THE ENTROPIC ERDŐS-KAC LIMIT THEOREM

S. G. BOBKOV[1,4], G. P. CHISTYAKOV[2,4], AND H. KÖSTERS[3,4]

Abstract. We prove entropic and total variation versions of the Erdős-Kac limit theorem for the maximum of the partial sums of i.i.d. random variables with densities.

1. Introduction

Let $X_1, X_2, X_3, \ldots$ be i.i.d. random variables with mean $\mathbb{E}X_1 = 0$ and variance $\mathbb{E}X_1^2 = 1$. Let $S_n := \sum_{k=1}^n X_k$ denote the associated partial sums, and let $\overline{S}_n := \max_{k=1,\ldots,n} S_k$ denote their maxima.

The classical central limit theorem states that

$$S_n/\sqrt{n} \Rightarrow Z \qquad \text{as } n \to \infty\,, \tag{1.1}$$

where $\Rightarrow$ denotes convergence in distribution and Z is a standard normal random variable. Barron [Ba] established an entropic version of this result, the so-called *entropic central limit theorem.* To formulate it, let us introduce some notation. Let $L(x) := x \log x$ for $x > 0$, and let $L(0) := 0$ by way of continuous extension. If X and Y are real random variables with densities p and ψ, respectively, and such that the distribution of X is absolutely continuous with respect to that of Y, the relative entropy of X with respect to Y is defined by

$$D(X\,|\,Y) := \int_{\{\psi(x)>0\}} L\left(\frac{p(x)}{\psi(x)}\right) \psi(x)\, dx\,. \tag{1.2}$$

(In case X does not have a density or the distribution of X is not absolutely continuous with respect to that of Y, put $D(X\,|\,Y) := \infty$.) The entropic central limit theorem by Barron [Ba] now states that

$$D(S_n/\sqrt{n}\,|\,Z) \to 0 \qquad \text{as } n \to \infty \tag{1.3}$$

if and only if $D(S_{n_0}\,|\,Z) < \infty$ for some $n_0 \in \mathbb{N}$.

Date: December 20th, 2012.

2000 *Mathematics Subject Classification.* 60E.

Key words and phrases. Entropy, entropic distance, Erdős-Kac limit theorem.

1) School of Mathematics, University of Minnesota, USA;
Email: bobkov@math.umn.edu.

2) Department of Mathematics, Bielefeld University, Germany;
Email: chistyak@math.uni-bielefeld.de.

3) Department of Mathematics, Bielefeld University, Germany;
Email: hkoesters@math.uni-bielefeld.de.

4) Research supported by CRC 701.
S.B. was also supported by DMS-1106530 and Simons Fellowship.

Let us note that by the Pinsker-Csiszár-Kullback inequality [Pi, Cs, Ku, FHT], we have

$$D(X \,|\, Z) \geq \tfrac{1}{2}\,(d_{TV}(X,Z))^2\,, \tag{1.4}$$

where d_{TV} is the total variation distance between (the distributions of) X and Z. Thus, $D(X \,|\, Z)$ represents a strong measure of distance between (the distributions of) X and Z, and (1.3) implies $d_{TV}(S_n/\sqrt{n}, Z) \to 0$ as $n \to \infty$. In particular, (1.1) follows from (1.3).

Furthermore, the relative entropy $D(X \,|\, Z)$ has an interesting interpretation from the viewpoint of information theory. Let us recall that for a random variable X with finite second moment, the entropy (also called Shannon entropy or differential entropy) is defined by

$$h(X) := -\int_{-\infty}^{+\infty} L(p(x))\,dx\,,$$

where $p(x)$ is the density of X. (If X does not have a density, put $h(X) := -\infty$.) Let X be a random variable with second moment $\mathbb{E}X^2 = \sigma^2$. Then it is well known that

$$h(\sigma Z) - h(X) = D(X \,|\, \sigma Z) \geq 0\,,$$

with equality if and only if X and σZ have the same distribution. Thus, the distribution of σZ (i.e. the centered normal distribution with second moment σ^2) maximizes entropy among all probability measures with the same second moment. Also, it is easy to see that for any $\tau > 0$,

$$D(X \,|\, \tau Z) = -h(X) + \tfrac{1}{2}\log(2\pi\tau^2) + \tfrac{1}{2}\sigma^2/\tau^2\,,$$

the right-hand side being minimized for $\tau = \sigma$. Thus, $D(X \,|\, \sigma Z)$ may be interpreted as a measure of deviation of the distribution of X from the class of centered normal distributions.

Barron's result has sparked much further research, see e.g. Artstein, Ball, Barthe and Naor [ABBN], Johnson and Barron [JB], Johnson [Jo], Bobkov, Chistyakov and Götze [BCG3, BCG4]. In view of these results, it seems natural to ask whether other results on convergence in distribution admit an entropic formulation as well. We investigate this question for the maxima of sums of i.i.d. random variables.

In this context the analogue of the classical central limit theorem is given by the Erdős-Kac limit theorem:

Theorem 1.1 (Erdős-Kac [EK]). *Suppose that the i.i.d. random variables X_j have mean 0 and variance 1. Then*

$$\overline{S}_n/\sqrt{n} \Rightarrow |Z| \qquad \text{as } n \to \infty, \tag{1.5}$$

where $\Rightarrow$ denotes convergence in distribution and Z is a standard normal random variable.

Further results in this direction were obtained, inter alia, by Nagaev [N1, N2], Aleshkyavichene [Al], and Naudziuniene [Na]. The aim of this paper is to establish an entropic variant of the Erdős-Kac theorem. Let X be a random variable such that $\mathbb{P}(X > 0) > 0$, and let Y be a positive random variable with density ψ. Then the distribution of X conditioned to be positive is given by the probability measure $\mathbb{P}(X \in \cdot \mid X > 0)$ on the positive half-line, and the relative entropy of X conditioned to be positive with respect to Y is defined by

$$D_+(X \mid Y) := D(\widetilde{X} \mid Y) \tag{1.6}$$

where $\widetilde{X}$ has the same distribution as X conditioned to be positive. In the sequel the random variable Y will always be given by $|Z|$, with density $\varphi_+(x) = \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}$ $(x > 0)$. The distribution of $|Z|$ is commonly called the one-sided (or reflected) standard normal distribution. Our main result is as follows:

Theorem 1.2. *Suppose that the i.i.d. random variables X_j with density have mean* 0 *and variance* 1*. Then*

$$D_+(\overline{S}_n/\sqrt{n} \mid |Z|) \to 0 \qquad \textit{as } n \to \infty \tag{1.7}$$

if and only if

$$D_+(X_1 \mid |Z|) < \infty\,. \tag{1.8}$$

Remark. The assumption that the X_j have a density is only for convenience and can be omitted. Note, however, that (1.8) implies that the X_j have a density on the positive half-line.

As is well known, under our moment assumptions,

$$\mathbb{P}(\overline{S}_n \leq 0) = \mathcal{O}(n^{-1/2}) \tag{1.9}$$

(see e.g. [Fe, pp. 414f]). Thus, the distribution of $\overline{S}_n/\sqrt{n}$ restricted to the positive half-line is almost a probability measure for large n, and it follows from the Pinsker-Csiszár-Kullback inequality that (1.7) implies $d_{TV}(\overline{S}_n/\sqrt{n}, |Z|) \to 0$ as $n \to \infty$.

In fact, in this weaker assertion the condition (1.8) does not play a role, since we have:

Theorem 1.3. *Suppose that the i.i.d. random variables X_j with density have mean* 0 *and variance* 1*. Then*

$$d_{TV}(\overline{S}_n/\sqrt{n}, |Z|) \to 0 \qquad \textit{as } n \to \infty.$$

Moreover, similar observations about the interpretation from the viewpoint of information theory as above can be made for probability measures on the positive half-line. Indeed, suppose that X is a positive random variable with finite second moment $\mathbb{E}X^2 = \sigma^2$. Then

$$h(\sigma|Z|) - h(X) = D_+(X \mid \sigma|Z|) \geq 0\,,$$

with equality if and only if X and $\sigma|Z|$ have the same distribution. Thus, the distribution of $\sigma|Z|$ (i.e. the one-sided normal distribution with second moment σ^2) maximizes entropy among all probability measures on the positive half-line with the same second moment. Also, for any $\tau > 0$,

$$D_+(X \mid \tau|Z|) = -h(X) + \tfrac{1}{2}\log(\tfrac{1}{2}\pi\tau^2) + \tfrac{1}{2}\sigma^2/\tau^2\,,$$

the right-hand side being minimized for $\tau = \sigma$. Therefore, similarly as above, $D_+(X \mid \sigma|Z|)$ may be interpreted as a measure of deviation of the distribution of X from the class of one-sided normal distributions.

It is worth noting here that

$$\mathbb{E}(\overline{S}_n/\sqrt{n})^2 = 1 + o(1) \qquad \text{as } n \to \infty. \tag{1.10}$$

(For instance, this follows from Proposition 6.1 below and Equation (1) in [Al].) Combining (1.9) and (1.10), we see that for large n, $\overline{S}_n/\sqrt{n}$ conditioned to be positive has second moment approximately equal to 1, so that the comparison with $|Z|$ in (1.7) is natural.

Finally, let us emphasize the following curious difference between Barron's entropic central limit theorem and our Theorem 1.2. Even if X_1 itself has density, Barron's characterization uses the finiteness of $D(S_{n_0} \,|\, Z)$ for some $n_0 \in \mathbb{N}$ (which may be any natural number), whereas our characterization uses $n_0 = 1$ at once. More precisely, it follows from our proof that $D_+(\overline{S}_{n_0} \,|\, |Z|) < \infty$ for some $n_0 \in \mathbb{N}$ if and only if this is true for $n_0 = 1$. In this respect, maxima of sums behave differently from sums.

Let us introduce some conventions for the rest of the paper. We assume that the random variables X_j are i.i.d. and have a density, mean 0 and variance 1. Unless otherwise indicated, we write p for their density, F for their distribution function and f for their characteristic function. Moreover, for $n = 1, 2, 3, \ldots$, let p_n, F_n, f_n and $\overline{p}_n, \overline{F}_n, \overline{f}_n$ denote the corresponding functions for the random variables S_n and $\overline{S}_n$. (Note that the densities exist due to our assumption that the X_j have a density.) Furthermore, we write p_n^* and $\overline{p}_n^*$ for the densities of the rescaled random variables $S_n/\sqrt{n}$ and $\overline{S}_n/\sqrt{n}$. Unless otherwise indicated, $\mathcal{O}$-bounds and o-bounds hold uniformly in the region under consideration. Finally, $C_1, C_2, C_3, \ldots$ denote positive constants which may depend on the distribution of the X_j and which may change from step to step.

The paper is organized as follows. Section 2 contains some preliminary remarks on relative entropy. Section 3–7 are devoted to the proof of the sufficiency part of Theorem 1.2, while the necessity part of Theorem 1.2 is proved in Section 8. Section 9 contains the proof of Theorem 1.3.

2. Some Remarks on Relative Entropy

Throughout this section, let ψ be a positive probability density on the positive half-line. Given a non-negative measurable function f on the real line, set

$$D(f \,|\, \psi) := \int_0^\infty L\left(\frac{f(x)}{\psi(x)}\right) \psi(x)\, dx\,, \tag{2.1}$$

where $L(x)$ is the function defined in the introduction. By abuse of terminology, we will call $D(f \,|\, \psi)$ *relative entropy* even when f is not a probability density on the positive half-line. Note that in this special case, we have $D(f \,|\, \psi) \geq 0$ by Jensen's inequality. If f is an arbitrary non-negative measurable function, this need not be true anymore, but we have at least $D(f \,|\, \psi) \geq \min\{L(x) : x \geq 0\} = -e^{-1}$.

