n''(y)dy = \frac{1}{n\sigma^2} \int_x^{x_n^*} (n\sigma^2 |q''(y)| - 1)dy. \quad (5.12)$$

For  $y \geq x_n^*/2$  the integrand stays bounded away from zero, hence

$$f'_n(x) \geq \frac{c(x_n^* - x)}{n\sigma^2} \quad (5.13)$$

for all  $x \geq x_n^*/2$  and some  $c > 0$ . Then

$$f_n(x_n^*) - f_n(0) = \int_0^{x_n^*} f'_n(y) dy \geq \frac{cx_n^{*2}}{8n\sigma^2} \quad (5.14)$$

and the statement follows from  $N_n^* = O(x_n^*)$  (Lemma 2.6) and  $f_n(0) = \frac{(N_n^*)^2}{2n\sigma^2}(1 + o(1))$ . The proof for  $a > 0$  is based on a similar estimate of  $f_n(x_n^*) - f_n(2a)$  and therefore omitted.  $\square$

*Proof of Prop. 2.12.* Fix  $r \in \mathbb{N}_0$ . We have

$$f'_{nr}(x) = f'_n(x) + O\left(\left(\frac{N_n - x}{n\sigma^2}\right)^2\right). \quad (5.15)$$

Clearly  $f'_{nr}(N_n) = q'(N_n) > 0$ . Fix  $\delta \in (0, \varepsilon)$  and  $x \in (x_n^*, (1 + \delta)x_n^*)$ . Remembering (2.10) and the monotonicity of  $q'$ , we obtain

$$\begin{aligned} f'_{nr}(x) &\leq \frac{N_n^* - x_n^*}{n\sigma^2} - (1 + o(1)) \frac{N_n - (1 + \delta)x_n^*}{n\sigma^2} \\ &= -\frac{1}{n\sigma^2} (N_n + o(N_n) - N_n^* - \delta x_n^*) \\ &\leq -\frac{1}{n\sigma^2} ((1 + o(1))\varepsilon N_n^* - \delta x_n^*) \end{aligned} \quad (5.16)$$

which is eventually negative because of  $x_n^* \leq N_n^*$  and  $\delta < \varepsilon$ . It follows that  $f'_{nr}$  does indeed have a zero  $x_{nr}$  which lies between  $(1 + \delta)x_n^*$  and  $N_n$ . On  $(x_n^*, (1 + \delta)x_n^*)$   $f_{nr}$  is strictly decreasing by (5.16), on  $((1 + \delta)x_n^*, N_n)$  the second derivative satisfies

$$f''_{nr}(x) = \frac{1}{n\sigma^2} \left(1 - n\sigma^2 |q''(x)| + O\left(\frac{N_n}{n}\right)\right) \quad (5.17)$$

which stays bounded away from 0, hence  $f_{nr}$  is strictly convex. It follows that  $x_{nr}$  is the unique zero of  $f'_{nr}$  and the maximizer of  $f_{nr}$  in  $(x_n^*, N_n)$ . Moreover

$$(1 + o(1)) \frac{N_n - x_{nr}}{n\sigma^2} = q'(x_{nr}) \leq q'(x_n^*) = \frac{N_n^* - x_n^*}{n\sigma^2} \quad (5.18)$$

hence  $x_{nr} = N_n - (1 + o(1))n\sigma^2 q'(x_{nr}) = N_n + O(N_n^*)$ . The asymptotic expression for  $f_{nr}(x_{nr})$  is easily checked.  $\square$

**5.2. Critical points.** We look for critical points  $(t_n, \xi_n)$  with  $t_n \searrow 0$  and  $\xi_n \in (a, \infty)$  of

$$\Phi_n(t, \xi) = -q(\xi) + n\varphi(t) - mt + t\xi \quad (5.19)$$

(remember  $m = \mu n + N_n$ ). Since our contour integrals involve integrals over  $z = e^t$ , it is convenient to work with functions of a single variable  $t$ : for  $t < \lim_{x \rightarrow a} |q''(x)|$ , let

$$\Psi_n(t) = \Phi_n(t, \xi(t)) = n\varphi(t) - mt + \psi(t). \quad (5.20)$$

where  $\xi(t)$  is the solution of  $q'(\xi(t)) = t$  as on p. 17. Then  $(t, \xi)$  is a critical point of  $\Phi_n(t, \xi)$  if and only if  $\xi = \xi(t)$  and  $\Psi'_n(t) = 0$ . So instead of looking for bivariate

critical points, we may look for zeros of  $\Psi'_n(t)$  in  $(0, \sup |q''(x)|)$ . For later purpose we note the relations

$$\Psi'_n(t) = n\varphi'(t) - m + \xi(t) = -(N_n - \xi(t)) + n\sigma^2 t(1 + O(t)) \quad (5.21)$$

$$\Psi''_n(t) = n\varphi''(t) + \psi''(t) = n\sigma^2(1 + O(t)) + \frac{1}{q''(\xi(t))}, \quad (5.22)$$

with  $\psi(t)$  defined in (4.9). The variable  $t$  and the analysis of  $\Psi_n$  are in some sense dual to the variable  $x$  (or  $\xi$ ) and the variational problem  $f_n(x) = \min$  analyzed in Section 5.1. The analysis becomes more involved, however, because we need to take into account correction terms from  $\sum_{j \geq 3} \kappa_j t^j / j!$ .

**Lemma 5.2** (Inflection point of  $\Psi_n$ ). *Let  $\delta > 0$  such that  $\inf_{(0, \delta)} \varphi''(t) > 0$ . Then for all sufficiently large  $n$ ,  $\Psi_n$  has an inflection point  $t_n^* \in (0, \delta)$ . Any inflection point satisfies  $t_n^* \sim q'(x_n^*)$ , and  $\Psi_n$  is concave below the smallest inflection point and convex above the largest inflection point.*

If the inflection point is unique, then  $\Psi_n$  is concave on  $(0, t_n^*)$  and  $(t_n^*, \delta)$ . In general we do not know whether the inflection point is unique, however the asymptotic behavior  $t_n^* \sim q'(x_n^*)$  is uniquely determined, and all statements below hold for every inflection point  $t_n^*$ .

*Remark 4.* It follows that  $t_n^*$  is of the order of  $N_n^*/n$ : from Lemma 5.2 and the definition of  $N_n^*$ , we have  $t_n^* \sim \frac{N_n^* - x_n^*}{n\sigma^2}$ , and then Lemma 2.6 yields

$$\delta \frac{N_n^*}{n} \leq t_n^* \leq C \frac{N_n^*}{n}. \quad (5.23)$$

*Proof of Lemma 5.2.* As a preliminary observation, we note that any solution  $t_n^*$  of  $\Psi'_n(t) = 0$  converges to zero: this is because  $-\psi''(t_n^*) = n\varphi''(t) \geq n \inf_{(0, \delta)} \varphi'' \rightarrow \infty$ .

As  $t \searrow 0$  at fixed  $n$ , we have  $\varphi''(t) \rightarrow \sigma^2$  and  $\psi''(t) \rightarrow -\infty$  hence  $\Psi''_n(t) \rightarrow -\infty$ . On the other hand we may choose  $\varepsilon_n$  in such a way that  $\varepsilon_n \searrow 0$  and  $\Psi''_n(\varepsilon_n) \rightarrow \infty$ : indeed  $\psi''(q'(x_n^*)) = 1/q''(x_n^*) = -n\sigma^2$  by definition of  $x_n^*$ , so choosing  $\varepsilon_n \gg q'(x_n^*)$  in such a way that  $|\psi''(\varepsilon_n)| \ll |\psi''(q'(x_n^*))| = n\sigma^2$  we find  $\Psi''_n(\varepsilon_n) = (1 + o(1))n\sigma^2 + o(n\sigma^2) \rightarrow \infty$ . It follows from the intermediate value theorem that  $\Psi''_n(t) = 0$  has a solution  $t_n^*$  in  $(0, \varepsilon_n)$ . It satisfies

$$\frac{q''(x_n^*)}{q''(\xi(t_n^*))} = -\frac{\psi''(t_n^*)}{n\sigma^2} = \frac{\varphi''(t_n^*)}{\sigma^2} = 1 + O(t_n^*) \rightarrow 1. \quad (5.24)$$

Assumption 2.1(iv) and its consequence (2.5) imply that  $\xi(t_n^*)/x_n^* \rightarrow 1$  and  $t_n^* = q'(\xi(t_n^*)) \sim q'(x_n^*)$ .  $\square$

$\Psi'_n(t)$  is positive for small  $t$  and decreasing on  $(0, t_n^*)$  and increasing on  $(t_n^*, \delta)$ . Define  $\xi_n^* = \xi(t_n^*)$  and notice  $\xi_n^* \sim x_n^*$  from the proof of the previous lemma.

For  $N_n \ll n$ , let  $\eta_n \sim N_n/(n\sigma^2)$  be the solution of (3.4). Notice  $\Psi'_n(\eta_n) = \xi(\eta_n) \rightarrow \infty$ .

**Lemma 5.3.** *Let  $\delta > 0$  be as in Lemma 5.2,  $\varepsilon > 0$ , and  $t_n^*$  an inflection point of  $\Psi_n$ . Let  $N_n \rightarrow \infty$  with  $(1 + \varepsilon)N_n^* \leq N_n \ll n$ . Then for sufficiently large  $n$ ,  $\Psi'_n$  has exactly two zeros in  $(0, \delta)$ , one zero  $t_n \in (0, t_n^*)$  and another  $t'_n \in (t_n^*, \eta_n)$ . Set  $\xi_n := \xi(t_n)$ . We have*

$$0 \leq N_n - \xi_n = O(N_n^*),$$

$\limsup_{n \rightarrow \infty} t_n/t_n^* < 1$ ,  $\liminf_{n \rightarrow \infty} \xi_n/x_n^* > 1$ , and  $\liminf_{n \rightarrow \infty} t'_n/\eta_n > 0$ .

*Remark 5.* When  $N_n \gg N_n^*$ , we use Eq. (2.4), Lemmas 2.6 and 5.2 and find

$$\xi_n \sim N_n, \quad t_n = q'(\xi_n) \sim q'(N_n) \ll q'(x_n^*) \sim t_n^* = O\left(\frac{N_n^*}{n}\right) \quad (5.25)$$

hence  $t_n = o(t_n^*)$ . When  $N_n = O(N_n^*)$ , we have instead  $c^{-1}t_n^* \leq t_n \leq ct_n^*$  for some  $c > 0$  and all sufficiently large  $n$ : The upper bound is part of Lemma 5.3. For the lower bound, we note that  $t_n = q'(\xi_n) \geq q'(N_n)$  because  $q'$  is decreasing and  $\xi_n \leq N_n$ . Since  $N_n$  is of the order of  $N_n^*$ , Eq. (2.4) shows that  $q'(N_n)$  is of the order of  $q'(N_n^*)$  which in turn is of the order of  $q'(x_n^*) \sim t_n^*$ .