Let us collect some basic properties of relative entropy which will be used later. (Some of the proofs are straightforward, which is why we omit them.)

Lemma 2.1. *Suppose that α is a positive real number and f is a non-negative measurable function with $\int_0^\infty f(x)\,dx < \infty$. Then*

$$D(\alpha f \,|\, \psi) = \alpha D(f \,|\, \psi) + L(\alpha) \int_0^\infty f(x)\,dx\,.$$

Lemma 2.2. *Suppose that $\alpha_1, \ldots, \alpha_n$ are positive real numbers and $f_1, \ldots, f_n$ are non-negative measurable functions with $\int_0^\infty f_k(x)\,dx < \infty$, $k = 1, \ldots, n$. Then*

$$D\left(\sum_{k=1}^n \alpha_k f_k \,\middle|\, \psi\right) \le \sum_{k=1}^n \alpha_k D(f_k \,|\, \psi) + \left(\log \sum_{k=1}^n \alpha_k\right) \sum_{k=1}^n \alpha_k \int_0^\infty f_k(x)\,dx\,.$$

Lemma 2.3. *Suppose that ψ is decreasing on the positive half-line and that f and g are probability densities on $(0, +\infty)$ and $(-\infty, 0)$, respectively. Then*

$$D(f * g \,|\, \psi) \le D(f \,|\, \psi) + e^{-1}\,.$$

Proof of Lemma 2.3. Since L is a convex function and g is a probability density on $(-\infty, 0)$, it follows from Jensen's inequality that

$$L\left(\int_{-\infty}^0 h(y)\,g(y)\,dy\right) \le \int_{-\infty}^0 L(h(y))\,g(y)\,dy$$

for any non-negative measurable function h. We therefore obtain

$$\begin{aligned} D(f * g|\psi) &= \int_0^\infty L\left(\int_{-\infty}^0 \frac{f(x-y)}{\psi(x)}\,g(y)\,dy\right)\psi(x)\,dx \\ &\le \int_0^\infty \int_{-\infty}^0 L\left(\frac{f(x-y)}{\psi(x)}\right) g(y)\,dy\,\psi(x)\,dx \\ &= \int_{-\infty}^0 \int_0^\infty f(x-y) \log\left(\frac{f(x-y)}{\psi(x)}\right) dx\,g(y)\,dy\,. \end{aligned}$$

Since $\psi(x)$ is decreasing in x, we have, for any $y < 0$,

$$\begin{aligned} \int_0^\infty f(x-y) \log\left(\frac{f(x-y)}{\psi(x)}\right) dx &\le \int_0^\infty f(x-y) \log\left(\frac{f(x-y)}{\psi(x-y)}\right) dx \\ = \int_0^\infty L\left(\frac{f(u)}{\psi(u)}\right)\psi(u)\,du &- \int_0^{-y} L\left(\frac{f(u)}{\psi(u)}\right)\psi(u)\,du \le D(f|\psi) + e^{-1}\,. \end{aligned}$$

Combining these estimates, we get

$$D(f * g \,|\, \psi) \le D(f \,|\, \psi) + e^{-1}\,,$$

and the lemma is proved. □

Lemma 2.4 (cf. [BCG3, Lemma 2.3])**.** *Suppose that f, g are probability densities on the positive half-line and α, β are positive real numbers with $\alpha + \beta = 1$. Then*

$$D(\alpha f + \beta g \,|\, \psi) \le \alpha D(f \,|\, \psi) + \beta D(g \,|\, \psi)$$

and

$$D(\alpha f + \beta g \,|\, \psi) \ge \alpha D(f \,|\, \psi) + \beta D(g \,|\, \psi) + L(\alpha) + L(\beta)\,.$$

In particular, it follows from Lemmas 2.1 and 2.4 that for any non-negative measurable functions f, g with $\int_0^\infty f(x)\,dx < \infty$, $\int_0^\infty g(x)\,dx < \infty$ and any $\alpha, \beta > 0$, we have

$$D(\alpha f + \beta g \,|\, \psi) < \infty \quad \text{if and only if} \quad D(f \,|\, \psi) < \infty \text{ and } D(g \,|\, \psi) < \infty\,. \tag{2.2}$$

Lemma 2.5. *Suppose that* (f_n) *and* (g_n) *are sequences of non-negative measurable functions on the positive half-line such that*

$$\int_0^\infty f_n(x)\,dx = 1 + o(1) \quad \textit{and} \quad \int_0^\infty g_n(x)\,dx = o(1)$$

as $n \to \infty$*. Then*

$$D(f_n + g_n \,|\, \psi) = D(f_n \,|\, \psi) + D(g_n \,|\, \psi) + o(1) \qquad \textit{as } n \to \infty\,.$$

Proof. Setting $\alpha_n := \int_0^\infty f_n(x)\,dx$ and $\beta_n := \int_0^\infty g_n(x)\,dx$ and using Lemmas 2.1 and 2.4, we have

$$\begin{aligned}
&D(f_n + g_n|\psi) \\
&= (\alpha_n{+}\beta_n)\, D(\tfrac{\alpha_n}{\alpha_n+\beta_n}\tfrac{f_n}{\alpha_n} + \tfrac{\beta_n}{\alpha_n+\beta_n}\tfrac{g_n}{\beta_n}|\psi) + L(\alpha_n + \beta_n) \\
&= (\alpha_n{+}\beta_n)\Big[\tfrac{\alpha_n}{\alpha_n+\beta_n} D(\tfrac{f_n}{\alpha_n}|\psi) + \tfrac{\beta_n}{\alpha_n+\beta_n} D(\tfrac{g_n}{\beta_n}|\psi) + \mathcal{O}(-L(\tfrac{\alpha_n}{\alpha_n+\beta_n})) + \mathcal{O}(-L(\tfrac{\beta_n}{\alpha_n+\beta_n}))\Big] \\
&\qquad + L(\alpha_n + \beta_n) \\
&= \alpha_n D(\tfrac{f_n}{\alpha_n}|\psi) + \beta_n D(\tfrac{g_n}{\beta_n}|\psi) + \mathcal{O}(-L(\tfrac{\alpha_n}{\alpha_n+\beta_n})) + \mathcal{O}(-L(\tfrac{\beta_n}{\alpha_n+\beta_n})) \\
&\qquad + L(\alpha_n + \beta_n) \\
&= D(f_n|\psi) + D(g_n|\psi) - L(\alpha_n) - L(\beta_n) + \mathcal{O}(-L(\tfrac{\alpha_n}{\alpha_n+\beta_n})) + \mathcal{O}(-L(\tfrac{\beta_n}{\alpha_n+\beta_n})) \\
&\qquad + L(\alpha_n + \beta_n) \\
&= D(f_n|\psi) + D(g_n|\psi) + o(1)
\end{aligned}$$

as $n \to \infty$. □

In the following sections, ψ will always be given by the probability density $\varphi_+(x) := \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}$ $(x > 0)$ or its rescaled version $\varphi_{n,+}(x) := \sqrt{\frac{2}{\pi n}}\, e^{-x^2/2n}$ $(x > 0)$, where $n \in \mathbb{N}$. Note that $\varphi_{n,+}$ is the density of the one-sided normal distribution with second moment n. It is easy to check that for any non-negative measurable function f, we have

$$D(\sqrt{n} f(\sqrt{n}\,\cdot\,) \,|\, \varphi_+) = D(f \,|\, \varphi_{n,+})\,. \tag{2.3}$$

3. Binomial Decomposition

In this section we start with the proof of sufficiency in Theorem 1.2. In the sequel, by a signed density we mean any measurable function $h(x)$ defined on the real line or on the positive half-line such that $\int_{-\infty}^\infty |h(x)|\,dx < \infty$. Since it is more convenient to work with bounded densities, we use a *binomial decomposition* of the density p to write the density $\overline{p}_n^*$ (restricted to the positive half-line) as the sum of two signed densities, a bounded term $\overline{q}_n^*$ and a remainder term $\overline{r}_n^*$. This representation will play an important role in the proof of the sufficiency part of Theorem 1.2.

Before we begin, let us remark that binomial decompositions are a well-known tool in the investigation of the classical central limit theorem, see e.g. [SM, IL]. In connection with entropic central limit theorems, they have recently been used in [BCG3].

Recall that p is the density of X_1. Write

$$p = (1-\varrho)q_1 + \varrho q_2\,, \tag{3.1}$$

where q_1 is a bounded probability density with $\int_0^\infty q_1(x)\,dx > 0$, q_2 is a potentially unbounded probability density, and $0 \le \varrho < \frac{1}{2}$. It follows that for any $n \ge 1$,

$$\begin{aligned} p_n(x) = p^{*n}(x) &= \left(\sum_{k=1}^{n} \tbinom{n}{k} (1-\varrho)^k \varrho^{n-k} \, (q_1^{*k} * q_2^{*(n-k)})(x) \right) + \varrho^n q_2^{*n}(x) \\ &=: (1-\varrho^n) q_{n,1}(x) + \varrho^n q_{n,2}(x)\,, \end{aligned} \tag{3.2}$$

where $q_{n,1}(x)$ and $q_{n,2}(x)$ are again probability densities.

We now need the following formula due to Nagaev [N3, Equation (0.8)]: For $n \in \mathbb{N}$ and $t \in \mathbb{R}$, we have

$$\mathbb{E}e^{it\overline{S}_n} = \sum_{k=1}^{n} f^k(t) \overline{\varphi}_{n-k}(t)\,, \tag{3.3}$$

where

$$\overline{\varphi}_0(t) := 1 \qquad \text{and} \qquad \overline{\varphi}_k(t) := \int_{-\infty}^{0} (1 - e^{itx})\, d\overline{F}_k(x) \qquad (k > 0)\,. \tag{3.4}$$

By (3.3) and the uniqueness theorem for Fourier transforms (of signed measures), it follows that the density of $\overline{S}_n := \max\{S_1, \ldots, S_n\}$ is given by

$$\overline{p}_n(x) = \sum_{k=1}^{n} (p^{*k} * G_{n-k})(x)\,,$$

where

$$G_0(dx) := \delta_0(dx), \quad G_k(dx) := \overline{F}_k(0)\delta_0(dx) - \overline{p}_k(x)\, \mathbf{1}_{(-\infty,0)}(x)\,dx \quad \text{for } k > 0$$

and $(p^{*k} * G_{n-k})(x) := \int p^{*k}(x-y)\, G_{n-k}(dy)$.

Using (3.2), we may write

$$\overline{p}_n(x) = \overline{q}_n(x) + \overline{r}_n(x)\,, \tag{3.5}$$

where

$$\overline{q}_n(x) := \sum_{k=1}^{n} (1-\varrho^k)(q_{k,1} * G_{n-k})(x) \quad \text{and} \quad \overline{r}_n(x) := \sum_{k=1}^{n} \varrho^k (q_{k,2} * G_{n-k})(x)\,. \tag{3.6}$$

Note that each $\overline{q}_n$ is bounded, since the $q_{k,1}$ are bounded and the G_{n-k} are finite signed measures. The main idea is to use $\overline{q}_n$ as a bounded approximation to $\overline{p}_n$.

Of course, we have to keep in mind that $\overline{q}_n$ and $\overline{r}_n$ are only signed densities in general. However, they may be represented as differences of non-negative densities by writing

$$\overline{q}_n(x) = \overline{q}_n^+(x) - \overline{q}_n^-(x) \quad \text{and} \quad \overline{r}_n(x) = \overline{r}_{n,1}(x) - \overline{r}_{n,2}(x)\,,$$

where $\overline{q}_n^+$ and $\overline{q}_n^-$ denote the positive and negative part of $\overline{q}_n$ and $\overline{r}_{n,1}$ and $\overline{r}_{n,2}$ are defined by

$$\overline{r}_{n,j}(x) := \sum_{k=1}^{n} \varrho^k (q_{k,2} * G_{n-k}^{\pm})(x)\,,$$

($j = 1, 2$), where $\pm = +$ for $j = 1$, $\pm = -$ for $j = 2$, and G_{n-k}^+ and G_{n-k}^- denote the positive and negative part of the signed measure G_{n-k}. Note that $\overline{r}_{n,1}$ and $\overline{r}_{n,2}$ are *not* the positive and negative part of $\overline{r}_n$ in general.

Thus, we obtain

$$\overline{p}_n = (\overline{q}_n^+ - \overline{q}_n^-) + (\overline{r}_{n,1} - \overline{r}_{n,2}) \tag{3.7}$$

or (equivalently)

$$\overline{p}_n + \overline{q}_n^- + \overline{r}_{n,2} = \overline{q}_n^+ + \overline{r}_{n,1}\,. \tag{3.8}$$

Let $\overline{p}_n^*(x) := \sqrt{n}\,\overline{p}_n(\sqrt{n}\,x)$, $\overline{q}_n^*(x) := \sqrt{n}\,\overline{q}_n(\sqrt{n}\,x)$, $\overline{r}_n^*(x) := \sqrt{n}\,\overline{r}_n(\sqrt{n}\,x)$, etc. denote the rescaled versions of the above densities.

We then have the following result:

Lemma 3.1.
(a) $\int_0^\infty |p_n^*(x) - q_n^*(x)|\,dx = \mathcal{O}(n^{-1/2})\,.$
(b) $\int_0^\infty x^2 |p_n^*(x) - q_n^*(x)|\,dx = \mathcal{O}(n^{-1/2})\,.$
(c) If (1.8) *holds then* $D(p_n^* \,|\, \varphi_+) = D((q_n^*)^+ \,|\, \varphi_+) + o(1)$ *as* $n \to \infty$.

Proof. Throughout this proof, for any measurable function p, we write

$$\|p\|_1 := \int_0^\infty |p(x)|\,dx$$

for the *total variation norm* (of the associated signed measure) and

$$\|p\|_\infty := \sup_{x\in(0,\infty)} |p(x)|$$

for the *supremum norm.* Furthermore, if p is non-negative, we write $D(p \,|\, \varphi_+)$ for the *relative entropy* as in (2.1). Recall the probability densities $\varphi_{n,+}$ introduced at the end of Section 2.

Analysis of $\overline{r}_{n,j}^*(x)$. By (1.9), $\overline{F}_n(0) = \mathcal{O}(n^{-1/2})$ as $n \to \infty$. Thus,

$$\|\overline{r}_{n,j}^*\|_1 \le \sum_{k=1}^{n} \overline{F}_{n-k}(0)\varrho^k \le \sum_{k=1}^{n} \frac{C_1 \varrho^k}{\sqrt{n-k+1}} = \mathcal{O}(n^{-1/2})\,, \tag{3.9}$$

$j = 1, 2$. Also, since $G_{n-k}^{\pm}$ is concentrated on $(-\infty, 0]$,

$$\int_0^\infty x^2 \overline{r}_{n,j}^*(x)\,dx \le \tfrac{1}{n} \sum_{k=1}^{n} \frac{C_1 \varrho^k}{\sqrt{n-k+1}} \int_{-\infty}^{+\infty} x^2 q_{k,2}(x)\,dx\,,$$

$j=1,2$. Let $Y_1,\dots,Y_k$ be i.i.d. random variables with density q_2. Then

$$\int_{-\infty}^{+\infty} x^2 q_{k,2}(x)\,dx = \|Y_1+\ldots+Y_k\|_2^2 \le k^2\,\|Y_1\|_2^2\,,$$

and we come to the conclusion that

$$\int_0^{\infty} x^2 \overline{r}^*_{n,j}(x)\,dx \le \tfrac{1}{n}\sum_{k=1}^{n} \frac{C_1\varrho^k}{\sqrt{n-k+1}} \int_{-\infty}^{+\infty} x^2 q_{k,2}(x)\,dx = \mathcal{O}(n^{-3/2})\,, \tag{3.10}$$

$j=1,2$. Clearly, (3.9) and (3.10) imply (a) and (b).

We will now show that if (1.8) holds then

$$D\left(\overline{r}^*_{n,j}\,|\,\varphi_+\right) = D\left(\sum_{k=1}^{n} \varrho^k\, q_{k,2} * G^{\pm}_{n-k}\,\middle|\,\varphi_{n,+}\right) = o(1)\,, \tag{3.11}$$

$j=1,2$. We provide the details for $\overline{r}^*_{n,2}$ only, the argument for $\overline{r}^*_{n,1}$ being similar but simpler.

Note that $G^-_0=0$. For $k=1,\dots,n-1$, write $G^-_{n-k}(dx) = \overline{F}_{n-k}(0)\, r_{n-k}(x)\,dx$, where $r_{n-k}(x) := \overline{p}_{n-k}(x)/\overline{F}_{n-k}(0)$ $(x<0)$ is a probability density on $(-\infty,0)$. Also, write $q_2 = \lambda_+ q_{2,+} + \lambda_- q_{2,-}$, where $\lambda_+,\lambda_- \ge 0$, $\lambda_+ + \lambda_- = 1$, and $q_{2,+}$ and $q_{2,-}$ are probability densities on $(0,+\infty)$ and $(-\infty,0)$, respectively. Then

$$q_{k,2} = \sum_{j=0}^{k} \tbinom{k}{j} \lambda_+^j\, \lambda_-^{k-j}\, q_{2,+}^{*j} * q_{2,-}^{*(k-j)}\,,$$

and it follows by a two-fold application of Lemma 2.2 that

$$\begin{aligned}
&D\left(\sum_{k=1}^{n} \varrho^k\, q_{k,2} * G^-_{n-k}\,\middle|\,\varphi_{n,+}\right)\\
&\quad\le \sum_{k=1}^{n-1} \varrho^k\, \overline{F}_{n-k}(0)\, D\left(q_{k,2} * r_{n-k}\,\middle|\,\varphi_{n,+}\right) + \mathcal{O}(\log n/\sqrt{n})\\
&\quad\le \sum_{k=1}^{n-1} \varrho^k\, \overline{F}_{n-k}(0) \sum_{j=1}^{k} \tbinom{k}{j} \lambda_+^j\, \lambda_-^{k-j}\, D\left(q_{2,+}^{*j} * q_{2,-}^{*(k-j)} * r_{n-k}\,\middle|\,\varphi_{n,+}\right) + \mathcal{O}(\log n/\sqrt{n})\,.
\end{aligned}$$

Using Lemma 2.3 with $f(x) := q_{2,+}^{*j}(x)$ and $g(x) := (q_{2,-}^{*(k-j)} * r_{n-k})(x)$, we get

$$D\left(q_{2,+}^{*j} * q_{2,-}^{*(k-j)} * r_{n-k}\,\middle|\,\varphi_{n,+}\right) \le D\left(q_{2,+}^{*j}\,\middle|\,\varphi_{n,+}\right) + e^{-1}\,.$$

Let μ and σ^2 denote the mean and variance of the probability density $q_{2,+}$, and let φ_{μ,σ^2} denote the density of the Gaussian distribution with mean μ and variance σ^2. As a consequence of the entropy power inequality (see e.g. Theorem 4 in [DCT]), we have

$$D(q_{2,+}^{*j}|\varphi_{j\mu,j\sigma^2}) \le D(q_{2,+}|\varphi_{\mu,\sigma^2})\,, \quad j\ge 1\,.$$

We therefore obtain

$$\begin{aligned} D(q_{2,+}^{*j}|\varphi_{n,+}) &= \int_0^\infty q_{2,+}^{*j} \log\left(\frac{q_{2,+}^{*j}}{\varphi_{j\mu,j\sigma^2}}\frac{\varphi_{j\mu,j\sigma^2}}{\varphi_{n,+}}\right) dx \\ &= D(q_{2,+}^{*j}|\varphi_{j\mu,j\sigma^2}) + \int_0^\infty q_{2,+}^{*j} \log\left(\frac{\varphi_{j\mu,j\sigma^2}}{\varphi_{n,+}}\right) dx \\ &\le D(q_{2,+}|\varphi_{\mu,\sigma^2}) + \mathcal{O}(\log n + j + 1) \\ &= \int_0^\infty q_{2,+} \log\left(\frac{q_{2,+}}{\varphi_+}\frac{\varphi_+}{\varphi_{\mu,\sigma^2}}\right) dx + \mathcal{O}(\log n + j + 1) \\ &= D(q_{2,+}|\varphi_+) + \int_0^\infty q_{2,+} \log\left(\frac{\varphi_+}{\varphi_{\mu,\sigma^2}}\right) dx + \mathcal{O}(\log n + j + 1) \\ &= \mathcal{O}(\log n + j + 1)\,, \end{aligned}$$

the implicit constants depending only on $q_{2,+}$. Here the last step follows from (1.8), see the remark below Lemma 2.4.

Combining the preceding estimates, it follows that

$$\begin{aligned} &D\left(\sum_{k=1}^n \varrho^k\, q_{k,2} * G_{n-k}^- \,\middle|\, \varphi_{n,+}\right) \\ &\qquad \le \sum_{k=1}^{n-1} \varrho^k\, \overline{F}_{n-k}(0)\, \mathcal{O}(\log n + k + 1) + \mathcal{O}(\log n/\sqrt{n}) = \mathcal{O}(\log n/\sqrt{n})\,, \end{aligned}$$

and the proof of (3.11) is complete.

Analysis of $(\overline{q}_n^*)^\pm(x)$. To complete the proof of part (c), we will show that the relative entropy of the main terms $\overline{p}_n^*$ and $(\overline{q}_n^*)^+$ in (3.8) is "stable" w.r.t. the addition of the error terms $\overline{r}_{n,1}^*$, $\overline{r}_{n,2}^*$ and $(\overline{q}_n^*)^-$.

To begin with, since $\|\overline{p}_n^*\|_1 = 1 - \overline{F}_n(0) = 1 + \mathcal{O}(1/\sqrt{n})$ and $\|\overline{r}_{n,j}^*\|_1 = \mathcal{O}(1/\sqrt{n})$ $(j = 1, 2)$, it follows from (3.8) that

$$\|(\overline{q}_n^*)^+\|_1 = 1 + \mathcal{O}(1/\sqrt{n}) \quad \text{and} \quad \|(\overline{q}_n^*)^-\|_1 = \mathcal{O}(1/\sqrt{n})\,. \tag{3.12}$$

Next we will show that

$$D((\overline{q}_n^*)^- \mid \varphi_+) = D(\overline{q}_n^- \mid \varphi_{n,+}) = o(1)\,. \tag{3.13}$$

Since q_1 is bounded by construction, $(1 - \varrho^k)q_{k,1}$ is bounded uniformly in $k \ge 1$, and we obtain

$$\|\overline{q}_n\|_\infty = \|\sum_{k=1}^n (1-\varrho^k) q_{k,1} * G_{n-k}\|_\infty = \mathcal{O}\left(\sum_{k=1}^n \frac{1}{\sqrt{n-k+1}}\right) = \mathcal{O}(\sqrt{n})\,.$$

Since $\|\overline{q}_n^-\|_1 = \mathcal{O}(1/\sqrt{n})$, it follows that

$$\begin{aligned} D(\overline{q}_n^-|\varphi_{n,+}) = \int_0^\infty \overline{q}_n^- \log(\overline{q}_n^-/\varphi_{n,+})\, dx &\le \int_0^\infty \overline{q}_n^- \log(C_1\sqrt{n}/\varphi_{n,+})\, dx \\ &= \mathcal{O}\Big(\frac{\log n}{\sqrt{n}}\Big) + \mathcal{O}\Big(\int_0^\infty \tfrac{1}{n} x^2\, \overline{q}_n^-(x)\, dx\Big)\,. \end{aligned}$$

Now, using (3.8) and (3.10), we have

$$\int_0^\infty \tfrac{1}{n}x^2\overline{q}_n^-(x)\,dx \le \int_0^\infty \tfrac{1}{n}x^2\overline{r}_{n,1}(x)\,dx = \int_0^\infty y^2\overline{r}^*_{n,1}(y)\,dy = \mathcal{O}(n^{-3/2})\,.$$

This completes the proof of (3.13).