*Proof of Lemma 5.3.* We check first that  $t_n^* < \eta_n$ . With  $C \geq 1$  as in Lemma 2.6, we have

$$t_n^* \sim q'(x_n^*) = \frac{N_n^* - x_n^*}{n\sigma^2} \leq (1 - C^{-1})\frac{N_n^*}{n\sigma^2} \quad (5.26)$$

so

$$t_n^*/\eta_n \leq (1 + o(1))(1 - C^{-1})N_n^*/N_n \leq (1 + o(1))\frac{1 - C^{-1}}{1 + \varepsilon} \quad (5.27)$$

is bounded away from 1. As  $t \rightarrow 0$  at fixed  $n$ ,  $\Psi'_n(t) \sim \xi(t) \rightarrow \infty$ , and as  $n \rightarrow \infty$ ,  $\Psi'_n(\eta_n) \rightarrow \infty$ . Furthermore

$$\begin{aligned} \Psi'_n(t_n^*) &= -(N_n - \xi_n^*) + (1 + o(1))n\sigma^2 q'(\xi_n^*) \\ &= -(N_n - (1 + o(1))x_n^*) + (1 + o(1))n\sigma^2 q'(x_n^*) \\ &= -(N_n - (1 + o(1))x_n^*) + (1 + o(1))(N_n^* - x_n^*) \\ &= -(N_n - N_n^*) + o(N_n^*) \rightarrow -\infty. \end{aligned} \quad (5.28)$$

The intermediate value theorem proves the existence of a zero  $t_n \in (0, t_n^*)$  and another zero  $t'_n \in (t_n^*, \eta_n)$ .

Suppose by contradiction that  $t_n/t_n^* \rightarrow 1$ . Then the identities  $t_n = q'(\xi_n)$ ,  $t_n^* = q'(\xi_n^*)$  and Eq. (2.4) imply  $\xi_n \sim \xi_n^* \sim x_n^*$  and an estimate analogous to (5.28) shows  $\Psi'_n(t_n) \rightarrow -\infty$ , in contradiction with  $\Psi'_n(t_n) = 0$ . It follows that  $t_n/t_n^*$  and  $\xi_n/x_n^*$  are bounded away from 1. In addition,  $t_n \leq t_n^* = O(N_n^*/n)$  and

$$N_n - \xi_n \sim n\sigma^2 t_n = O(N_n^*). \quad (5.29)$$

For the lower bound of  $t'_n$ , we use  $t'_n \geq t_n^* \sim q'(x_n^*)$  and Assumption 2.1 to get  $\xi'_n \leq (1 + o(1))x_n^*$  where  $q'(\xi'_n) = t'_n$ . Since  $t'_n = O(\eta_n) \rightarrow 0$ ,  $\Psi'_n(t'_n) = 0$  together with (5.21) yields

$$t'_n \sim \frac{N_n - \xi'_n}{n\sigma^2} \sim \eta_n \left(1 - \frac{\xi'_n}{N_n}\right) \geq \eta_n \left(1 - \frac{N_n^*}{N_n}\right) \quad (5.30)$$

and  $N_n \geq (1 + \varepsilon)N_n^*$  implies  $\liminf t'_n/\eta_n > 0$ . Notice that, in view of (5.27), we have a fortiori  $\liminf t'_n/t_n^* > 0$ .

We have actually shown that for every inflection point  $t_n^*$ ,  $t/t_n^* \leq 1$  stays bounded away from 1. In particular, in case of non-uniqueness of  $t_n^*$  we may choose  $t_n^*$  as the *smallest* inflection point of  $\Psi_n$ . Then  $\Psi'_n$  is strictly increasing on  $(0, t_n^*)$ , consequently the zero  $t_n$  is unique. A similar argument shows that  $t'_n$  is unique.  $\square$

When  $\liminf_{n \rightarrow \infty} N_n/n > 0$ , the sequence  $\eta_n$  is either no longer defined or it does not converge to zero. The previous lemma is modified as follows.

**Lemma 5.4.** *Assume  $N_n \rightarrow \infty$  with  $\liminf N_n/n > 0$ . Then there exists  $\delta_0 > 0$  such that for large  $n$ ,  $\Psi_n$  has exactly one critical point  $t_n$  in  $(0, \delta_0)$ . The critical point lies in  $(0, t_n^*)$  and it satisfies  $0 \leq N_n - \xi(t_n) = O(N_n^*)$ .*



*Proof.* The existence and uniqueness of a critical point in  $(0, t_n^*)$  as well as the properties of  $\xi(t_n)$  are proven as in the previous lemma. Fix  $\delta_0 > 0$  such that  $\varphi'(\delta_0) \leq \mu + \frac{1}{2} \liminf(N_n/n) =: \mu + \varepsilon/2$ . Then

$$\Psi_n'(\delta_0) = -n \left( \mu + \frac{N_n}{n} - \varphi'(\delta_0) \right) + \xi(\delta_0) \leq -n \left( \frac{\varepsilon}{2} + o(1) \right) + \xi(\delta_0) \rightarrow -\infty. \quad (5.31)$$

It follows that  $\Psi_n' < 0$  on  $(t_n^*, \delta_0)$ .  $\square$

**5.3. Hessians.** Let  $t_n \in (0, t_n^*)$  be the critical point of  $\Psi_n(t)$ , and  $\xi_n = \xi(t_n)$ . Thus  $(t_n, \xi_n)$  is a critical point of  $\Phi_n(t, \xi)$ . Lemma B.1 shows

$$\Psi_n''(t_n) = -\frac{\det \text{Hess } \Phi_n(t_n, \xi_n)}{|q''(\xi_n)|} = -\frac{1 - n \text{Re } \varphi''(t_n) |q''(\xi_n)|}{|q''(\xi_n)|}. \quad (5.32)$$

**Lemma 5.5.** *Assume  $N_n \rightarrow \infty$  with  $\liminf_{n \rightarrow \infty} (N_n/N_n^*) > 1$  and let  $(t_n, \xi_n)$  be the unique critical point of  $\Phi_n$  in  $(0, t_n^*) \times (a, \infty)$ .*

- (a) *If  $N_n \gg N_n^*$ , then  $\det \text{Hess } \Phi_n(t_n, \xi_n) \rightarrow -1$ .*
- (b) *If  $N_n = O(N_n^*)$ , then  $\det \text{Hess } \Phi_n(t_n, \xi_n) = -(1 - n\sigma^2 |q''(\xi_n)|) + o(1)$  and it stays bounded away from zero.*

*Proof.* (a) If  $N_n \gg N_n^*$ , then by Lemmas 5.3 and 5.4, we have  $\xi_n = N_n + O(N_n^*) \sim N_n$ , hence in particular  $\xi_n \gg N_n^* \geq x_n^*$ . Exploiting the monotonicity and the convexity of  $q'$ , we have

$$0 \geq q''(\xi_n) \geq \frac{q'(\xi_n) - q'(x_n^*)}{\xi_n - x_n^*} = \frac{O(N_n^*/(n\sigma^2))}{N_n(1 + o(1))} = -\frac{1}{n\sigma^2} O\left(\frac{N_n^*}{N_n}\right) \quad (5.33)$$

from which we get

$$\det \text{Hess } \Phi_n(t_n, \xi_n) = -1 + n\sigma^2(1 + O(t_n)) |q''(\xi_n)| = -1 + O\left(\frac{N_n^*}{N_n}\right) \rightarrow -1. \quad (5.34)$$

(b) If  $N_n = O(N_n^*)$ : By Lemma 5.3,  $\xi_n \geq (1 + \varepsilon)x_n^*$  for some  $\varepsilon > 0$ . Consequently  $|q''(\xi_n)| \leq (1 - \delta) |q''(x_n^*)| = (1 - \delta)/(n\sigma^2)$  for some  $\delta > 0$  and large  $n$ , from which we deduce that

$$1 - n \text{Re } \varphi''(t_n) |q''(\xi_n)| \geq 1 - (1 + O(t_n))(1 - \delta) = 1 - \delta + O(t_n), \quad (5.35)$$

in particular the expression stays bounded away from zero. We also have

$$\det \text{Hess } \Phi_n(t_n, \xi_n) = -(1 - n\sigma^2 |q''(\xi_n)|) + o(t_n) n\sigma^2 |q''(\xi_n)| \quad (5.36)$$

Since  $\xi_n > x_n^*$ , the estimate (5.33) still holds true and

$$t_n n\sigma^2 |q''(\xi_n)| = t_n O\left(\frac{N_n^*}{N_n}\right) = O(t_n) \rightarrow 0. \quad (5.37)$$

$\square$

**5.4. Gaussian approximation.** Next we address the Gaussian approximation for the evaluation of  $H_n$ . Because of Theorem 4.4, we need not deal with a bivariate integral and instead may use

$$\frac{1}{\pi} \int_0^{\varepsilon_n} e^{n \text{Re } \varphi(t) - mt} \text{Im } G(e^t) dt \sim \frac{1}{\sqrt{2\pi}} \int_0^{\varepsilon_n} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt \quad (5.38)$$

as  $n \rightarrow \infty$  and  $\varepsilon_n \searrow 0$ . Remember that  $\Psi_n(t_n) = \Phi_n(t_n, \xi_n)$  and from Eq. (5.32) and Lemma 5.5,

$$\frac{\psi_n''(t_n)}{\Psi_n''(t_n)} = \frac{1}{1 - n\sigma^2 |q''(\xi_n)|} \quad (5.39)$$

with a denominator bounded away from zero. The following technical lemma helps estimate the prefactor  $\sqrt{|\psi''(t)|}$ . Set

$$R_n(t) := \Psi_n(t) + \log \sqrt{|\psi''(t)|}. \quad (5.40)$$

**Lemma 5.6.** *Let  $N_n \rightarrow \infty$  with  $\liminf(N_n/N_n^*) > 1$  and  $t_n \in (0, t_n^*)$  the zero of  $\Psi'_n(t)$  from Lemma 5.3. Then  $R'_n$  has at least one zero  $s_n \in (0, t_n)$ . Moreover there exists a sequence  $\delta_n \searrow 0$  such that every such zero lies in  $((1 - \delta_n)t_n, t_n)$ .*

*Proof.* As  $t \searrow 0$  at fixed  $n$ , using Lemma 4.3, we have

$$\begin{aligned} R'_n(t) &= \Psi'_n(t) + \frac{1}{2} \frac{\psi^{(3)}(t)}{\psi''(t)} = n(\operatorname{Re} \varphi'(t) - \mu) - N_n + \psi(t) + \frac{1}{2} \frac{\psi^{(3)}(t)}{\psi''(t)} \\ &= n(\sigma^2 t + O(t^2)) - N_n + \psi'(t) + O(1/t) \\ &= -N_n + o(1) + (1 + o(1))\psi'(t) \rightarrow \infty, \end{aligned} \quad (5.41)$$

hence  $t \mapsto R_n(t)$  is initially increasing. At  $t = t_n$  we have

$$R'_n(t_n) = \Psi'_n(t_n) + \frac{1}{2} \frac{\psi^{(3)}(t_n)}{\psi''(t_n)} = \frac{1}{2} \frac{\psi^{(3)}(t_n)}{\psi''(t_n)} < 0. \quad (5.42)$$

The intermediate value theorem guarantees the existence of a zero  $s_n$  of  $R'_n$ . Set

$$y_n := \psi'(s_n) + \frac{1}{2} \frac{\psi^{(3)}(s_n)}{\psi''(s_n)}. \quad (5.43)$$

In view of Lemma 4.3, we have  $y_n \sim \psi'(s_n)$  and by Eqs. (2.4) and (4.10),  $q'(y_n) \sim s_n$ . From  $s_n \leq t_n \leq t_n^*$  and Lemma 2.7 we get that  $y_n$  is larger than  $x_n^*$  and actually bounded away from it. By the definition of  $s_n$ ,

$$0 = n(\operatorname{Re} \varphi'(s_n) - \mu) - N_n + y_n = n\sigma^2 s_n(1 + O(s_n)) - [N_n - y_n] \quad (5.44)$$

hence

$$\frac{N_n - y_n}{n\sigma^2} \sim s_n \sim q'(y_n). \quad (5.45)$$

Let  $x_n \in (x_n^*, N_n)$  be the solution of  $q'(x_n) = (N_n - x_n)/(n\sigma^2)$ . Now

$$\frac{d}{dy} \left( q'(y) - \frac{N_n - y}{n\sigma^2} \right) = q''(y) + \frac{1}{n\sigma^2} = \frac{1 - n\sigma^2 |q''(y)|}{n\sigma^2} \quad (5.46)$$

and  $n\sigma^2 |q''(y)|$  stays bounded away from 1 when  $y \geq (1 + \delta)x_n^*$ , we find that for some  $c > 0$ , we have

$$\left| q'(y_n) - \frac{N_n - y}{n\sigma^2} \right| \geq \frac{c|y_n - x_n|}{n\sigma^2} \quad (5.47)$$

hence  $y_n - x_n = o(N_n - y_n) = o(N_n)$ . If  $N_n \gg N_n^*$ , Lemma 2.7 says  $x_n \sim N_n$  and we deduce  $y_n - x_n = o(x_n)$ . If  $N_n = O(N_n^*)$ , we use  $x_n \geq x_n^*$  in conjunction with Lemma 2.6 and find  $y_n - x_n = o(N_n^*) = o(x_n^*) = o(x_n)$ .

Thus we have checked that  $y_n \sim x_n$ , which in turn yields  $s_n \sim q'(x_n)$ . An entirely similar argument shows  $t_n \sim q'(x_n)$ , so we must have  $s_n \sim t_n$ . This holds for every zero in  $(0, t_n)$ , in particular the smallest one, and the lemma follows.  $\square$

**Lemma 5.7.** *Let  $N_n \rightarrow \infty$  with  $\liminf(N_n/N_n^*) > 1$  and  $N_n = O(N_n^*)$ . Then*

$$\int_0^{t'_n} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt \sim \sqrt{\frac{2\pi}{1 - n\sigma^2 q''(\xi_n)}} e^{\Psi_n(t_n)}.$$

*Proof.* By Lemma 5.2 and 5.3,  $\Psi_n$  is increasing on  $(0, t_n)$  and decreasing on  $(t_n, t'_n)$ . We use a Gaussian approximation to  $\Psi_n$  around  $t_n$  and adapt [17, Lemma 2.1]. First we check that the window  $(t_n - \varepsilon_n, t_n + \varepsilon_n)$  contributing most to the Gaussian integral fits amply into  $(0, t'_n)$ , i.e.,

$$\varepsilon_n := \frac{1}{\sqrt{|\Psi_n''(t_n)|}} = o(\min(t_n, t'_n - t_n)). \quad (5.48)$$

By Lemma 5.3, we have  $t_n \leq (1 - \delta)t_n^* \leq (1 - \delta)t'_n$  for some  $\delta \in (0, 1)$ , hence  $t'_n - t_n \geq \delta t_n$  and

$$\min(t_n, t'_n - t_n) \geq \delta t_n. \quad (5.49)$$

Eq. (5.39) and Lemma 5.5 show that  $\Psi_n''(t_n)$  is of the order of  $\psi''(t_n)$ ,  $\varepsilon_n^2$  of the order of  $1/|\psi_n''(t_n)|$ . By Lemma 4.3,  $t_n^2|\psi_n''(t_n)| \rightarrow \infty$  hence  $\varepsilon_n^2/t_n^2 \rightarrow 0$  and (5.48) follows. The same argument shows that for

$$c_n \rightarrow \infty \quad \text{with} \quad c_n \varepsilon_n = o(t_n) \quad (5.50)$$

we still get  $c_n \varepsilon_n = o(\min(t_n, t'_n - t_n))$ . Second, we observe that the prefactor  $\psi''(t)$  is essentially constant on the relevant window: Because of  $\psi^{(3)}(u)/\psi''(u) = O(1/u)$  (Lemma 4.3), we have

$$\frac{\psi''(t)}{\psi''(t_n)} = \exp\left(-\int_{t_n}^t \frac{\psi^{(3)}(u)}{\psi''(u)} du\right) = \exp\left(O\left(\log \frac{t}{t_n}\right)\right) \quad (5.51)$$

and

$$\sup\left\{\left|\frac{\psi''(t)}{\psi''(t_n)} - 1\right| \mid |t - t_n| \leq c_n \varepsilon_n\right\} = O(c_n \varepsilon_n) \rightarrow 0. \quad (5.52)$$

Third, we note that cubic corrections can be neglected:

$$\Psi_n(t) = \Psi_n(t_n) + \frac{1}{2}(1 + o(1))\Psi_n''(t_n)(t - t_n)^2 \quad (5.53)$$

uniformly in  $|t - t_n| \leq c_n \varepsilon_n$ . To this aim write, with the help of (5.52),

$$\begin{aligned} \Psi_n''(t) &= n(\varphi''(t_n) + O(c_n \varepsilon_n)) + \psi_n''(t_n)(1 + O(c_n \varepsilon_n)) \\ &= \Psi_n''(t_n) + \psi_n''(t_n)\left(O\left(\frac{nc_n \varepsilon_n}{\psi_n''(t_n)}\right) + O(c_n \varepsilon_n)\right) \end{aligned}$$

Now  $n/\psi''(t_n) = nq''(\xi_n) = O(1)$  by Lemma 5.5 and  $\psi''(t_n) = O(\Psi_n''(t_n))$  by Eq. (5.39) and the same lemma, hence

$$\Psi_n''(t) = (1 + O(c_n \varepsilon_n))\Psi_n''(t_n) \quad (5.54)$$

in  $|t - t_n| \leq c_n \varepsilon_n$  and (5.53) follows. Eqs. (5.52), (5.53) and (5.39) yield

$$\int_{t_n - c_n \varepsilon_n}^{t_n + c_n \varepsilon_n} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt \sim \sqrt{\frac{2\pi\psi''(t_n)}{\Psi_n''(t_n)}} \sim \sqrt{\frac{2\pi}{1 - n\sigma^2 q''(\xi_n)}} e^{\Psi_n(t_n)}. \quad (5.55)$$

Our next task is to estimate the integral on  $(0, t_n - c_n \varepsilon_n)$  and  $(t_n + c_n \varepsilon_n, t'_n)$ , taking into account that the prefactor  $\sqrt{|\psi''(t)|}$  goes to infinity. On both intervals we have

$$\Psi_n(t) \leq \Psi_n(t_n) - \frac{1}{2}(1 + O(c_n \varepsilon_n))c_n^2. \quad (5.56)$$

For  $t \geq t_n + c_n \varepsilon_n$  we have  $-\psi''(t) \leq -\psi''(t_n)$  hence

$$\sqrt{|\psi''(t)|} \exp(\Psi_n(t)) \leq \exp\left(\Psi_n(t_n) - \frac{1}{2}(1 + o(1))c_n^2 + \frac{1}{2} \log |\psi''(t_n)|\right). \quad (5.57)$$

By Lemma 5.6 we may choose  $c_n \varepsilon_n \searrow 0$  in such a way that  $R_n$  is increasing on  $(0, t_n(1 - c_n \varepsilon_n))$ , which yields

$$\Psi_n(t) + \log \sqrt{|\psi''(t)|} \leq \Psi_n(t_n - c_n \varepsilon_n) + \log \sqrt{|\psi''(t_n - c_n \varepsilon_n)|}. \quad (5.58)$$

This estimate, combined with Eqs. (5.52) and (5.56), shows that (5.57) holds true not only in  $(t_n + c_n \varepsilon_n, t'_n)$  but also in  $(0, t_n - c_n \varepsilon_n)$ . Next we check that we can choose  $c_n \rightarrow \infty$  so that not only  $c_n \varepsilon_n = o(t_n)$  but in fact

$$-(1 + o(1))c_n^2 + \log |\psi''(t_n)| \rightarrow -\infty. \quad (5.59)$$

By Lemma 4.3 (iii), as  $t \searrow 0$ , we may estimate  $|\psi''(t)|$  as follows: fix  $t_0 > 0$  and take  $t \in (0, t_0)$ , then

$$\begin{aligned} \log(-\psi''(t)) &= \log(-\psi''(t_0)) - \int_t^{t_0} \frac{\psi^{(3)}(u)}{\psi''(u)} du \leq \log(-\psi''(t_0)) + \int_t^{t_0} \frac{C}{u} du \\ &= \log(-\psi''(t_0)) + C \log \frac{t_0}{t} = O(|\log t|). \end{aligned} \quad (5.60)$$

In particular  $\log |\psi''(t_n)| = O(\log t_n)$ . On the other hand

$$\frac{c_n^2}{|\log t_n|} = \left(\frac{c_n \varepsilon_n}{t_n}\right)^2 \frac{\Psi_n''(t_n) t_n^2 |\psi''(t_n)|}{\psi''(t_n) |\log t_n|} \quad (5.61)$$

The ratio  $\Psi_n''(t_n)/\psi''(t_n)$  stays bounded away from 0 and by Lemma 4.3,

$$\frac{t^2 |\psi''(t)|}{|\log t|} \rightarrow \infty. \quad (5.62)$$

Thus we may find a function  $\omega(t) \rightarrow 0$  such that  $\omega(t)^2 t^2 |\psi''(t)|/|\log t|$  still goes to infinity as  $t \searrow 0$ , set  $c_n \varepsilon_n = t_n \omega(t_n)$ , and then Eq. (5.59) holds true. The bound (5.57) then shows

$$\int_0^{t_n - c_n \varepsilon_n} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt + \int_{t_n + c_n \varepsilon_n}^{t'_n} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt = o\left(e^{\Psi_n(t_n)}\right). \quad (5.63)$$

Eqs. (5.55) and (5.63) complete the proof of the lemma.  $\square$

**Lemma 5.8.** *Let  $N_n \rightarrow \infty$  with  $N_n^* \ll N_n \ll n$ . Then*

$$\int_0^{t'_n} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt \sim \sqrt{2\pi} e^{\Psi_n(t_n)}.$$

*Proof.* The proof of Lemma 5.7 applies without any changes, the end result simplifies because  $1 - n\sigma^2 q''(\xi_n) \rightarrow 1$  by Lemma 5.5.  $\square$

**Lemma 5.9.** *Let  $N_n \rightarrow \infty$  with  $\liminf(N_n/n) > 0$ . Let  $\delta_0 > 0$  as in Lemma 5.4. Then*

$$\int_0^{\delta_0} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} dt \sim \sqrt{2\pi} e^{\Psi_n(t_n)}.$$

*Proof.* By Lemma 5.4,  $\Psi_n(t)$  has a unique critical point  $t_n$  in  $(0, \delta_0)$  and we may use a Gaussian approximation on this interval—there is no need to restrict to an interval whose length goes to zero. Apart from this difference, the proof is identical to Lemmas 5.7 and 5.8.  $\square$

## 6. EVALUATION OF CONTOUR INTEGRALS. PROOF OF THE MAIN THEOREMS

The proof of the main theorems starts from the decomposition

$$\mathbb{P}(S_n = n\mu + N_n) = H_n + V_n \quad (6.1)$$

where  $H_n$  and  $V_n$  are defined as in (3.7). The correctness of (6.1) is checked as in Section 3, building on the properties of  $G(z)$  proven in Section 4. For the proof of Theorem 2.10 it is convenient to decompose  $H_n$  further as  $H_n = H_n^1 + H_n^2$  where

$$H_n^1 = \frac{1}{\pi} \int_0^{t'_n} e^{n\operatorname{Re} \varphi(t) - (\mu n + N_n)t} \sin(n\operatorname{Im} \varphi(t)) dt \quad (6.2)$$

and  $H_n^2$  is a similar integral, but with integration from  $t'_n$  to  $\eta_n$ . In the proof of Theorem 2.11 we adopt slightly modified definitions and replace the sequence  $\eta_n \sim N_n/(n\sigma^2)$  in the domain of integration by another sequence  $\varepsilon_n \searrow 0$  or by some fixed small  $\varepsilon > 0$ .