Using (3.8) – (3.13) as well as Lemma 2.5, we now obtain

$$\begin{aligned} D(\overline{p}^*_n|\varphi_+) &= D(\overline{p}^*_n + (\overline{q}^*_n)^- + \overline{r}^*_{n,2}|\varphi_+) + o(1) \\ &= D((\overline{q}^*_n)^+ + \overline{r}^*_{n,1}|\varphi_+) + o(1) \\ &= D((\overline{q}^*_n)^+|\varphi_+) + o(1)\,, \end{aligned}$$

and Lemma 3.1 is proved. □

4. The Proof of Sufficiency in Theorem 1.2

This section contains the main part of the proof of sufficiency in Theorem 1.2. It relies on two auxiliary results whose proof is postponed to the following sections.

Proof of Sufficiency in Theorem 1.2. Suppose that (1.8) holds. Recall that $\overline{p}^*_n(x)$ is the density of $\overline{S}_n/\sqrt{n}$ (i.e. with the proper rescaling), and $\varphi_+(x) = \sqrt{2/\pi}\,e^{-x^2/2}$ $(x > 0)$. Using (1.9), it is easy to see that

$$D_+(\overline{S}_n/\sqrt{n}\,|\,|Z|) \to 0 \quad \text{if and only if} \quad D(\overline{p}^*_n\,|\,\varphi_+) \to 0\,. \tag{4.1}$$

Indeed, since $\overline{S}_n/\sqrt{n}$ conditioned to be positive has the density $\overline{p}^*_n(x)/(1-\overline{F}_n(0))$ $(x > 0)$, it follows from our definitions and Lemma 2.1 that

$$D(\overline{p}^*_n\,|\,\varphi_+) = (1-\overline{F}_n(0))\,D_+(\overline{S}_n/\sqrt{n}\,|\,|Z|) + L(1-\overline{F}_n(0))\,,$$

so that (4.1) follows from (1.9).

Since $D_+(\overline{S}_n/\sqrt{n}\,|\,|Z|) \ge 0$, it also follows from the preceding argument that

$$\liminf_{n\to\infty} D(\overline{p}^*_n\,|\,\varphi_+) \ge 0\,.$$

Thus, it remains to show that

$$\limsup_{n\to\infty} D(\overline{p}^*_n\,|\,\varphi_+) \le 0\,.$$

Recall that $\overline{q}^*_n(x) := \sqrt{n}\,\overline{q}_n(\sqrt{n}\,x)$, where $\overline{q}_n$ is defined in (3.6). By Lemma 3.1 (c), it is sufficient to show that for any $\varepsilon_0 > 0$,

$$D((\overline{q}^*_n)^+\,|\,\varphi_+) = \int_0^\infty L\left(\frac{(\overline{q}^*_n)^+(x)}{\varphi_+(x)}\right)\varphi_+(x)\,dx \le 3\varepsilon_0$$

for all sufficiently large $n \in \mathbb{N}$. Fix $\varepsilon_0 > 0$, and let C and c be positive real numbers with $0 < c < 1 < C < \infty$. (The precise choices will be specified below.) Then

$$\int_0^\infty L\left(\frac{(\overline{q}^*_n)^+(x)}{\varphi_+(x)}\right)\varphi_+(x)\,dx = E_1 + E_2 + E_3\,,$$

where E_1, E_2, E_3 denote the integrals over the intervals $(0, c)$, (c, C), (C, ∞), respectively. (Note that E_1, E_2, E_3 implicitly depend on n.) We will complete the proof of Theorem 1.2 by showing that if $C \in (1, \infty)$ is sufficiently large and $c \in (0, 1)$ is sufficiently small, then $E_j \le \varepsilon_0$ for all sufficiently large $n \in \mathbb{N}$.

To this end we will need the following results, which do not depend on condition (1.8):

Proposition 4.1. *For any $\varepsilon > 0$, there exists a constant $C > 0$ such that*

$$\int_C^\infty x^2 \,\overline{p}_n^*(x)\,dx \leq \varepsilon$$

for all sufficiently large $n \in \mathbb{N}$.

Proposition 4.2. *Under the assumptions of Theorem 1.2, there exist signed densities $r_n(x)$ such that $\|r_n\|_1 = \mathcal{O}(1/\sqrt{n})$, $\|r_n\|_\infty = \mathcal{O}(1)$ and the following holds:*

(a) For $x \in (0,\infty)$,

$$\overline{q}_n^*(x) = \varphi_+(x) + r_n(x) + o(1/x) \qquad \text{as } n \to \infty\,,$$

uniformly in $x \in (0,\infty)$.

(b) For $x \in (0,e^{-1})$,

$$\overline{q}_n^*(x) = \varphi_+(x) + r_n(x) + \mathcal{O}\left(\log n \wedge \tfrac{1}{\sqrt{n}x}\right) + \mathcal{O}\left(\log x^{-1}\right) \qquad \text{as } n \to \infty\,,$$

uniformly in $x \in (0,e^{-1})$.

Here the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are defined as in the proof of Lemma 3.1.

It should be mentioned that the proof of Proposition 4.2 closely follows that in Aleshkyavichene [Al].

In the special case where the random variables X_j have a bounded density $p(x)$, we could take

$$\overline{q}_n^*(x) := \overline{p}_n^*(x) \quad \text{and} \quad r_n(x) := \overline{F}_{n-1}(0)\,\sqrt{n}\,p(\sqrt{n}\,x) \quad (x > 0)\,,$$

and part (a) specializes to the following result from the literature:

Theorem 4.3 (Aleshkyavichene [Al])**.** *If the random variables X_j have a bounded density $p(x)$, we have $\overline{p}_n^*(x) = \varphi_+(x) + \overline{F}_{n-1}(0)\,\sqrt{n}p(\sqrt{n}x) + o(1/x)$, uniformly in $x \in (0,\infty)$.*

This may be regarded as a local version of the Erdős-Kac theorem 1.1. Moreover, part (b) is a refinement of part (a) which yields a better estimate for the error term for $x \approx 0$. Although this estimate is still unbounded, it is integrable near the origin. This is the crucial point for our purposes.

In the general case, the definition of the signed densities $r_n(x)$ is more complicated, see Equation (7.5) below.

Remark. In [Al] Theorem 4.3 is stated somewhat differently (for any $x_0 > 0$, the last term is of order $o(1)$ uniformly in $x > x_0$), but a careful analysis of the proof shows that after some minor modifications (similar to those in the proof of part (a) of Proposition 4.2 below), it also yields the result stated above.

Let us now explain how Propositions 4.1 and 4.2 may be used to complete the proof of sufficiency in Theorem 1.2.

Estimating E_3^+. By Proposition 4.2 (a), there exists a constant $M > 1$ (not depending on n) such that for $n \geq n_0$ and $x \geq 1$, $|\overline{q}_n^*(x)| \leq M$. It follows that

$$E_3 \leq \int_C^\infty |\overline{q}_n^*(x)|(\log M + \tfrac{1}{2}\log\tfrac{\pi}{2} + \tfrac{1}{2}x^2)\,dx \leq C_1 \int_C^\infty x^2|\overline{q}_n^*(x)|\,dx\,,$$

where C_1 is a constant depending only on M. By Proposition 4.1, there exists a constant $C > 1$ such that

$$\int_C^\infty x^2|\overline{p}_n^*(x)|\,dx < \varepsilon_0/C_1 \tag{4.2}$$

for all sufficiently large $n \in \mathbb{N}$. By Lemma 3.1 (b), this implies

$$\int_C^\infty x^2|\overline{q}_n^*(x)|\,dx < \varepsilon_0/C_1 \tag{4.3}$$

for all sufficiently large $n \in \mathbb{N}$. Thus, for C sufficiently large, we have $E_3 \leq \varepsilon_0$ for all sufficiently large $n \in \mathbb{N}$.

Estimating E_1^+. Suppose that $c \in (0, e^{-1})$. Setting

$$v_n(x) := \frac{(\overline{q}_n^*)^+(x) - \varphi_+(x)}{\varphi_+(x)} \qquad (x > 0)$$

and using that $L(y) \leq 0$ for $y \in [0,1]$ and $L(1+y) \leq y + \frac{1}{2}y^2$ for $y \in (1,\infty)$, we get

$$E_1 = \int_0^c L\left(1 + v_n(x)\right)\varphi_+(x)\,dx \leq \int_0^c \left(|v_n(x)| + \tfrac{1}{2}|v_n(x)|^2\right)\varphi_+(x)\,dx\,.$$

Using Proposition 4.2 (b), it follows that

$$\begin{aligned} E_1 &\leq \int_0^c |\overline{q}_n^*(x) - \varphi_+(x)| + \tfrac{1}{2}|\overline{q}_n^*(x) - \varphi_+(x)|^2/\varphi_+(x)\,dx \\ &\leq C_2\left(\int_0^c |r_n(x)|\,dx + \int_0^c (\log n \wedge \tfrac{1}{\sqrt{nx}})\,dx + \int_0^c (\log x^{-1})\,dx\right) \\ &\quad + C_3\left(\int_0^c |r_n(x)|^2\,dx + \int_0^c (\log n \wedge \tfrac{1}{\sqrt{nx}})^2\,dx + \int_0^c (\log x^{-1})^2\,dx\right). \end{aligned}$$

By Cauchy-Schwarz inequality, it remains to control the integrals in the last line. Now, for any fixed $c \in (0, e^{-1})$, we have

$$\int_0^c |r_n(x)|^2\,dx \leq \|r_n\|_1\|r_n\|_\infty = o(1)\,,$$

$$\int_0^c (\log n \wedge \tfrac{1}{\sqrt{nx}})^2\,dx = \frac{\log n}{\sqrt{n}} + \frac{1}{n}(-c^{-1} + \sqrt{n}\log n) = o(1)\,,$$

$$\int_0^c (\log x^{-1})^2\,dx = \int_{\log(1/c)}^\infty y^2 e^{-y}\,dy < \infty\,,$$

Thus, for c sufficiently small, we have $E_1 \leq \varepsilon_0$ for all sufficiently large $n \in \mathbb{N}$.

Estimating E_2^+. Let $C \in (1,\infty)$ and $c \in (0,1)$ be the constants fixed above. The same argument as for E_1^+ yields

$$E_2 = \int_c^C L\,(1+v_n(x))\,\varphi_+(x)\,dx \le \int_c^C \left(|v_n(x)| + \tfrac{1}{2}|v_n(x)|^2\right)\varphi_+(x)\,dx\,.$$

Using Proposition 4.2 (a), it follows that

$$\begin{aligned} E_2 &\le \int_c^C |\overline{q}_n^*(x) - \varphi_+(x)| + \tfrac{1}{2}|\overline{q}_n^*(x) - \varphi_+(x)|^2/\varphi_+(x)\,dx \\ &\le C_4\bigg(\int_c^C |r_n(x)|\,dx + o(1)\int_c^C x^{-1}\,dx\bigg) \\ &\qquad + C_5\exp(C^2/2)\bigg(\int_c^C |r_n(x)|^2\,dx + o(1)\int_c^C x^{-2}\,dx\bigg) \\ &\le C_4\Big(\|r_n\|_1 + o(1)(\log C - \log c)\Big) \\ &\qquad + C_5\exp(C^2/2)\Big(\|r_n\|_1 \cdot \|r_n\|_\infty + o(1)(c^{-1} - C^{-1})\Big)\,. \end{aligned}$$

Thus, $E_2^+ = o(1)$ as $n \to \infty$.