**6.1. Evaluation of  $V_n$ .** Theorem 2.11 only needs upper bounds for  $V_n$ , provided in Lemmas 6.1 and 6.2. Theorems 2.9 and 2.10 requires the full asymptotic behavior of  $V_n$  proven in Lemma 6.3.

**Lemma 6.1.** *Suppose  $N_n \rightarrow \infty$  and  $N_n \geq \delta n$  for some  $\delta > 0$  and all  $n \in \mathbb{N}$ . Then for suitable  $C = C_\delta > 0$ , every sufficiently small  $\eta > 0$ , and all  $n \in \mathbb{N}$*

$$\frac{1}{\pi} \int_0^\pi e^{n\operatorname{Re} \varphi(\eta+i\theta) - (\mu n + N_n)\eta} d\theta \leq e^{-Cn\eta}.$$

*Proof.* Let  $\mathcal{S}_+$  be the half-strip  $\{t \in \mathbb{C} \mid \operatorname{Re} t \geq 0, \operatorname{Im} t \in [0, \pi]\}$ . By Theorem 4.2, we know that as  $t \rightarrow 0$  in  $\mathcal{S}_+$

$$\operatorname{Re}(\varphi(t) - \mu t - \frac{N_n}{n}t) = -\frac{N_n}{n}\operatorname{Re} t + \frac{1}{2}\sigma^2((\operatorname{Re} t)^2 - (\operatorname{Im} t)^2) + O(t^3). \quad (6.3)$$

Then for sufficiently small  $\varepsilon_1 > 0$  and all  $t \in \mathcal{S}_+$  with  $\max(|\operatorname{Re} t|, |\operatorname{Im} t|) \leq \varepsilon_1$ , the right side of Eq. (6.3) is smaller than  $-\delta \operatorname{Re} t/2$ , which shows

$$\frac{1}{\pi} \int_0^{\varepsilon_1} e^{n\operatorname{Re} \varphi(\eta+i\theta) - (\mu n + N_n)\eta} d\theta \leq \frac{\varepsilon_1}{\pi} e^{-n\delta\eta/2} \quad (6.4)$$

for all  $\eta \in (0, \varepsilon_1)$ . On the other hand on the unit circle  $G(z)$  is given by a power series with strictly positive coefficients and therefore  $|G(z)|$  has a unique maximum at  $z = 1$ . It follows that for all  $t = i\theta$  with  $\theta \in [\varepsilon_1, 2\pi - \varepsilon_1]$ , we know  $\operatorname{Re} \varphi(t) < 0$ . By continuity this extends to some thin vertical strip  $\operatorname{Im} t \in [\varepsilon_1, 2\pi - \varepsilon_1]$ ,  $\operatorname{Re} t \in [0, \varepsilon_2]$  so that  $\operatorname{Re}(\varphi(t) - \mu t) \leq -\mu \operatorname{Re} t$  and

$$\frac{1}{\pi} \int_{\varepsilon_1}^\pi e^{n\operatorname{Re} \varphi(\eta+i\theta) - (\mu n + N_n)\eta} d\theta \leq \frac{\pi - \varepsilon_1}{\pi} e^{-n\mu\eta} \quad (6.5)$$

for all  $\eta \in (0, \varepsilon_2)$ . To conclude, we let  $C_\delta := \min(\delta/2, \mu)$ .  $\square$

**Lemma 6.2.** *Assume  $N_n = m - n\mu \rightarrow \infty$  and  $N_n = o(n)$ . Let  $\varepsilon_n \searrow 0$  with  $\varepsilon_n \leq \eta_n = (1 + o(1))\frac{N_n}{n\sigma^2}$ . Then for suitable constant  $C > 0$ , as  $n \rightarrow \infty$ ,*

$$\left| \frac{1}{2\pi} \int_0^\pi e^{n\operatorname{Re} \varphi(\varepsilon_n+i\theta) - m\varepsilon_n} d\theta \right| \leq e^{-(1+o(1))N_n\varepsilon_n/2}.$$

*Proof.* The proof is analogous to Lemma 6.1. We start from the estimate (6.3). In a sufficiently small  $\varepsilon_1$ -neighborhood of the corner 0 of the half-strip  $\mathcal{S}_+$ , we have

$$\operatorname{Re} \left( \varphi(t) - \frac{m}{n}t \right) \leq - \left( \frac{N_n}{n} - \frac{1}{2}\sigma^2 \operatorname{Re} t \right) \operatorname{Re} t. \quad (6.6)$$

Notice that  $\varepsilon_1$  can be chosen  $n$ -independent: we only need  $-\sigma^2(\operatorname{Im} t)^2 + O((\operatorname{Im} t)^3) \leq 0$  for  $|\operatorname{Im} t| \leq \varepsilon_1$ . When  $\operatorname{Re} t = \varepsilon_n$  with  $\varepsilon_n \leq \frac{N_n}{n\sigma^2}$ , the upper bound in (6.6) is in turn bounded by  $-(1+o(1))N_n\varepsilon_n/2$ .

When  $\operatorname{Re} t = \varepsilon_n$  is small but  $\operatorname{Im} t$  is bounded away from 0 and  $2\pi$ , we estimate

$$\operatorname{Re} \left( \varphi(t) - \frac{m}{n}t \right) \leq -\frac{m}{n} \operatorname{Re} t = - \left( \mu + \frac{N_n}{n} \right) t \leq -\frac{N_n}{n} t. \quad (6.7)$$

and we conclude as in Lemma 6.1.  $\square$

The proof of the next lemma is closely related to the treatment of moderate deviations for random variables with generating functions analytic beyond  $z = 1$  given by Ibragimov and Linnik [12].

**Lemma 6.3.** *Let  $N_n \rightarrow \infty$  along  $\sqrt{n} \ll N_n = O(n^{1-\gamma})$  for some  $\gamma > 0$ . Define  $\eta_n \sim N_n/(n\sigma^2)$  by (3.4) and  $V_n$  as in (3.7). Then Eq. (3.11) holds true.*

*Proof.* Since  $\operatorname{Re} \varphi'(t) = \mu + \sigma^2 t + O(t^2)$  is strictly increasing for small  $t > 0$  and  $N_n/n \rightarrow 0$ , we may fix  $\delta > 0$  small enough so that for large  $n \geq n_\delta$ , the equation (3.4) has indeed a unique solution  $\eta_n \in (0, \delta)$ , which satisfies  $\eta_n \sim N_n/(n\sigma^2)$ . Arguments analogous to the proof of Eq. (6.5) show

$$\frac{1}{\pi} \left| \int_\delta^\pi e^{n(\varphi(\eta_n + i\theta)) - (\mu n + N_n)(\eta_n + i\theta)} \cos(n \operatorname{Im} \varphi(\eta_n + i\theta)) d\theta \right| \leq e^{-n\mu\eta_n}. \quad (6.8)$$

We also have, uniformly in  $s \in (0, \delta_n)$ ,

$$\begin{aligned} \varphi(\eta_n + is) &= \varphi(\eta_n) + \varphi'(\eta_n)is - \frac{1}{2}\varphi''(\eta_n)s^2 + O(s^3) \\ &= \varphi(\eta_n) + i\left(\mu + \frac{N_n}{n}\right)s - \operatorname{Im} \varphi'(\eta_n)s - \frac{1}{2}\varphi''(\eta_n)s^2 + O(s^3) \end{aligned} \quad (6.9)$$

and therefore

$$\operatorname{Re} \varphi(\eta_n + is) - \left( \mu + \frac{N_n}{n} \right) \eta_n = \operatorname{Re} \varphi(\eta_n) - \frac{1}{2}\varphi''(\eta_n)s^2(1 + O(s)) - \operatorname{Im} \varphi'(\eta_n)s. \quad (6.10)$$

Since  $\varphi''(\eta_n) \rightarrow \sigma^2$ , standard arguments show

$$\frac{1}{\pi} \int_0^\delta e^{n \operatorname{Re} \varphi(\eta_n) - \frac{n}{2} \operatorname{Re} \varphi''(\eta_n) s^2 (1 + O(s))} ds \sim \frac{\exp(n \operatorname{Re} \varphi(\eta_n))}{\sqrt{2\pi n \sigma^2}}, \quad (6.11)$$

moreover the contribution to the interval from  $s \geq \delta_n := (\log n)/\sqrt{n}$  is negligible and Eq. (6.11) holds true with  $\delta$  replaced by  $\delta_n$ . By Theorem 4.2 and the relation  $\operatorname{Im} G(e^t) = |G(e^t)| \operatorname{Im} \varphi(t)$  the imaginary part of  $\varphi'(\eta_n)$  vanish faster than any power of  $\eta_n = O(n^{-\gamma})$ , hence  $n \operatorname{Im} \varphi'(\eta_n)$  can be neglected; the same argument works for  $n \operatorname{Im} \varphi''(\eta_n)$ . For the cosine, we look separately at  $(0, \delta_n)$  and  $(\delta_n, \delta)$ . On  $(0, \delta_n)$ , again by Theorem 4.2,  $\sup_{s \in (0, \delta_n)} |\operatorname{Im} \varphi(\eta_n + is)|$ , vanishes faster than any power of  $\max(\eta_n, \delta_n)$  hence it can be neglected. On  $(\delta_n, \delta)$  we simply bound the

cosine by 1. As the contribution from  $(\delta_n, \delta)$  to the integral (6.11) is negligible, combining with (6.8) we find in the end

$$V_n = (1 + o(1)) \frac{\exp(n \operatorname{Re} \varphi(\eta_n) - \mu n - N_n)}{\sqrt{2\pi n \sigma^2}} + O(e^{-n\mu\eta_n}). \quad (6.12)$$

By Theorem 4.2 and Definition 2.8, we have

$$n \operatorname{Re} \varphi(\eta_n) - \mu n - N_n = \left(1 + O\left(\frac{N_n}{n}\right)\right) \frac{N_n^2}{2n\sigma^2} \quad (6.13)$$

with correction terms expressed in terms of the Cramér series. In particular, the exponent goes to  $-\infty$  as  $-N_n^2/n$ , i.e., slower than the term  $-n\mu\eta_n = -N_n\mu$  in the second term. Therefore the second term in Eq. (6.12) is negligible compared to the first and Eqs. (3.10) and (3.11) hold true.  $\square$

**6.2. Evaluation of  $H_n$ .** Here we focus on the regime  $\liminf N_n/N_n^* > 1$ ; the case  $\limsup N_n/N_n^* \leq 1$  is treated in the proof of Theorem 2.9. In order to apply the Gaussian approximation from Section 5.4, we need to drop the sine and replace  $\operatorname{Im} \varphi(t)$  with  $G(e^t)$ .

**Lemma 6.4.** *Let  $\varepsilon_n \searrow 0$  faster than some power of  $n$ , i.e.,  $n\varepsilon_n^p \rightarrow 0$  for some  $p > 0$ . Then*

$$\sin(n \operatorname{Im} \varphi(t)) \sim n \operatorname{Im} G(e^t)$$

uniformly for  $t \in [0, \varepsilon_n]$ .

*Proof.* We have

$$\operatorname{Im} G(e^t) = \operatorname{Im} e^{\varphi(t)} = e^{\operatorname{Re} \varphi(t)} \operatorname{Im} \varphi(t) = |G(e^t)| \operatorname{Im} \varphi(t). \quad (6.14)$$

We know that

$$\sup_{t \in [0, \varepsilon_n]} |\operatorname{Re} G(e^t) - 1| = O(\varepsilon_n) \rightarrow 0, \quad (6.15)$$

and the imaginary part vanishes faster than any power, in particular

$$\sup_{t \in [0, \varepsilon_n]} |\operatorname{Im} G(e^t)| = O(\varepsilon_n^p) \rightarrow 0. \quad (6.16)$$

It follows that a similar bound applies to  $\operatorname{Im} \varphi(t)$ . As a consequence

$$\begin{aligned} \sup_{t \in [0, \varepsilon_n]} \left| \frac{n \operatorname{Im} \varphi(t) - \sin(n \operatorname{Im} \varphi(t))}{\sin(n \operatorname{Im} \varphi(t))} \right| &= O\left(n^2 \sup_{t \in [0, \varepsilon_n]} |\operatorname{Im} \varphi(t)|^2\right) \\ &= O\left(n^2 \varepsilon_n^{2p}\right) \rightarrow 0, \end{aligned} \quad (6.17)$$

Thus  $\sin(n \operatorname{Im} \varphi(t)) \sim n \operatorname{Im} \varphi(t)$ , uniformly in  $[0, \varepsilon_n]$ . Eq. (6.14) in turn shows  $\operatorname{Im} \varphi(t) \sim \operatorname{Im} G(e^t)$  uniformly in  $[0, \varepsilon_n]$ .  $\square$

The dominant contribution to the Gaussian integral in Lemmas 5.7-5.9 comes from windows of width  $o(t_n)$  around  $t_n$ ; by Remark 5 and Lemma 2.6,  $t_n = O(t_n^*) = O(N_n^*/n) = O(n^{-[1-\alpha]/[2-\alpha]})$ . This latter bound vanishes like some negative power of  $n$ , hence Lemma 6.4 is applicable on the interval contributing most to the Gaussian integrals. Outside we use the inequality  $|\sin(n \operatorname{Im} \varphi(t))| \leq n \operatorname{Im} \varphi(t) = n(1 + o(1)) \operatorname{Im} G(e^t)$ , and we find: For  $N_n \rightarrow \infty$  with  $\liminf_{n \rightarrow \infty} N_n/N_n^* > 1$ , we have

$$H_n^1 \sim \frac{n}{\sqrt{1 - n\sigma^2 q''(\xi_n)}} e^{\Psi_n(t_n)}. \quad (6.18)$$

For  $N_n \rightarrow \infty$  with  $N_n \gg N_n^*$ ,

$$H_n^1 \sim n e^{\Psi_n(t_n)}. \quad (6.19)$$

$H_n^2$  is estimated in the proof of Theorem 2.10 and is not needed in the proof of Theorem 2.11. Finally for  $\liminf N_n/n > 0$  and  $\delta_0 > 0$  small enough as in Lemma 5.4,

$$\frac{1}{\pi} \int_0^{\delta_0} e^{n\operatorname{Re} \varphi(t) - mt} \sin(n\operatorname{Im} \varphi(t)) dt \sim n e^{\Psi_n(t_n)}. \quad (6.20)$$

Remember that  $\Psi_n(t_n) = \Phi_n(t_n, \xi_n) = -f_{nr}(x_{nr}) + o(1)$  and  $1 - n\sigma^2|q''(\xi_n)| \sim 1 - n\sigma^2|q''(x_{nr})|$  by the definition of  $\Psi_n$  and Eq. (2.22), so the right-hand sides in Eqs. (6.18)–(6.20) correspond to the relevant contribution in Theorems 2.10 and 2.11.

**6.3. Critical scale: proof of Theorem 2.10.** Let  $N_n \rightarrow \infty$  with  $\liminf N_n/N_n^* > 1$  and  $N_n = O(n^{1/(2-\alpha)})$ . Let  $\eta_n \sim N_n/(n\sigma^2)$  be the solution of (3.4), define  $H_n$  and  $V_n$  as in (3.7), and  $H_n^1$  and  $H_n^2$  as in (6.2). Thus we have

$$\mathbb{P}(S_n = n\mu + N_n) = V_n + H_n^1 + H_n^2. \quad (6.21)$$

By Lemma 6.3, the asymptotics of  $V_n$  is given in terms of the Cramér series as in Eq. (3.11).  $H_n^1$  is evaluated with Lemma 5.7 and Lemma 5.8 as the right-hand side of Eq. (3.22). The proof is complete once we show  $H_n^2 = o(V_n)$ .

On  $(t'_n, \eta_n)$  the function  $\Psi_n(t)$  is increasing by Lemma 5.3,

$$\sup_{t \in (t'_n, \eta_n)} \Psi_n(t) \leq \Psi_n(\eta_n) = \left( n\operatorname{Re} \varphi(\eta_n) - (n\mu + N_n)\eta_n \right) + \psi(\eta_n). \quad (6.22)$$

The term in big parentheses can be reexpressed with the Cramér series and is exactly equal to the exponent in the evaluation of  $V_n$  (see Eqs. (3.10) and (3.11)), while  $\psi(\eta_n) \rightarrow -\infty$ . The prefactor satisfies

$$\sup_{t \in (t'_n, \eta_n)} \sqrt{|\psi''(t)|} \leq \sqrt{|\psi''(t'_n)|} = \exp\left(O(\log t'_n)\right) = \exp\left(O(\log n)\right). \quad (6.23)$$

(remember (5.4) and  $t'_n \geq t_n^*$ , with  $t_n^*$  of the order of  $N_n^*/n \gg 1/\sqrt{n}$ ). On the other hand  $\eta_n \searrow 0$  faster than some power of  $n$ , so by Lemma 4.3 we have for some constant  $C$

$$|\psi(\eta_n)| \gg |\log \eta_n| \geq C \log n \gg \log \sqrt{|\psi''(t'_n)|}, \quad (6.24)$$

whence

$$\psi(\eta_n) + \log n + \sup_{t \in (t'_n, \eta_n)} \log \sqrt{|\psi''(t)|} \rightarrow -\infty. \quad (6.25)$$

In view of (6.22), (3.10) and (3.11), we obtain

$$H_n^2 = \frac{n}{\pi} \int_{t'_n}^{\eta_n} e^{n\operatorname{Re} \varphi(t) - (n\mu + N_n)t} \operatorname{Im} G(e^t) dt = o(V_n). \quad (6.26)$$

which concludes the proof.  $\square$



**6.4. Big jump: proof of Theorem 2.11.** Let  $N_n \rightarrow \infty$  with  $N_n \gg n^{1/(2-\alpha)}$ . Notice that, by Lemma 2.6, we then have  $N_n \gg N_n^*$ . We distinguish two cases.

*Case 1:*  $\liminf_{n \rightarrow \infty} N_n/n > 0$ . Fix  $\delta_0 > 0$  as in Lemma 5.4 and define  $H_n$  and  $V_n$  as in (3.7) but with  $\delta_0$  instead of  $\eta_n$ . The asymptotic behavior of  $H_n$  is given by Eq. (6.20). The proof is complete once we check  $V_n = o(H_n)$ .

$V_n$  is exponentially small in  $n$  by Lemma 6.1. Remembering Eq. (2.22) and (??) we get

$$\begin{aligned} \Psi_n(t_n) &= -f_{nr}(x_{nr}) + o(1) = -q(N_n) + \frac{1}{2}(1 + o(1))n\sigma^2 q'(N_n)^2 + o(1) \\ &\geq -q(N_n) + o(1) \geq -CN_n^\alpha. \end{aligned} \quad (6.27)$$

It follows that  $H_n \sim n \exp(\Psi_n(t_n))$  goes to zero slower than  $\exp(-cn^\alpha)$  for some  $c > 0$ , hence  $V_n = o(H_n)$  and

$$\mathbb{P}(S_n = n\mu + N_n) \sim H_n \sim n e^{-f_{nr}(x_{nr})}. \quad (6.28)$$

If  $n^{1/(2-\alpha)} \ll N_n \ll n$ , set  $\varepsilon_n = t'_n$ . By Remark 5, the critical point  $t'_n \in (t_n^*, \eta_n)$  is bounded from below by some constant times  $\eta_n \sim N_n/(n\sigma^2)$ . We define  $H_n$  and  $V_n$  as in (3.7) but with  $\varepsilon_n$  instead of  $\eta_n$ .  $H_n$  is evaluated as in (6.19), which yields  $H_n \sim n \exp(\Psi_n(t_n))$ ; Eq. (6.27) stays valid.  $V_n$  is estimated by Lemma 6.2, which yields  $V_n = O(\exp(-N_n t'_n/2))$ . Now  $N_n t'_n \rightarrow \infty$  much faster than  $N_n^\alpha$ . Indeed  $t'_n$  is bounded from below by some constant times  $N_n/n$ , hence

$$\frac{N_n^\alpha}{N_n t'_n} = O(nN_n^{\alpha-2}) \rightarrow 0. \quad (6.29)$$

It follows that  $V_n = o(H_n)$  and Eq. (6.28) stays true.  $\square$

**6.5. Small steps: proof of Theorem 2.9.** The proof of Theorem 2.9 requires two more technical lemmas, proven at the end of this section. Remember the function  $R_n(t)$  from (5.40).

**Lemma 6.5.** *If  $\limsup_{n \rightarrow \infty} N_n/N_n^* < 1$ , then  $R'_n > 0$  on  $(0, \eta_n)$  for all sufficiently large  $n$ .*

**Lemma 6.6.** *Let  $N_n \rightarrow \infty$  with  $N_n \sim N_n^*$ . Set  $f_n^*(x) := q(x) + \frac{(N_n^* - x)^2}{2n\sigma^2}$ . Suppose that there are infinitely many  $n \in \mathbb{N}$  for which the equation  $R'_n(t) = 0$  has a solution  $s_n \in (0, \eta_n)$ . Then  $s_n \sim t_n^*$  and*

$$R_n(s_n) = -f_n^*(x_n^*) + o\left(\frac{(N_n^*)^2}{n}\right) + O(\log n).$$

The zero  $s_n$  need not be unique—in case of non-uniqueness the lemma applies to every choice of  $s_n$ .

*Proof of Theorem 2.9.* Let  $N_n \rightarrow \infty$  along  $\sqrt{n} \ll N_n \leq (1 + o(1))N_n^*$ . Let  $\eta_n \sim N_n/(n\sigma^2)$  be the solution of (3.4) and define  $V_n$  and  $H_n$  as in (3.7). By Lemma 6.3, the asymptotic behavior of  $V_n$  is given by Eqs. (3.10) and 3.11, so it remains to verify that  $H_n = o(V_n)$ . We estimate

$$H_n \leq \frac{n\eta_n}{\pi} \sup_{t \in (0, \eta_n)} \sqrt{|\psi''(t)|} e^{\Psi_n(t)} = \frac{n\eta_n}{\pi} \sup_{t \in (0, \eta_n)} e^{R_n(t)}. \quad (6.30)$$

Just as in Lemma 5.6, one checks that  $R_n(t)$  is increasing for small  $t$ . We distinguish two cases.