This completes the proof of sufficiency in Theorem 1.2 (up to the proof of Propositions 4.1 and 4.2). □

5. Some Auxiliary Results

Let us collect some results from the literature which will be needed for the proofs of Propositions 4.1 and 4.2.

Let $\overline{a}_k := \int_{-\infty}^0 x\,d\overline{F}_k(x)$ and $\overline{b}_k := \int_{-\infty}^0 x^2\,d\overline{F}_k(x)$, $k \ge 1$. It is known that under our standing moment assumptions the functions $\overline{\varphi}_k(t)$ introduced in (3.4) satisfy the following estimates:

$$|\overline{\varphi}_k(t)| \le 2\overline{F}_k(0)\,, \tag{5.1}$$

$$|\overline{\varphi}_k(t)| \le |\overline{a}_k||t|\,, \tag{5.2}$$

$$|\overline{\varphi}_k'(t)| \le |\overline{a}_k|\,, \tag{5.3}$$

$$|\overline{\varphi}_k(t) - (-it\overline{a}_k)| \le \tfrac{1}{2}|\overline{b}_k||t|^2\,, \tag{5.4}$$

$$|\overline{\varphi}_k'(t) - (-i\overline{a}_k)| \le |\overline{b}_k||t|\,, \tag{5.5}$$

$$|\overline{\varphi}_k''(t)| \le |\overline{b}_k|\,, \tag{5.6}$$

(see e.g. [Al, Equations (26) and (46)]), where

$$\overline{F}_k(0) = \mathcal{O}(k^{-1/2}) \tag{5.7}$$

(see e.g. [Al, Equation (39)]),

$$\overline{a}_k = -(2\pi k)^{-1/2} + o(k^{-1/2}) \quad \text{and} \quad \overline{b}_k = o(1) \tag{5.8}$$

(see e.g. [Al, Equation (1)]). Let us note that the implicit constants may depend on the distribution of X_1.

Furthermore, we need the following classical approximations for characteristic functions of sums of i.i.d. random variables and their derivatives:

Given i.i.d. random variables $X_1, X_2, X_3, \ldots$ with mean 0, variance 1, density p and characteristic function f, there exist positive real numbers $\gamma, \delta_1, \delta_2, \delta_3, \ldots$ (depending on the distribution of X_1) with $\lim_{n\to\infty} \delta_n = 0$ such that for $n \in \mathbb{N}$, $|t| \le \gamma n^{1/2}$ and $j = 0, 1, 2$,

$$\left| \tfrac{d^j}{dt^j} (f^n(t/\sqrt{n}) - e^{-t^2/2}) \right| \le \delta_n \, e^{-t^2/4} .$$

See e.g. [BR, Theorem 9.12]. Replacing n with k and t with $t\sqrt{k/n}$ in this estimate, we obtain, for $1 \le k \le n$, $|t| \le \gamma n^{1/2}$ and $j = 0, 1, 2$,

$$\left| \tfrac{d^j}{dt^j} (f^k(t/\sqrt{n}) - e^{-kt^2/2n}) \right| \le \delta_k \, (k/n)^{j/2} e^{-kt^2/4n} . \tag{5.9}$$

Furthermore, let $\eta \in (0, 1)$ be a constant such that

$$|t| \ge \gamma \quad \Rightarrow \quad |f(t)| \le \eta . \tag{5.10}$$

The existence of such a constant follows from the Riemann-Lebesgue lemma, X_1 having a density.

Besides that, we will repeatedly use the fact that for any $\alpha > 0$ and $n \ge k \ge 1$,

$$\sup_{t\in\mathbb{R}} (kt^2/n)^{\alpha/2} \, e^{-kt^2/4n} = \mathcal{O}_\alpha(1) , \tag{5.11}$$

and

$$\int_{-\infty}^{+\infty} (kt^2/n)^{\alpha/2} \, e^{-kt^2/4n} \, dt = \mathcal{O}_\alpha(\sqrt{\tfrac{n}{k}}) \tag{5.12}$$

with implicit constants depending only on α.

In addition to that, we will use the following (well-known) Gaussian tail bounds: For any $\alpha > 0$ and $t > 0$ we have

$$\int_t^\infty e^{-\alpha x^2/2} \, dx \le \sqrt{\tfrac{\pi}{2\alpha}} \wedge \tfrac{1}{\alpha t} e^{-\alpha t^2/2} , \tag{5.13}$$

$$\int_t^\infty \sqrt{\alpha} x e^{-\alpha x^2/2} \, dx \le \tfrac{1}{\sqrt{\alpha}} e^{-\alpha t^2/2} , \tag{5.14}$$

$$\int_t^\infty \alpha x^2 e^{-\alpha x^2/2} \, dx \le \sqrt{\tfrac{\pi}{2\alpha}} \wedge \tfrac{1}{\alpha t}(\alpha t^2 + 1) e^{-\alpha t^2/2} . \tag{5.15}$$

Moreover, we will repeatedly use the fact that

$$\sum_{k=1}^{n-1} \frac{1}{\sqrt{k(n-k)}} = \mathcal{O}\left(\frac{1}{\sqrt{n}} \sum_{1\le k\le n/2} \frac{1}{\sqrt{k}} \right) + \mathcal{O}\left(\frac{1}{\sqrt{n}} \sum_{n/2\le k\le n-1} \frac{1}{\sqrt{n-k}} \right) = \mathcal{O}(1) . \tag{5.16}$$

A similar decomposition shows that if $(t_n)_{n\in\mathbb{N}}$ is a sequence of real numbers with $\lim_{n\to\infty} t_n = 0$, we have

$$\sum_{k=1}^{n-1} \frac{t_k}{\sqrt{k(n-k)}} = o(1) \qquad \text{and} \qquad \sum_{k=1}^{n-1} \frac{t_{n-k}}{\sqrt{k(n-k)}} = o(1) . \tag{5.17}$$

Finally, we will need the observation that the Fourier transform of the density $\varphi_+(x) := \sqrt{2/\pi}e^{-x^2/2}$ $(x > 0)$ is given by

$$\hat{\varphi}_+(t) = e^{-t^2/2} + \frac{it}{\sqrt{2\pi n}} \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \tag{5.18}$$

(see [Al, page 452]). It follows from this that for any $x > 0$,

$$\varphi_+(x) = \frac{1}{2\pi} \lim_{R\to\infty} \int_{-R}^{+R} e^{-itx} \left[e^{-t^2/2} + \frac{it}{\sqrt{2\pi n}} \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \right] dt \tag{5.19}$$

(see [Al, page 452]).

6. The proof of Proposition 4.1

Proposition 4.1 will be deduced from the following result concerning characteristic functions:

Proposition 6.1. *For $k = 0, 1, 2$, we have*

$$\frac{d^k}{dt^k} \left[\mathbb{E}(e^{it\overline{S}_n/\sqrt{n}}) - \hat{\varphi}_+(t) \right] = o(1)$$

as $n \to \infty$, uniformly in $|t| \le \gamma n^{1/2}$.

Remarks 6.2.

(a) The Erdős-Kac theorem is equivalent to the statement that $\mathbb{E}(e^{it\overline{S}_n/\sqrt{n}}) \to \hat{\varphi}_+(t)$ for any fixed $t \in \mathbb{R}$. Thus, the Erdős-Kac theorem follows from Proposition 6.1. Let us emphasize that we do not need the existence of densities in this section.

(b) For our "application" (namely the proof of Proposition 4.1), the result for the second derivative is relevant. Indeed, for this application, it would be be sufficient to prove Proposition 6.1 for $t = \mathcal{O}(1)$.

Proof of Proposition 6.1. Similarly as in [Al, Na], using (3.3) and (5.18), we have the following decomposition:

$$\begin{aligned}
D(t) &:= \mathbb{E}(e^{it\overline{S}_n/\sqrt{n}}) - \hat{\varphi}_+(t) \\
&= \left(f^n(t/\sqrt{n}) - e^{-t^2/2} \right) \\
&+ \left(\frac{it}{\sqrt{2\pi n}} \left(\sum_{k=3}^{n-1} e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} - \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \right) \right) \\
&+ \left(\sum_{k=3}^{n-1} \left(f^k(t/\sqrt{n}) - e^{-kt^2/2n} \right) \overline{\varphi}_{n-k}(t/\sqrt{n}) \right) \\
&+ \left(\sum_{k=3}^{n-1} e^{-kt^2/2n} \left(\overline{\varphi}_{n-k}(t/\sqrt{n}) - (-\overline{a}_{n-k})\, it/\sqrt{n} \right) \right) \\
&+ \left(\sum_{k=3}^{n-1} e^{-kt^2/2n} \left((-\overline{a}_{n-k}) - \frac{1}{\sqrt{2\pi(n-k)}} \right) it/\sqrt{n} \right) \\
&+ \left(f^2(t/\sqrt{n})\overline{\varphi}_{n-2}(t/\sqrt{n}) + f(t/\sqrt{n})\overline{\varphi}_{n-1}(t/\sqrt{n}) \right).
\end{aligned}$$

Denote the differences in the large round brackets by $D_1(t), \dots, D_6(t)$. (Note that all these differences implicitly depend on n.) We will show that for $j = 1, \dots, 6$, uniformly in $|t| \le \gamma n^{1/2}$, $D_j(t), D_j'(t), D_j''(t) \to 0$ as $n \to \infty$.

Convention: We always assume that $n \ge 4$ and $|t| \le \gamma n^{1/2}$. $\mathcal{O}$- and o-bounds hold uniformly in this region (unless otherwise mentioned), and they may depend on the constants $\gamma, \delta_1, \delta_2, \delta_3, \dots$ and η introduced in Section 5.

On the Difference D_1. For the difference $D_1(t)$ and its first two derivatives, the claim is immediate from (5.9) (with $k = n$).

On the Difference D_2. For fixed $n \in \mathbb{N}$, $t \in \mathbb{R}$ and $\beta \in \{0, 1, 2, \dots\}$, put

$$h_\beta(u) := (u/n)^\beta \, e^{-ut^2/2n} \, \tfrac{1}{\sqrt{n-u}} \qquad (0 < u < n)\,.$$

Then, for $1 \le v \le w \le n-1$, we have

$$\begin{aligned} |h_\beta(w) - h_\beta(v)| &= \left| \int_v^w h_\beta'(u)\, du \right| = \left| \int_v^w \left(\tfrac{\beta}{u} - \tfrac{t^2}{2n} + \tfrac{1}{2(n-u)} \right) h_\beta(u)\, du \right| \\ &\le (w-v) \left(\tfrac{\beta}{v} + \tfrac{t^2}{2n} + \tfrac{1}{2(n-w)} \right) (w/n)^\beta \, e^{-vt^2/2n} \, \tfrac{1}{\sqrt{n-w}}\,. \end{aligned}$$

Hence, for the difference $D_2(t)$, we get (using the above estimate with $\beta = 0$)

$$\begin{aligned} &\left| \frac{it}{\sqrt{2\pi n}} \left(\sum_{k=3}^{n-1} e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} - \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \right) \right| \\ &\le \frac{|t|}{\sqrt{2\pi n}} \sum_{k=3}^{n-2} \left| \int_k^{k+1} e^{-kt^2/2n} \tfrac{1}{\sqrt{n-k}} - e^{-ut^2/2n} \tfrac{1}{\sqrt{n-u}}\, du \right| + \mathcal{O}(n^{-1/2}) \\ &\le \frac{|t|}{\sqrt{2\pi n}} \sum_{k=3}^{n-2} \left(\tfrac{1}{(n-k-1)^{3/2}} e^{-kt^2/2n} + \tfrac{1}{(n-k-1)^{1/2}} \tfrac{t^2}{2n} e^{-kt^2/2n} \right) + \mathcal{O}(n^{-1/2}) \\ &= \mathcal{O}\left(\sum_{k=3}^{n-2} \left(\tfrac{1}{k^{1/2}(n-k-1)^{3/2}} + \tfrac{1}{k^{3/2}(n-k-1)^{1/2}} \right) \right) + \mathcal{O}(n^{-1/2}) = \mathcal{O}(n^{-1/2})\,. \end{aligned}$$

Here we have used the fact that $(k/n)^{1/2}\, |t|\, e^{-kt^2/2n}$ and $(k/n)^{3/2}\, |t|^3\, e^{-kt^2/2n}$ are uniformly bounded. In particular, this fact is also used in the first step to absorb the summand for $k = n-1$ and the integral over $u \in [n-1, n]$ into the $\mathcal{O}(n^{-1/2})$-term.