*Case 1:*  $\liminf_{n \rightarrow \infty} N_n/N_n^* < 1$ . Then Lemma 6.5 shows, for large  $n$ ,

$$\sup_{t \in (0, \eta_n)} R_n(t) \leq R_n(\eta_n) = \left( n \operatorname{Re} \varphi(\eta_n) - (n\mu + N_n)\eta_n \right) + \psi(\eta_n) + \log \sqrt{|\psi''(\eta_n)|} \quad (6.31)$$

and we deduce from (3.10)

$$\begin{aligned} \frac{H_n}{V_n} &\leq \frac{n\eta_n}{\pi} \sqrt{2\pi n\sigma^2} \exp\left(\psi(\eta_n) + \log \sqrt{|\psi''(\eta_n)|}\right) \\ &= \exp\left(\psi(\eta_n) + \log \sqrt{|\psi''(\eta_n)|} + O(\log n)\right) \end{aligned} \quad (6.32)$$

which goes to zero by an estimate analogous to (6.24).

*Case 2:*  $N_n \sim N_n^*$ . If  $R'_n$  reaches its maximum at  $t = \eta_n$ , the estimate (6.32) still applies. If along some subsequence  $(n_j)$ ,  $R'_n$  reaches its maximum at some interior point  $s_n \in (0, \eta_n)$ , then we must have  $R'_n(s_n) = 0$  and by Lemmas 6.6 and 5.1

$$\sup_{t \in (0, \eta_n)} R_n(t) \leq -(1 + \varepsilon + o(1)) \frac{N_n^{*2}}{2n\sigma^2} + O(\log n). \quad (6.33)$$

for some  $\varepsilon > 0$  and all large  $n$ . Since  $V_n = \exp(-(1 + o(1)) \frac{N_n^{*2}}{2n\sigma^2} + O(\log n))$  we get

$$\frac{H_n}{V_n} \leq \exp\left(-(\varepsilon + o(1)) \frac{N_n^{*2}}{2n\sigma^2} + O(\log n)\right). \quad (6.34)$$

By Assumption 2.1 and Lemma 2.6,

$$\frac{1}{n\sigma^2} = |q''(x_n^*)| \gg \frac{\log x_n^*}{x_n^{*2}} \geq \frac{\log n + O(1)}{x_n^{*2}} \quad (6.35)$$

so  $\log n = o(N_n^{*2}/n)$  and the right-hand of (6.34) goes to zero.

We have checked in both cases that  $H_n = o(V_n)$ , which concludes the proof.  $\square$

*Proof of Lemma 6.5.* Suppose that the equation  $R'_n(s_n) = 0$  has a solution  $s_n \in (0, \eta_n)$  for infinitely many  $n$ . Thus  $(s_n)$  may be defined only along some subsequence  $(n_j)$ , which we suppress from the notation. Define  $y_n \sim \psi'(s_n)$  as in (5.43). Proceeding as in Lemma 5.6, we find that  $q'(y_n) \sim (N_n - y_n)/(n\sigma^2)$ . By the convexity of  $q'$  and the definition of  $N_n^*$ , we have

$$(1 + o(1)) \frac{N_n - y_n}{n\sigma^2} = q'(y_n) \geq \frac{N_n^* - y_n}{n\sigma^2} = \frac{N_n^* - N_n}{n\sigma^2} + \frac{N_n - y_n}{n\sigma^2} \quad (6.36)$$

hence  $N_n^* - N_n \leq o(N_n - y_n) = o(N_n)$  and  $\limsup N_n^*/N_n \leq 1$  i.e.  $\liminf N_n/N_n^* > 1$ .

So if  $\limsup_{n \rightarrow \infty} N_n/N_n^* < 1$ , we must have  $R'_n \neq 0$  on  $(0, \eta_n)$  except possibly for finitely many  $n$ . From the proof of Lemma 5.6 we know that  $\lim_{t \searrow 0} R'_n(t) = \infty$  for all  $n \in \mathbb{N}$ , and Lemma 6.5 follows.  $\square$

*Proof of Lemma 6.6.* For  $t \in (0, \eta_n)$  with  $\eta_n \sim N_n/(n\sigma^2)$  we have

$$\begin{aligned} R'_n(t) &= n(\operatorname{Re} \varphi'(t) - \mu) - N_n + \psi'(t) + \frac{1}{2} \frac{\psi^{(3)}(t)}{\psi''(t)} \\ &= n\sigma^2 t - (N_n^* - \xi(t)) + o(nt) + (N_n^* - N_n) + O\left(\frac{1}{t}\right) \\ &= n\sigma^2 t - (N_n^* - \xi(t)) + o(N_n^*) + o(\xi(t)). \end{aligned} \quad (6.37)$$

In terms of the dual variable  $\xi = \xi(t) = \psi'(t)$ , the equation  $R'_n(t) = 0$  reads

$$n\sigma^2 q'(\xi) - (N_n^* - \xi) = o(N_n^*) + o(\xi). \quad (6.38)$$

At  $\xi = x_n^*$ , the left-hand side of (6.38) vanishes by definition of  $N_n^*$ . It follows that

$$n\sigma^2 q'(\xi) - (N_n^* - \xi) = \int_{x_n^*}^{\xi} (1 + n\sigma^2 q''(x)) dx. \quad (6.39)$$

Fix  $\gamma > 0$ . On  $((1 + \gamma)x_n^*, \infty)$ , we have  $1 - n\sigma^2 |q''(x)| \geq 1 - n\sigma^2 |q''((1 + \gamma)x_n^*)| =: c_\gamma > 0$  and for all  $x \geq x_n^*$ , the integrand is non-negative. As a consequence,

$$n\sigma^2 q'(\xi) - (N_n^* - \xi) \geq c_\gamma (\xi - (1 + \gamma)x_n^*) \quad (6.40)$$

for all  $\xi \geq (1 + \gamma)x_n^*$ . Suppose that Eq. (6.38) has a solution  $\xi'_n$  with  $\xi'_n \geq (1 + 2\gamma)x_n^*$  along some subsequence. Then  $N_n^* = O(x_n^*) = o(\xi'_n)$  and

$$c_\gamma \gamma \xi'_n \leq c_\gamma (\xi'_n - (1 + \gamma)x_n^*) = o(\xi'_n), \quad (6.41)$$

which is a contradiction (remember  $\xi'_n \geq x_n^* \rightarrow \infty$ ). It follows that for every  $\gamma > 0$ , there are at most finitely many  $n$  for which  $\xi'_n > (1 + 2\gamma)x_n^*$ , hence  $\limsup \xi'_n/x_n^* \leq 1$ .

The case  $\psi'(\eta_n) \leq \xi'_n \leq (1 - 2\gamma)x_n^*$  is treated in an analogous fashion, based on two observations: first,  $n\sigma^2 q''(x) + 1 \leq -c_\gamma < 0$  for all  $x \in (a, (1 - \gamma)x_n^*)$  and some  $c_\gamma > 0$ . Second, since  $\eta_n$  is of the order of  $N_n^*/n$  i.e., of the order of  $t_n^* \sim q'(x_n^*)$ , the estimate (2.4) shows that  $\xi'_n$  is bounded from below by some constant times  $N_n^*$ , i.e., we still have  $N_n^* = O(\xi'_n)$  and  $\xi'_n \rightarrow \infty$ . We find  $\liminf \xi'_n/x_n^* \geq 1$ , hence altogether  $\lim \xi'_n/x_n^* = 1$ . This applies in particular to  $\xi'_n := \psi'(s_n)$  i.e.  $s_n = q'(\xi'_n)$ . Eq. 2.4 and  $t_n^* \sim q'(x_n^*)$  (Lemma 5.2) yield  $\lim s_n/t_n^* = 1$ .

By Remark 5 and Lemma 2.6 we have  $s_n \sim t_n^* \geq \text{const } N_n^*/n \gg 1/\sqrt{n}$  hence  $\log s_n = O(\log n)$  and by (5.4)

$$\log |\psi''(s_n)| = O(\log s_n) = O(\log n). \quad (6.42)$$

Furthermore for  $t \in (0, \eta_n)$

$$\begin{aligned} \Psi_n(t) &= \frac{1}{2}n(\sigma^2 t^2 + o(\eta_n^2)) - N_n t + t\xi(t) - q(\xi(t)) \\ &= \left\{ -q(\xi(t)) + \frac{1}{2}n\sigma^2 t^2 - (N_n^* - \xi(t))t \right\} + O((N_n^* - N_n)\eta_n) + o(n\eta_n^2) \end{aligned} \quad (6.43)$$

The two remainders are  $o((N_n^*)^2/n)$  by our choice of  $N_n$ . Write  $g_n(t)$  for the term in curly braces. At  $t = q'(x_n^*) = [N_n^* - x_n^*]^2/[n\sigma^2]$  we have  $\xi(t) = x_n^*$ ,  $g_n(t) = -f_n^*(x_n^*)$  and  $g'_n(t) = 0$ . Moreover for  $t \in (0, \eta_n)$

$$g''_n(t) = n\sigma^2 + \xi'(t) = n\sigma^2 - \frac{1}{|q''(\xi(t))|} = O(n\sigma^2). \quad (6.44)$$

Here we have used that  $t \leq \eta_n$  implies that  $\xi(t)$  is bounded from below by a constant times  $N_n^*$  or equivalently,  $x_n^*$  and therefore  $|q''(\xi(t))|$  is bounded from below by some constant times  $1/(n\sigma^2)$ . We deduce

$$|g_n(s_n) + f_n^*(x_n^*)| \leq \frac{1}{2}(t_n - q'(x_n^*))^2 \sup_{0, \eta_n} |g''_n| = o(t_n^2 n\sigma^2) = o(n\eta_n^2) = o\left(\frac{(N_n^*)^2}{n}\right). \quad (6.45)$$

The estimate on  $R_n(s_n)$  follows from Eqs. (6.42), (6.43) and (6.45).  $\square$

## 7. EXTENSIONS

The methods described in this paper depend on being able to extend weight functions defined on integers to analytic functions on some region in the complex plane (Assumption 2.2). Here we explain how one might proceed when the existence of an analytic extension is not a priori given.

**7.1. Probability. Log-convexity.** A first option, for weights satisfying Assumption 2.1, is to proceed as in [17]. At the end of the section we comment on the role of log-convexity and a possibly broader perspective. As shown in [17, Lemma 2.1], Assumption 2.1 is enough to guarantee that for every  $\delta > 0$ , as  $n \rightarrow \infty$ ,

$$p(n) \sim \int_0^\delta e^{-nt} \sqrt{2\pi|\psi''(t)|} e^{\psi(t)} dt, \quad (7.1)$$

a statement in some sense inverse to Theorem 4.4. To simplify formulas, consider laws on  $\mathbb{N}_0$  rather than  $\mathbb{N}$ . Let us pretend for a moment that  $p(n)$  is in fact completely monotone, i.e.,

$$p(n) = \int_0^\infty e^{-n\lambda} g(\lambda) d\lambda \quad (7.2)$$

for some non-negative integrable function  $g(\lambda)$ . This holds true for the stretched-exponential case [19]. The probabilistic interpretation is that the law  $p$  is a mixture of geometric laws. Eq. (7.2) can be adopted as a definition for  $p(\xi)$  when  $\operatorname{Re} \xi \geq 0$  and thus provides an analytic continuation of the probability weights. Moreover we have, for  $|z| \leq 1$ ,

$$G(z) = \int_0^\infty \frac{g(\lambda)}{1 - ze^{-\lambda}} d\lambda, \quad (7.3)$$

and the right-hand side provides an analytic continuation for  $G(z)$  from the unit disk to the slit plane  $\mathbb{C} \setminus [1, \infty)$ . For  $z = e^{t+i\varepsilon}$  with  $t \in \mathbb{R}$  and  $\varepsilon \rightarrow 0$  we note

$$\frac{1}{1 - ze^{-\lambda}} = \frac{1}{\lambda - t - i\varepsilon} \frac{\lambda - t}{1 - e^{t-\lambda}} (1 + O(\varepsilon)). \quad (7.4)$$

The function  $h_t(\lambda) = (\lambda - t)/(1 - e^{t-\lambda})$  is smooth and satisfies  $h_t(t) = 1$ . Assuming that  $g$  is smooth—which by (7.2) and Paley-Wiener theory directly translates into the decay of  $p(\xi)$  along lines  $\operatorname{Re} \xi = \text{const}$ —we may expect from Plemelj's formula that

$$\lim_{\varepsilon \searrow 0} G(e^{t+i\varepsilon}) = \lim_{\varepsilon \searrow 0} \int_0^\infty \frac{h_t(\lambda)g(\lambda)}{\lambda - t - i\varepsilon} d\lambda = i\pi g(t) + \mathcal{P} \left( \int \frac{g(\lambda)h_t(\lambda)}{\lambda - t} d\lambda \right) \quad (7.5)$$

where the second term is the Cauchy principal value of the possibly divergent integral. So in particular, we expect that the limit  $G(e^t) = \lim_{\varepsilon \searrow 0} G(e^{t+i\varepsilon})$  exists, moreover

$$\operatorname{Im} G(e^t) = \pi g(t), \quad p(n) = \frac{1}{\pi} \int_0^\infty e^{-nt} \operatorname{Im} G(e^t) dt \quad (7.6)$$

and we recover Eq. (4.6) and some of the results from Section 4 by starting from the relation (7.2) rather than Assumption 2.2 and the Lindelöf integral. In view of (7.1) we would like to conclude that  $\operatorname{Im} G(e^t) \sim \frac{1}{2} \sqrt{2\pi|\psi''(t)|} e^{\psi(t)}$  as in Theorem 4.4, however doing so would require a kind of Tauberian theorem (we had worked with Assumption 2.3 instead).

In general Assumption 2.1 might not be enough to guarantee a representation as in (7.2), so following [17] one may take the approximate relation (7.1) as a starting point. We introduce a truncation sequence  $(y_n)$  and define

$$\tilde{p}_n(k) := \begin{cases} p(k), & k \leq y_n, \\ \int_0^\delta e^{-kt} \sqrt{2\pi|\psi''(t)|} e^{\psi(t)} dt, & k > y_n. \end{cases} \quad (7.7)$$

Notice that  $\sum_{k=1}^\infty \tilde{p}_n(k)$  might be slightly different from 1. By the considerations sketched above, for each fixed  $n$ , the associated generating function  $\tilde{G}_n(z) = \sum_{k=0}^\infty \tilde{p}_n(k)z^k$  extends to the slit plane  $\mathbb{C} \setminus [1, \infty)$ , and coefficients can be expressed by contour integrals as we have done. The principal idea is then to show that for a judicious choice of  $(y_n)$  coefficients of large powers  $[z^m]G(z)^n$  behave approximately like  $[z^m]\tilde{G}_n(z)^n$ . The necessary estimates have been carried out in [17] in a slightly different context—inversion formulas for Laplace transforms instead of contour integrals of generating functions—but it should be possible to adapt them to the present context.

Key to the present approach is a property we might term “asymptotic complete monotonicity”, where the probability weights are asymptotically equivalent to the Laplace transform of some integrable non-negative function. This property is closely related to log-convexity: every completely monotone function is log-convex [14, Lemma 4.3]. The converse is, in general false [14], but Eq. (7.1) shows that if log-convexity is strengthened to Assumption 2.1, then at least  $p(n)$  is asymptotically equivalent to a completely monotone sequence. It would be interesting to extend the method to more general log-convex sequences, under weaker conditions than Assumption 2.1.

**7.2. Complex analysis.** Another option, especially when the weights may not be log-convex, is to try to obtain an analytic interpolation based on the Paley-Wiener theorem (known as *the Sampling Formula*).

**Theorem 7.1.** [6] *If  $f$  is an entire analytic function of type  $\leq T$  whose restriction to  $\xi \in \mathbb{R}$  is  $L^2$ , then*

$$\begin{aligned} f(\xi) &= \sum_{n=-\infty}^{\infty} \frac{1}{T} f\left(\frac{n\pi}{T}\right) \frac{\sin[(\xi - n\pi/T)T]}{\xi - n\pi/T} \quad \text{and} \\ \|f\|_2^2 &= \sum_{n=-\infty}^{\infty} \frac{\pi}{T} \left| f\left(\frac{n\pi}{T}\right) \right|^2. \end{aligned}$$

To see how to use this, let  $G(z)$  be a probability generating function with probability weights  $p(n) = c_n$ ,  $n \in \mathbb{N}$ . Then one may define the analytic symbol for this case by

$$p(\xi) \doteq \sum_{n=1}^{\infty} c_n \frac{\sin[\pi(\xi - n)]}{\pi(\xi - n)} = \frac{1}{2} \sum_{n=-\infty}^{\infty} c_n \frac{\sin[\pi(\xi - n)]}{\pi(\xi - n)} \quad (7.8)$$

where  $c_{-n} = c_n$  for  $n \neq 0$  and  $c_0 = 0$ . Equivalently,  $p(\xi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{it}) e^{-i\xi t} dt$ . It is straightforward to check that  $p(\xi)$  so defined is an entire analytic function of type  $\leq \pi$ . Moreover, since  $(c_n)$  is square summable, the sampling theorem establishes that the restriction of  $p(\xi)$  to the real line gives an  $L^2$  function with  $L^2$

norm

$$\|p\|_2^2 = \frac{1}{4} \sum_{n=-\infty}^{\infty} c_n^2 = \frac{1}{2} \sum_{n=1}^{\infty} c_n^2. \quad (7.9)$$

In order to apply such an analytic representation to the framework outlined in this paper one needs to assume that Assumption 2.2(ii) holds on some appropriate contours in the complex plane which is needed to ensure that the Lindelöf integral exists and has an analytic continuation to slit complex plane. The above representation, in terms of entire analytic functions is probably too rigid to apply in the case of classes such as regularly varying hazard functions which extend the specific case of a sub-exponential distribution. We are currently exploring other versions of interpolation on restricted domains that may provide more flexibility in extending our results.

#### APPENDIX A. PROOFS OF LEMMAS 2.14 AND 2.15

*Proof of Lemma 2.14.* The proof of Assumption 2.1 is straightforward and left to the reader. The function  $q(\xi) = \xi^\alpha$  is analytic in  $\operatorname{Re} \xi > 0$  and  $p(\xi) = \exp(-\xi^\alpha)$  satisfies, for all  $k \in \mathbb{N}$ ,

$$|\xi^k p(\xi)| = |\xi|^k e^{-|\xi|^\alpha \cos(\alpha \arg(\xi))} \leq |\xi|^k e^{-|\xi|^\alpha \cos(\alpha \pi/2)}. \quad (A.1)$$

Since  $\alpha \in (0, 1)$ , we have  $\cos \frac{\alpha \pi}{2} > 0$  and Eq. (A.1) shows that  $|\xi^k p(\xi)|$  is integrable along  $\operatorname{Re} \xi = 1/2$  and that  $p(\xi)$  grows slower than any exponential  $\exp(\varepsilon|\xi|)$ . This proves Assumption 2.2.

The equation  $q''(x_n^*) = -1/(n\sigma^2)$  can be solved explicitly.  $N_n^*$  and  $N_n^{**}$  are best determined with the scaling relation (2.24). They have already been determined in [15], we omit the proof. For the insensitivity scale, we notice that

$$n\sigma^2 q'(N_n)^2 = n\sigma^2 \alpha^2 N_n^{2\alpha-2} \quad (A.2)$$

which goes to zero if and only if  $N_n \gg n^{-1/(2-2\alpha)}$ .  $\square$

*Proof of Lemma 2.15.* The function  $q(x) = -\log c + (\log x)^\beta$  is clearly smooth on  $(1, \infty)$ . Then  $q'(x) = \beta(\log x)^{\beta-1}/x$  and as  $x \rightarrow \infty$ ,

$$q''(x) \sim -\frac{\beta(\log x)^{\beta-1}}{x^2}, \quad q^{(3)}(x) \sim \frac{2\beta(\log x)^{\beta-1}}{x^3}. \quad (A.3)$$

Assumption 2.1 is easily checked. For Assumption 2.2 we note  $q(\xi) = c + (\log \xi)^\beta$  is analytic in  $\operatorname{Re} \xi > 1$ . Fix  $b > 1$  and write  $\xi = r \exp(i\theta)$ . As  $|\xi| \rightarrow \infty$  along  $\operatorname{Re} \xi = b$  i.e.  $\xi = b + iy$ , the argument  $\theta$  goes to  $\pm\pi/2$  and we have

$$\operatorname{Re} (\log |\xi| + i\theta)^\beta = \operatorname{Re} \left( \log |y| + \frac{1}{2} \log \left( 1 + \frac{b^2}{y^2} \right) + i\theta \right)^\beta \quad (A.4)$$

$$= (\log |y|)^\beta + o(1), \quad (A.5)$$

conditions(i) and (ii) in Assumption 2.2 are easily checked. Assumption 2.3(iii) follows from a computation similar to (A.3). Set  $y_0 = \sqrt{r^2 - b^2}$ . We have for  $\xi = b + iy$ ,  $y \geq y_r$ , uniformly in  $r$ ,

$$\begin{aligned} \operatorname{Re} q(\xi) - \operatorname{Re} q(z_r) &= (\log y)^\beta - (\log y_0)^\beta + o(1) \\ &\geq \beta(\log y_0)^{\beta-1} \log \frac{y}{y_0} + o(1) \end{aligned} \quad (A.6)$$

hence

$$\begin{aligned} \int_{y_r}^{\infty} e^{-\operatorname{Re} q(b+iy)} dy &\leq e^{-\operatorname{Re} q(b+iy_0)+o(1)} \int_1^{\infty} e^{-(\log y_0)^{\beta-1} \log s} y_0 ds \\ &\sim e^{-\operatorname{Re} q(b+iy_0)} \frac{y_0}{(\log y_0)^{\beta-1}} = e^{-\operatorname{Re} q(b+iy_0)+O(\log r)}, \end{aligned} \quad (\text{A.7})$$

which proves Assumption 2.3(i). Next let  $\xi \in \mathbb{C}$  with  $\operatorname{Re} \xi > 1$ , write  $\xi = r \exp(i\theta)$ , then

$$\xi q'(\xi) = \beta(\log \xi)^{\beta-1} = \beta(\log r + i\theta)^{\beta-1} = q'(r) \left(1 + \frac{i\theta}{\log r}\right)^{\beta-1} \quad (\text{A.8})$$

and for large  $r$  and  $\theta \in (0, \pi/2)$ ,

$$\frac{\operatorname{Im} \xi q'(\xi)}{\operatorname{Im} \xi q'(r)} \sim \frac{\beta-1}{r} \frac{\theta}{\log r \sin \theta} \rightarrow 0 \quad (\text{A.9})$$

and so Assumption 2.3(ii) holds.