Furthermore, similar estimates hold for the first two derivatives of $D_2(t)$. Indeed, these derivatives are finite linear combinations of expressions of the form

$$\frac{it^\alpha}{\sqrt{2\pi n}} \left(\sum_{k=3}^{n-1} (k/n)^\beta e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} - \int_0^n (u/n)^\beta e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \right)$$

(with $\alpha, \beta \in \{0, 1, 2, 3, \ldots\}$ and $\alpha \le \beta + 1$), and, by similar arguments as above,

$$\begin{aligned}
&\left| \frac{it^\alpha}{\sqrt{2\pi n}} \Big(\sum_{k=3}^{n-1} (k/n)^\beta e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} - \int_0^n (u/n)^\beta e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \Big) \right| \\
&\le \frac{|t|^\alpha}{\sqrt{2\pi n}} \sum_{k=3}^{n-2} \left| \int_k^{k+1} (k/n)^\beta e^{-kt^2/2n} \tfrac{1}{\sqrt{n-k}} - (u/n)^\beta e^{-ut^2/2n} \tfrac{1}{\sqrt{n-u}} \, du \right| + \mathcal{O}_\beta(n^{-1/2}) \\
&\le \frac{|t|^\alpha}{\sqrt{2\pi n}} \sum_{k=3}^{n-2} \Big(\tfrac{((k+1)/n)^\beta}{(n-k-1)^{3/2}} e^{-kt^2/2n} + \tfrac{((k+1)/n)^\beta}{(n-k-1)^{1/2}} \tfrac{t^2}{2n} e^{-kt^2/2n} + \tfrac{((k+1)/n)^\beta}{(n-k-1)^{1/2}} \tfrac{\beta}{k} e^{-kt^2/2n} \Big) \\
&\qquad + \mathcal{O}_\beta(n^{-1/2}) \\
&= \mathcal{O}_\beta\Big(\sum_{k=3}^{n-2} \Big(\tfrac{1}{k^{1/2}(n-k-1)^{3/2}} + \tfrac{1}{k^{3/2}(n-k-1)^{1/2}} + \tfrac{1}{k^{3/2}(n-k-1)^{1/2}} \Big) \Big) + \mathcal{O}_\beta(n^{-1/2}) \\
&= \mathcal{O}_\beta(n^{-1/2}) \,.
\end{aligned}$$

On the Difference D_3. For the difference $D_3(t)$, the claim follows from (5.9) (with $k < n$), (5.2), (5.8) and (5.17), since

$$\begin{aligned}
&\sum_{k=3}^{n-1} \Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n} \Big) \overline{\varphi}_{n-k}(t/\sqrt{n}) \\
&\qquad = \mathcal{O}\left(\sum_{k=3}^{n-1} \frac{\delta_k \, (k/n)^{1/2} |t| \, e^{-kt^2/4n}}{\sqrt{k(n-k)}} \right) = \mathcal{O}\left(\sum_{k=3}^{n-1} \frac{\delta_k}{\sqrt{k(n-k)}} \right) = o(1) \,.
\end{aligned}$$

Similar estimates hold for the first two derivatives. Indeed, using (5.9) (with $k < n$), (5.2) – (5.3), (5.6), (5.8) and (5.17), we get

$$\begin{aligned}
&\sum_{k=3}^{n-1} \frac{d}{dt} \Big[\Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n} \Big) \overline{\varphi}_{n-k}(t/\sqrt{n}) \Big] \\
&= \sum_{k=3}^{n-1} \Big[\frac{d}{dt} \Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n} \Big) \overline{\varphi}_{n-k}(t/\sqrt{n}) \\
&\qquad\qquad + \Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n} \Big) \overline{\varphi}'_{n-k}(t/\sqrt{n})/\sqrt{n} \Big] \\
&= \mathcal{O}\left(\sum_{k=3}^{n-1} \left[\frac{\delta_k \, (k/n)^{1/2} |t| \, e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{\delta_k \, e^{-kt^2/4n}}{\sqrt{n(n-k)}} \right] \right) \\
&= \mathcal{O}\left(\sum_{k=3}^{n-1} \left[\frac{\delta_k}{\sqrt{n(n-k)}} + \frac{\delta_k}{\sqrt{n(n-k)}} \right] \right) = o(1)
\end{aligned}$$

as well as

$$\begin{aligned}
&\sum_{k=3}^{n-1} \frac{d^2}{dt^2}\Big[\Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n})\Big] \\
&= \sum_{k=3}^{n-1}\Big[\frac{d^2}{dt^2}\Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n}) \\
&\qquad + 2\frac{d}{dt}\Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}'_{n-k}(t/\sqrt{n})/\sqrt{n} \\
&\qquad\qquad + \Big(f^k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}''_{n-k}(t/\sqrt{n})/n\Big] \\
&= \mathcal{O}\left(\sum_{k=3}^{n-1}\left[\frac{\delta_k\,(k/n)\,|t|e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{\delta_k\,(k/n)^{1/2}\,e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{\delta_k\,|b_{n-k}|\,e^{-kt^2/4n}}{n}\right]\right) \\
&= \mathcal{O}\left(\sum_{k=3}^{n-1}\left[\frac{\delta_k\sqrt{k/n}}{\sqrt{n(n-k)}} + \frac{\delta_k\sqrt{k/n}}{\sqrt{n(n-k)}} + \frac{\delta_k|b_{n-k}|}{n}\right]\right) = o(1)\,.
\end{aligned}$$

On the Difference D_4. Let $(m_n)_{n\in\mathbb{N}}$ be a sequence of natural numbers such that $\lim_{n\to\infty} m_n = \infty$ and $\lim_{n\to\infty}(m_n/n) \to 0$. Then, by (5.4), (5.2) and (5.8), we have

$$\begin{aligned}
&\sum_{k=3}^{n-1} e^{-kt^2/2n}\Big|\overline{\varphi}_{n-k}(t/\sqrt{n}) - (-\overline{a}_{n-k})\,it/\sqrt{n}\Big| \\
&\qquad \le \sum_{k=3}^{n-m_n}(t^2/n)e^{-kt^2/2n}|b_{n-k}| + \sum_{k=n-m_n}^{n-1}(|t|/\sqrt{n})e^{-kt^2/2n}|a_{n-k}| \\
&\qquad\qquad = o\left(\sum_{k=3}^{n-m_n}(t^2/n)e^{-kt^2/2n}\right) + \mathcal{O}\left(\sum_{k=n-m_n}^{n-1}\frac{1}{\sqrt{k(n-k)}}\right)\,.
\end{aligned}$$

Since $\sum_{k=3}^{\infty} xe^{-kx}$ is uniformly bounded in $x > 0$, it follows that $D_4(t) = o(1)$.

Similar estimates hold for the first two derivatives. Indeed, to this end, we have to bound, among other terms,

$$\sum_{k=3}^{n-1} e^{-kt^2/2n}\Big(\overline{\varphi}'_{n-k}(t/\sqrt{n})/\sqrt{n} - (-\overline{a}_{n-k})\,i/\sqrt{n}\Big)$$

and

$$\sum_{k=3}^{n-1} e^{-kt^2/2n}\Big(\overline{\varphi}''_{n-k}(t/\sqrt{n})/n\Big)\,.$$

(For the other terms we get similar bounds as for lower-order derivatives but with extra factors kt/n, which are easily controlled due to the exponential factor $e^{-kt^2/2n}$.)

But, using (5.5), (5.3), (5.6), and (5.8), we get

$$\sum_{k=3}^{n-1} e^{-kt^2/2n} \Big| \overline{\varphi}'_{n-k}(t/\sqrt{n})/\sqrt{n} - (-\overline{a}_{n-k})\, i/\sqrt{n} \Big|$$
$$\leq \sum_{k=3}^{n-m_n} (|t|/n) e^{-kt^2/2n} |b_{n-k}| + \sum_{k=n-m_n}^{n-1} e^{-kt^2/2n} |a_{n-k}|/\sqrt{n}$$
$$= o\Bigg(\sum_{k=3}^{n-m_n} \frac{t^2+1}{n} e^{-kt^2/2n} \Bigg) + \mathcal{O}\Bigg(\sum_{k=n-m_n}^{n-1} \frac{1}{\sqrt{n(n-k)}} \Bigg) = o(1)$$

as well as

$$\sum_{k=3}^{n-1} e^{-kt^2/2n} \Big| \overline{\varphi}''_{n-k}(t/\sqrt{n})/n \Big| \leq \sum_{k=3}^{n-1} (1/n) e^{-kt^2/2n} |b_{n-k}| \leq \frac{1}{n} \sum_{k=3}^{n-1} |b_{n-k}| = o(1)\,.$$

On the Difference D_5. Similarly as above, let $(m_n)_{n\in\mathbb{N}}$ be a sequence of natural numbers such that $\lim_{n\to\infty} m_n = \infty$ and $\lim_{n\to\infty}(m_n/n) \to 0$. Then, using (5.8), we have

$$\sum_{k=3}^{n-1} e^{-kt^2/2n} \Big| (-\overline{a}_{n-k}) - \frac{1}{\sqrt{2\pi(n-k)}} \Big| \, \Big| it/\sqrt{n} \Big|$$
$$= o\left(\sum_{k=3}^{n-m_n} \frac{1}{\sqrt{k(n-k)}} \right) + \mathcal{O}\left(\sum_{k=n-m_n}^{n-1} \frac{1}{\sqrt{k(n-k)}} \right) = o(1)\,.$$

Again, for the derivatives, we have similar estimates involving lower powers of t and / or additional factors kt/n.

On the Difference D_6. For fixed k, we have

$$|(f^k)(t)| = \mathcal{O}_k(1)\,, \qquad |\tfrac{d}{dt}(f^k)(t)| = \mathcal{O}_k(1)\,, \qquad |\tfrac{d^2}{dt^2}(f^k)(t)| = \mathcal{O}_k(1)$$

(as follows from our assumption $\mathbb{E}X_1^2 < \infty$), and as $n \to \infty$,

$$\overline{\varphi}_n(t) = o(1)\,, \qquad \overline{\varphi}'_n(t) = o(1)\,, \qquad \overline{\varphi}''_n(t) = o(1)$$

(as follows from (5.1), (5.3) and (5.6) – (5.8)). The claim for the difference $D_6(t)$ and its first two derivatives follows immediately from these relations.