We now turn to the asymptotic behavior of the sequences  $x_n^*$ ,  $N_n^*$  and  $N_n^{**}$ . Since  $q''(x_n) = -\frac{1}{n\sigma^2}$ , it is clear that  $x_n \rightarrow \infty$  as  $n \rightarrow \infty$ . The equation is

$$\frac{1}{n\sigma^2} \sim \frac{\beta(\log x_n^*)^{\beta-1}}{(x_n^*)^2}. \quad (\text{A.10})$$

Consequently,

$$\begin{aligned} (x_n^*)^2 &\sim \beta n \sigma^2 \left(\frac{1}{2} \log(x_n^*)\right)^{\beta-1} \\ &\sim \beta n \sigma^2 \left(\frac{1}{2} \log n\right)^{\beta-1} \left(1 + \frac{\log \beta \sigma^2 + (\beta-1) \log \log x_n^*}{\log n}\right)^{\beta-1}. \end{aligned} \quad (\text{A.11})$$

The last bracket is asymptotically equal to 1 and we get the expression for  $x_n^*$ . Next, we have from (2.10)

$$N_n^* = x_n^* + n\sigma^2 \frac{\beta(\log x_n^*)^{\beta-1}}{x_n^*} \sim 2x_n^*. \quad (\text{A.12})$$

The last asymptotics follows from (A.10).

We now turn to  $N_n^{**}$ . It is asymptotically given by the solution of the equations

$$\frac{N_n^2}{2n\sigma^2} = q(x_n) + \frac{(N_n - x_n)^2}{2n\sigma^2}, \quad (\text{A.13})$$

$$q'(x_n) = \frac{N_n - x_n}{n\sigma^2}. \quad (\text{A.14})$$

Eq. (A.14) is equivalent to

$$x_n^2 - N_n x_n + \beta n \sigma^2 (\log x_n)^{\beta-1} = 0. \quad (\text{A.15})$$

The relevant solution is

$$\begin{aligned} x_n &= \frac{1}{2} \left( N_n + \sqrt{N_n^2 - 4\beta n \sigma^2 (\log x_n)^{\beta-1}} \right) \\ &= N_n \left( 1 - \frac{\beta n \sigma^2}{N_n^2} (\log x_n)^{\beta-1} (1 + o(1)) \right). \end{aligned} \quad (\text{A.16})$$

It follows that  $N_n - x_n \sim \frac{\beta n \sigma^2}{N_n} (\log x_n)^{\beta-1}$ . We insert this in (A.13); using  $\log x_n \sim \log N_n$ , we get

$$\frac{N_n^4}{2n\sigma^2} - N_n^2 (\log N_n)^\beta - \frac{1}{2} \beta^2 n \sigma^2 (\log N_n)^{2\beta-2} = o(1). \quad (\text{A.17})$$

The relevant solution is

$$\begin{aligned}
N_n^2 &\sim n\sigma^2 \left[ (\log N_n)^\beta + \sqrt{(\log N_n)^{2\beta} + \beta^2 (\log N_n)^{2\beta-2}} \right] \\
&\sim 2n\sigma^2 \left( \frac{1}{2} \log N_n^2 \right)^\beta \\
&\sim 2^{1-\beta} n\sigma^2 (\log n + \log 2\sigma^2 + \beta \log \log N_n)^\beta.
\end{aligned} \tag{A.18}$$

Only the term  $\log n$  matters in the last bracket and the result follows.

The last part of the lemma on insensitivity sequence is shown as in [4, Section 8.3] the proof is therefore omitted.  $\square$

#### APPENDIX B. BIVARIATE HESSIAN

As explained in Step 4 of the proof outline, the Hessian at  $(t_n, \xi_n)$  has determinant  $-1 + o(1)$  and is a saddle point of  $\Phi_n(t, \xi)$ , considered as a function of two real variables  $t, \xi > 0$ . In order to get rid of off-diagonal elements in the Hessian and to give all eigenvalues the same sign, we take complex  $\xi$  and reparametrize, as sketched in Step 5.

**Lemma B.1.** *Let  $\xi(t)$  be the unique solution of  $(\partial_\xi \Phi_n)(t, \xi) = t - q'(\xi) = 0$ . Set*

$$F_n : (0, \infty) \times \mathbb{R} \rightarrow \mathbb{C}, \quad F_n(t, s) = \Phi_n(t, \xi(t) + is).$$

Then  $(\nabla F_n)(t_n, 0) = 0$ ,

$$\text{Hess } F_n(t, 0) = \begin{pmatrix} \beta(t) & 0 \\ 0 & q''(\xi(t)) \end{pmatrix}, \quad \beta(t) = -\frac{\det(\text{Hess } \Phi_n)(t, \xi(t))}{q''(\xi(t))}.$$

Note that  $\xi_n = \xi(t_n)$ , so as  $n \rightarrow \infty$

$$\det \text{Hess } F_n(t_n, 0) = \beta(t_n) q''(\xi_n) = -\det(\text{Hess } \Phi_n)(t_n, \xi_n) = 1 + o(1). \tag{B.1}$$

*Proof.* We have

$$\begin{aligned}
\partial_t F_n(t, s) &= (\partial_t \Phi_n)(t, \xi(t) + is) + (\partial_\xi \Phi_n)(t, \xi(t) + is) \xi'(t), \\
\partial_s F_n(t, s) &= i \partial_\xi \Phi_n(t, \xi(t) + is).
\end{aligned} \tag{B.2}$$

At  $t = t_n$ ,  $s = 0$ , we have  $\xi(t) = \xi_n$  and  $(\nabla F_n)(t_n, 0) = \nabla \Phi_n(t_n, \xi_n) = 0$ . For the Hessian, we compute

$$\begin{aligned}
\partial_s^2 F_n(t, 0) &= -\partial_\xi^2 \Phi_n(t, \xi(t)) = q''(\xi(t)), \\
\partial_t \partial_s F_n(t, 0) &= i(\partial_t \partial_\xi \Phi_n)(t, \xi(t)) + i \partial_\xi^2 \Phi_n(t, \xi(t)) \\
\partial_t^2 F_n(t, 0) &= \frac{d^2}{dt^2} \Phi_n(t, \xi(t)) = \beta(t).
\end{aligned} \tag{B.3}$$

By definition of  $\xi(t)$ ,

$$0 = \frac{d}{dt} \partial_\xi \Phi_n(t, \xi(t)) = (\partial_t \partial_\xi \Phi_n)(t, \xi(t)) + \partial_\xi^2 \Phi_n(t, \xi(t)) \xi'(t). \tag{B.4}$$

It follows that  $\partial_t \partial_s F_n(t, 0) = 0$ , and

$$\begin{aligned}
\beta(t) &= \frac{d^2}{dt^2} \Phi_n(t, \xi(t)) = \frac{d}{dt} (\partial_t \Phi_n)(t, \xi(t)) \\
&= (\partial_t^2 \Phi_n)(t, \xi(t)) + (\partial_t \partial_\xi \Phi_n)(t, \xi(t)) \xi'(t).
\end{aligned} \tag{B.5}$$

We solve for  $\xi'(t)$  in Eq. (B.4), insert into Eq. (B.5), and obtain the formula for  $\beta(t)$ .  $\square$



**Acknowledgement.** The authors wish to thank the Laboratoire Jean Dieudonné of the University of Nice and the Institut Henri Poincaré (during the program of the spring 2013 organized by M. Esteban and M. Lewin) for the kind hospitality and for the opportunity to discuss this project. S. J. thanks V. Wachtel for pointing out S. V. Nagaev’s article [17].

## REFERENCES

- [1] R. Arratia, A. D. Barbour, S. Tavaré, *Logarithmic combinatorial structures*, EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich (2003)
- [2] I. Armendáriz, S. Grosskinsky, M. Loulakis, *Zero-range condensation at criticality*, Stoch. Proc. Appl. 123, 3466–3496 (2013)
- [3] E. A. Bender, *Asymptotic Methods in Enumeration*, SIAM Rev. 16, 485–515 (1974)
- [4] D. Denisov, A. B. Dieker, V. Shneer, *Large deviations for random walks under subexponentiality: the big-jump domain*, Ann. Probab. 36, 1946–1991 (2008)
- [5] G. Doetsch, *Theorie und Anwendung der Laplace-Transformationen*. Springer, Berlin (1937)
- [6] H. Dym, H. P. McKean, *Gaussian Processes, function theory and the inverse spectral problem*. Probability and Mathematical Statistics, Vol. 31. Academic Press (Harcourt Brace Jovanovich, Publishers), New York-London (1976).
- [7] P. Embrechts, C. Klüppelberg, T. Mikosch, *Modelling extremal events. For insurance and finance*, Applications of Mathematics (New York), 33. Springer-Verlag, Berlin (1997)
- [8] N.M. Ercolani, S. Jansen, D. Ueltschi, *Random partitions in statistical mechanics*, Electron. J. Probab. 19, no. 82, 1-37 (2014)
- [9] W. Feller, *An Introduction to Probability theory and its Applications*, Vol. 2, John Wiley & Sons, Inc., New York-London-Sydney (1966)
- [10] P. Flajolet, S. Gerhold, B. Salvy, *Lindelöf representations and (non-)holonomic sequences*, Electr. J. Combin. 17, #R3 (2010)
- [11] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge (2009)
- [12] I. A. Ibragimov, Y. U. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing Groningen (1971)
- [13] E. Lindelöf, *Le calcul des résidus et ses applications à la théorie des fonctions*, Collection de monographies sur la théorie des fonctions, publiée sous la direction de M. Émile Borel. Gauthier-Villars, Paris, 1905. Reprinted by Gabay, Paris (1989)
- [14] M. Merkle, *Completely monotone functions: a digest*. Analytic number theory, approximation theory, and special functions, 347364, Springer, New York (2014)
- [15] A. V. Nagaev, *Local limit theorems with regard to large deviations when Cramér’s condition is not satisfied*, Litovsk. Mat. Sb. 8, 553–579 (1968). Selected Transl. in Math. Statist. and Probability 11, 249–278 (1973)
- [16] A. V. Nagaev, *Letter to the editor: A remark concerning the article “Local limit theorems with regard to large deviations”*, Litovsk. Mat. Sb. 10, 206 (1970); Selected Transl. in Math. Statist. and Probability 11, 279 (1973)
- [17] S. V. Nagaev, *Large deviations for sums of random variables*, Transactions of the Sixth Prague Conference on Information Theory, Statistical Decision Functions, Random Processes (Tech. Univ., Prague, 1971), 657–674. Academia, Prague (1973)
- [18] S. V. Nagaev, *Large deviations of sums of independent random variables*, Ann. Probab. 7, 745–789 (1979)
- [19] H. Pollard, *The Representation of  $e^{-x^\lambda}$  as a Laplace Integral*, Bull. Am. Math. Soc., **52**, 908–910 (1946)
- [20] L.V. Rozovskii, *Probabilities of large deviations on the whole axis*, Teor. Veroyatnost. i Primenen. 38, 79–109 (1993). Transl. in Theory Probab. Appl. 38, 53–79 (1993)