The proof of Proposition 6.1 is complete now. □

Proof of Proposition 4.1. To deduce Proposition 4.1 from Proposition 6.1, we use the following well-known observation: If X is a real random variable with $\mathbb{E}(X^{2k}) < \infty$, induced distribution $\mathbb{P}_X$ and characteristic function f_X, then, for any $T > 0$,

$$\int_{[-T,+T]^C} x^{2k}\, \mathbb{P}_X(dx) \leq \frac{T}{2} \int_{-2/T}^{+2/T} (-1)^k (f_X^{(2k)}(0) - f_X^{(2k)}(t))\, dt\,.$$

For the convenience of the reader, let us recall the argument: Using that $|\sin a/a| \le 1$ for $a \in \mathbb{R}$ and $|\sin a/a| \le \frac{1}{2}$ for $|a| \ge 2$, we get

$$\begin{aligned}
&\frac{T}{2}\int_{-2/T}^{+2/T}(-1)^k(f_X^{(2k)}(0) - f_X^{(2k)}(t))\,dt \\
&\qquad = \frac{T}{2}\int_{-2/T}^{+2/T}\int_{-\infty}^{+\infty} x^{2k}(1-e^{itx})\,d\mathbb{P}_X(x)\,dt \\
&\qquad = \int_{-\infty}^{+\infty} x^{2k}\frac{T}{2}\int_{-2/T}^{+2/T}(1-e^{itx})\,dt\,d\mathbb{P}_X(x) \\
&\qquad = \int_{-\infty}^{+\infty} x^{2k}\left(2 - 2\sin(\tfrac{2x}{T})/(\tfrac{2x}{T})\right)\,d\mathbb{P}_X(x) \\
&\qquad \ge \int_{[-T,+T]^C} x^{2k}\,d\mathbb{P}_X(x)\,.
\end{aligned}$$

Applying this inequality with $X = \overline{S}_n/\sqrt{n}$ and $T = C$, we get

$$\begin{aligned}
\int_C^\infty x^2\,\overline{p}_n^*(x)\,dx &\le \frac{C}{2}\int_{-2/C}^{+2/C}(\mathbb{E}(\overline{S}_n/\sqrt{n})^2 + \tfrac{d^2}{dt^2}\mathbb{E}(e^{it\overline{S}_n/\sqrt{n}}))\,dt \\
&\le 2\sup_{|t|\le 2/C}\left|f''_{\overline{S}_n/\sqrt{n}}(0) - f''_{\overline{S}_n/\sqrt{n}}(t)\right|\,.
\end{aligned}$$

Using Proposition 6.1, it follows that for any fixed $C > 0$, we have

$$\int_C^\infty x^2\,\overline{p}_n^*(x)\,dx \le 2\sup_{|t|\le 2/C}\left|\hat{\varphi}''_+(0) - \hat{\varphi}''_+(t)\right| + o(1)\,.$$

as $n \to \infty$. Since $\hat{\varphi}''_+(t)$ is continuous at zero, we may conclude that for $C = C(\varepsilon)$ sufficiently large, we have

$$\int_C^\infty x^2\overline{p}_n^*(x)\,dx \le \varepsilon$$

for all sufficiently large $n \in \mathbb{N}$, and the proof of Proposition 4.1 is complete. □

7. The proof of Proposition 4.2

Proof of Proposition 4.2. Let $p = (1-\varrho)q_1 + \varrho q_2$ be as in (3.1), and let g_1 and g_2 be the Fourier transforms of q_1 and q_2, respectively. For $k \ge 3$, put

$$\tilde{p}_k(x) := \sum_{j=3}^k \binom{k}{j}(1-\varrho)^j\,\varrho^{k-j}(q_1^{*j} * q_2^{*(k-j)})(x)\,.$$

and

$$\tilde{f}_k(t) := \sum_{j=3}^k \binom{k}{j}(1-\varrho)^j\,\varrho^{k-j}g_1^j(t)g_2^{k-j}(t)\,.$$

Note that $\tilde{f}_n(t)$ is the Fourier transform of $\tilde{p}_n(x)$ and that $\tilde{p}_n(x)$ can be recovered from $\tilde{f}_n(t)$ by means of Fourier inversion. This follows from the fact that $g_1 \in L^2$ (being the Fourier transform of a bounded probability density) and $g_2 \in L^\infty$ (being the Fourier transform of a probability measure).

Using our moment assumptions, it is easy to see for $k \geq 3$ and $t \in \mathbb{R}$,

$$\Big|\tfrac{d^j}{dt^j}\big(f^k(t/\sqrt{n}) - \tilde{f}_k(t/\sqrt{n})\big)\Big| = \mathcal{O}(n^{-j/2}\, 2^{-k}), \qquad j = 0,1,2.$$

It therefore follows from (5.9) that for $3 \leq k \leq n$, $|t| \leq \gamma n^{1/2}$ and $j = 0,1,2$,

$$\Big|\tfrac{d^j}{dt^j}(\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n})\Big| \leq \delta_k\,(k/n)^{j/2}\, e^{-kt^2/4n} + \mathcal{O}(n^{-j/2}\, 2^{-k}). \tag{7.1}$$

Furthermore, there exist a constant $C_0 > 0$ and a constant $\eta \in (0,1)$ such that for $k \geq 3$ and $|t| \geq \gamma$,

$$|\tilde{f}_k(t)| \leq \sum_{j=3}^{k} \tbinom{k}{j}(1-\varrho)^j \varrho^{k-j} \Big| \quad \Big[g_1^j(t) g_2^{k-j}(t)\Big]\Big| \leq C_0 \ \ \eta^{k-2}|g_1(t)|^2. \tag{7.2}$$

$$|\tilde{f}_k'(t)| \leq \sum_{j=3}^{k} \tbinom{k}{j}(1-\varrho)^j \varrho^{k-j} \Big|\tfrac{d}{dt}\Big[g_1^j(t) g_2^{k-j}(t)\Big]\Big| \leq C_0 k \eta^{k-3}|g_1(t)|^2. \tag{7.3}$$

This follows from the fact that g_1 and g_2 also satisfy (5.10) (possibly with some modified constant η) and that g_1' and g_2' are bounded, q_1 and q_2 being probability measures with finite moments.

Using the non-negative densities $\tilde{p}_k$ introduced above, we may write

$$\overline{q}_n^*(x) = \sqrt{n} \sum_{k=3}^{n} (\tilde{p}_k * G_{n-k})(\sqrt{n}\, x) + r_n(x), \tag{7.4}$$

where the remainder term $r_n(x)$ is given by

$$\begin{aligned} r_n(x) := \sqrt{n} \sum_{k=1}^{n} \tbinom{k}{1}(1-\varrho)\ \varrho^{k-1}(q_1 \ * q_2^{*(k-1)} * G_{n-k})(\sqrt{n}\, x) \\ + \sqrt{n} \sum_{k=2}^{n} \tbinom{k}{2}(1-\varrho)^2 \varrho^{k-2}(q_1^{*2} * q_2^{*(k-2)} * G_{n-k})(\sqrt{n}\, x). \end{aligned} \tag{7.5}$$

The functions $r_n(x)$ are the signed densities occurring in Proposition 4.2. It is easy to see that they have the asserted properties. Indeed, because q_1 and q_2 are probability densities, q_1 is bounded and the total variation norm of G_n is of order $\mathcal{O}(1/\sqrt{n})$, we have

$$\|q_1^{*j} * q_2^{*(k-j)} * G_{n-k}\|_1 \leq \frac{C_1}{\sqrt{n-k+1}} \qquad (j = 1,2)$$

and

$$\|q_1^{*j} * q_2^{*(k-j)} * G_{n-k}\|_\infty \leq \frac{C_1}{\sqrt{n-k+1}} \qquad (j = 1,2),$$

so that the asserted properties of the densities $r_n(x)$ follow from the estimate

$$\sum_{k=j}^{n} \frac{\binom{k}{j}(1-\varrho)^j \varrho^{k-j}}{\sqrt{n-k+1}} \leq \sum_{k=j}^{n} \frac{k^j \varrho^{k-j}}{\sqrt{n-k+1}} = \mathcal{O}(n^{-1/2}) \qquad (j = 1,2).$$

Observe that all the terms in the big sum in (7.4) contain the "factor" $q_1^{*2}(\sqrt{n}\,x)$ and therefore have Fourier transforms in L^1. Hence, similarly as in [Al], using Fourier inversion and (5.19), we obtain the representation, for $x > 0$,

$$\begin{aligned}
&\overline{q}_n^*(x) - \sqrt{\tfrac{2}{\pi}}\, e^{-x^2/2} - r_n(x) \\
&\quad = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \Big(\tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \Big)\, dt \\
&\quad + \frac{1}{2\pi} \lim_{R\to\infty} \int_{-R}^{+R} e^{-itx}\, \frac{it}{\sqrt{2\pi n}} \bigg(\sum_{k=3}^{n-1} e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} - \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \bigg)\, dt \\
&\quad + \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \bigg(\sum_{k=3}^{n-1} \Big(\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n} \Big) \overline{\varphi}_{n-k}(t/\sqrt{n}) \bigg)\, dt \\
&\quad + \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \bigg(\sum_{k=3}^{n-1} e^{-kt^2/2n} \Big(\overline{\varphi}_{n-k}(t/\sqrt{n}) - (-\overline{a}_{n-k})\, it/\sqrt{n} \Big) \bigg)\, dt \\
&\quad + \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \bigg(\sum_{k=3}^{n-1} e^{-kt^2/2n} \Big((-\overline{a}_{n-k}) - \frac{1}{\sqrt{2\pi(n-k)}} \Big)\, it/\sqrt{n} \bigg)\, dt\,.
\end{aligned}$$

Denote the integrals on the right-hand side by $I_1, \ldots, I_5$. Note that all the integrals implicitly depend on n and x. We will consider each of them separately.

Convention: We always assume that $n \geq 4$ and $x \in (0,\infty)$ (part (a)) or $x \in (0, e^{-1})$ (part (b)). $\mathcal{O}$- and o-bounds hold uniformly in these regions (unless otherwise mentioned), and they may depend on the constants $\gamma, \delta_1, \delta_2, \delta_3, \ldots$ introduced in Section 5, on the constants C_0 and η in (7.2) and (7.3), and on the L^2-norm of the function g_1.

7.1. **The proof of part (a).** Throughout this subsection we assume that $n \geq 4$ and $x \in (0,\infty)$. The proof is very similar to that of Theorem 1.1 in [Al].

On the Integral I_1. Using integration by parts, we get

$$\begin{aligned}
|I_1| &= \frac{1}{x} \left| \int_{\mathbb{R}} e^{-itx} \frac{d}{dt} \Big[\tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \Big]\, dt \right| \\
&\leq \frac{1}{x} \int_{(-\gamma\sqrt{n},\gamma\sqrt{n})} \left| \frac{d}{dt} \Big[\tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \Big] \right| dt \\
&\quad + \frac{1}{x} \int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} \left| \frac{d}{dt} \Big[\tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \Big] \right| dt\,.
\end{aligned}$$

By (7.1), the first integral on the right is of the order $\mathcal{O}(\delta_n + 2^{-n}) = o(1)$. Furthermore, by (7.3), (5.14) and the fact that $g_1 \in L^2$, the second integral on the right is of the order

$$\begin{aligned}
&\mathcal{O}\bigg(\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} \Big(n\eta^{n-3} |g_1(t/\sqrt{n})|^2 (1/\sqrt{n}) + |t| e^{-t^2/2} \Big)\, dt \bigg) \\
&\qquad\qquad = \mathcal{O}(n\eta^{n-3} + e^{-n\gamma^2/2}) = o(1)\,.
\end{aligned}$$

Thus, $I_1 = o(1/x)$.

On the Integral I_2. By [Al, Equation (24)], we have $I_2 = \mathcal{O}(1/(\sqrt{n}x))$.

On the Integral I_3. For $k = 3, \ldots, n-1$, let

$$I_{3,k} := \int_{\mathbb{R}} e^{-itx}\Big(\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n})\,dt\,.$$

Then, similarly as in [Al], it follows via integration by parts that

$$|I_{3,k}| = \frac{1}{x}\left|\int_{\mathbb{R}} e^{-itx}\frac{d}{dt}\Big[\Big(\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n})\Big]\,dt\right| \\ \leq x^{-1}|I_{3,k,1}| + x^{-1}|I_{3,k,2}|\,,$$

where $I_{3,k,1}$ and $I_{3,k,2}$ denote the integrals over the sets $(-\gamma\sqrt{n}, \gamma\sqrt{n})$ and $(-\gamma\sqrt{n}, \gamma\sqrt{n})^c$, respectively. It follows from (7.1), (5.1) – (5.3), (5.7) and (5.8) that

$$\begin{aligned}|I_{3,k,1}| \leq &\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})}\Big|\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big|\Big|\overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n})\Big|\,dt\\ &+\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})}\Big|\tfrac{d}{dt}\Big[\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big]\Big|\Big|\overline{\varphi}_{n-k}(t/\sqrt{n})\Big|\,dt\\ &=\mathcal{O}\left(\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})}\left[\frac{\delta_k\,e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{2^{-k}}{\sqrt{n(n-k)}}\right]dt\right)\\ &+\mathcal{O}\left(\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})}\left[\frac{\delta_k(k/n)^{1/2}|t|\,e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{2^{-k}}{\sqrt{n(n-k)}}\right]dt\right)\\ &=\mathcal{O}\left(\frac{\delta_k}{\sqrt{k(n-k)}} + \frac{2^{-k}}{\sqrt{n-k}}\right)\,.\end{aligned}$$

Also, using (5.1), (5.3), (5.7), (5.8), (7.2) and (7.3), the Gaussian tail estimates (5.13) – (5.15) and the fact that $g_1 \in L^2$, we get

$$\begin{aligned}|I_{3,k,2}| \leq &\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c}|\tilde{f}_k(t/\sqrt{n})||\overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n})|\,dt\\ &+\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c}|\tilde{f}'_k(t/\sqrt{n})(1/\sqrt{n})||\overline{\varphi}_{n-k}(t/\sqrt{n})|\,dt\\ &+\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c}|e^{-kt^2/2n}||\overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n})|\,dt\\ &+\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c}|e^{-kt^2/2n}(kt/n)||\overline{\varphi}_{n-k}(t/\sqrt{n})|\,dt\\ &=\mathcal{O}\left(\frac{\eta^{k-2} + k\eta^{k-3} + \frac{1}{k\gamma}e^{-k\gamma^2/2} + e^{-k\gamma^2/2}}{\sqrt{n-k}}\right)\,.\end{aligned}$$

Therefore,

$$I_{3,k} = \mathcal{O}\left(\frac{1}{x}\frac{\delta_k + \widetilde{\eta}^k}{\sqrt{k(n-k)}}\right)\,, \tag{7.6}$$

where $\widetilde{\eta} := \frac{1}{2}(1+\max\{\frac{1}{2}, \eta, e^{-\gamma^2/2}\}) \in (0,1)$. Hence, using (5.17), we get $I_3 = o(1/x)$.

On the Integral I_4. It follows from [Al, Equation (47)] that $I_4 = o(1/x)$.

On the Integral I_5. It is shown in [Al, Equation (48)] that $I_5 = o(1/x)$.

Clearly, combining the estimates for $I_1, \dots, I_5$, we get part (a) of Proposition 4.2.

7.2. **The proof of part (b).** Throughout this subsection we assume that $n \geq 4$ and $x \in (0, e^{-1})$. For these values of x, we can obtain somewhat better estimates by avoiding the integration-by-parts step.

On the Integral I_1. We have

$$\begin{aligned} |I_1| &= \left| \int_{\mathbb{R}} e^{-itx} \left[\tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \right] dt \right| \\ &\leq \int_{(-\gamma\sqrt{n}, \gamma\sqrt{n})} \left| \tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \right| dt \\ &\quad + \int_{(-\gamma\sqrt{n}, \gamma\sqrt{n})^c} \left| \tilde{f}_n(t/\sqrt{n}) - e^{-t^2/2} \right| dt \,. \end{aligned}$$

By (7.1), the first integral on the right is of the order $\mathcal{O}(\delta_n + \sqrt{n}\, 2^{-n}) = o(1)$. Furthermore, by (7.2), (5.13) and the fact that $g_1 \in L^2$, the second integral on the right is of the order

$$\begin{aligned} &\mathcal{O}\left(\int_{(-\gamma\sqrt{n}, \gamma\sqrt{n})^c} \left(\eta^{n-2} \, |g_1(t/\sqrt{n})|^2 + e^{-t^2/2} \right) dt \right) \\ &\qquad\qquad = \mathcal{O}(\sqrt{n} \eta^{n-2} + \tfrac{1}{\sqrt{n}} e^{-n\gamma^2/2}) = o(1) \,. \end{aligned}$$

Thus, $I_1 = o(1)$.

On the Integral I_2. We have already mentioned that $I_2 = \mathcal{O}(1/(\sqrt{n}x))$. Now, using (5.12) and (5.19), we also have

$$\begin{aligned} |I_2| &= \left| \lim_{R \to \infty} \int_{-R}^{+R} e^{-itx} \frac{it}{\sqrt{2\pi n}} \left(\sum_{k=3}^{n-1} e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} - \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \right) dt \right| \\ &\leq \sum_{k=3}^{n-1} \int_{-\infty}^{+\infty} \frac{|t|}{\sqrt{2\pi n}} e^{-kt^2/2n} \frac{1}{\sqrt{n-k}} \, dt \\ &\qquad\qquad + \left| \lim_{R \to \infty} \int_{-R}^{+R} e^{-itx} \frac{it}{\sqrt{2\pi n}} \int_0^n e^{-ut^2/2n} \frac{du}{\sqrt{n-u}} \, dt \right| \\ &= \mathcal{O}\left(\sum_{k=3}^{n-1} \frac{n}{k} \frac{1}{\sqrt{n(n-k)}} \right) + \varphi(x) \\ &= \mathcal{O}\left(\sum_{1 \leq k \leq n/2} \frac{1}{k} \right) + \mathcal{O}\left(\sum_{n/2 \leq k \leq n} \frac{1}{\sqrt{n(n-k)}} \right) + \mathcal{O}(1) = \mathcal{O}(\log n) \,. \end{aligned}$$

Thus, $I_2 = \mathcal{O}((\log n) \wedge (1/(\sqrt{n}x)))$.

On the Integral I_3. For $k = 3, \dots, n-1$, we can estimate the integral

$$I_{3,k} := \int_{\mathbb{R}} e^{-itx}\Big(\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n})\,dt\,.$$

in two different ways.

On the one hand, using integration by parts, we obtain

$$I_{3,k} = \mathcal{O}\left(\frac{1}{x}\frac{1}{\sqrt{k(n-k)}}\right)\,, \tag{7.7}$$

see (7.6).

On the other hand, similar estimates (without integration by parts) yield

$$\begin{aligned}
|I_{3,k}| &\le \int_{\mathbb{R}} \Big|\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big|\Big|\overline{\varphi}_{n-k}(t/\sqrt{n})\Big|\,dt \\
&\le \int_{(-\gamma\sqrt{n},\gamma\sqrt{n})} \left[\frac{\delta_k\,|t|\,e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{2^{-k}}{\sqrt{n-k}}\right] dt \\
&\quad + \int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} \Big|\tilde{f}_k(t/\sqrt{n})\Big|\Big|\overline{\varphi}_{n-k}(t/\sqrt{n})\Big|\,dt \\
&\quad + \int_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} \Big|e^{-kt^2/2n}\Big|\Big|\overline{\varphi}_{n-k}(t/\sqrt{n})\Big|\,dt \\
&= \mathcal{O}\left(\frac{n}{k}\frac{\delta_k}{\sqrt{n(n-k)}} + 2^{-k}\frac{\sqrt{n}}{\sqrt{n-k}}\right) \\
&\quad + \mathcal{O}\left(\eta^{k-2}\frac{\sqrt{n}}{\sqrt{n-k}}\right) + \mathcal{O}\left(\frac{1}{k\gamma}e^{-k\gamma^2/2}\frac{\sqrt{n}}{\sqrt{n-k}}\right)\,,
\end{aligned}$$

whence

$$I_{3,k} = \mathcal{O}\left(\frac{n}{k}\frac{1}{\sqrt{n(n-k)}}\right)\,. \tag{7.8}$$

Using (7.7) for $k \le nx^2$ and (7.8) for $k \ge nx^2$, it follows that

$$\begin{aligned}
I_3 = \mathcal{O}\left(\frac{1}{\sqrt{n}x}\sum_{1\le k\le nx^2}\frac{1}{\sqrt{k}} + \sum_{nx^2\le k\le n/2}\frac{1}{k} + \sum_{n/2\le k\le n}\frac{1}{\sqrt{n(n-k)}}\right) \\
= \mathcal{O}(1) + \mathcal{O}(-\log x) + \mathcal{O}(1) = \mathcal{O}(-\log x)\,.
\end{aligned}$$

Thus, $I_3 = \mathcal{O}(-\log x)$.

On the Integral I_4. For $k = 3, \dots, n-1$, we can estimate the integral

$$I_{4,k} := \int_{\mathbb{R}} e^{-itx}e^{-kt^2/2n}\Big(\overline{\varphi}_{n-k}(t/\sqrt{n}) - (-\overline{a}_{n-k})\,it/\sqrt{n}\Big)\,dt$$

in two different ways. On the one hand, using integration by parts and (5.2), (5.3) and (5.8), we have

$$|I_{4,k}| \le \frac{1}{x}\int 2|a_{n-k}|e^{-kt^2/2n}\Big(\tfrac{k}{n}|t|\,|t|/\sqrt{n} + 1/\sqrt{n}\Big)\,dt = \mathcal{O}\left(\frac{1}{x\sqrt{k(n-k)}}\right)\,.$$

On the other hand, also using (5.2) and (5.8) (but without integration by parts), we have

$$|I_{4,k}| \leq \int 2|a_{n-k}|e^{-kt^2/2n}\Big(|t|/\sqrt{n}\Big)\,dt = \mathcal{O}\left(\frac{n}{k}\frac{1}{\sqrt{n(n-k)}}\right).$$

Thus, the same argument as for I_3 leads to the conclusion that $I_4 = \mathcal{O}(-\log x)$.

On the Integral I_5. For $k = 3, \ldots, n-1$, we can estimate the integral

$$I_{5,k} := \int_{\mathbb{R}} e^{-itx}\,e^{-kt^2/2n}\Big((-\overline{a}_{n-k}) - \frac{1}{\sqrt{2\pi(n-k)}}\Big)\,it/\sqrt{n}\,dt$$

in two different ways. On the one hand, using integration by parts and (5.8), we get

$$|I_{5,k}| = \mathcal{O}\left(\frac{1}{x}\int e^{-kt^2/2n}\Big(\frac{\frac{k}{n}|t|\,|t|}{\sqrt{n(n-k)}} + \frac{1}{\sqrt{n(n-k)}}\Big)\,dt\right) = \mathcal{O}\left(\frac{1}{x}\frac{1}{\sqrt{k(n-k)}}\right).$$

On the other hand, using (5.8) (but without integration by parts), we get

$$|I_{5,k}| = \mathcal{O}\left(\int e^{-kt^2/2n}\Big(\frac{|t|}{\sqrt{n(n-k)}}\Big)\,dt\right) = \mathcal{O}\left(\frac{n}{k}\frac{1}{\sqrt{n(n-k)}}\right).$$

Thus, the same argument as for I_3 leads to the conclusion that $I_5 = \mathcal{O}(\log x^{-1})$.

The proof of part (b) of Proposition 4.2 is completed by combining the previous estimates. □

8. Proof of Necessity in Theorem 1.2

Proof of Necessity in Theorem 1.2. Let us quote some well-known results from the literature: Suppose that $|s| < 1$. By Spitzer's formula (see e.g. [Fe, p. 618]), we have

$$\sum_{n=0}^{\infty} s^n \mathbb{E}(e^{it\overline{S}_n^+}) = \frac{1}{1-s}\exp\left(\sum_{k=1}^{\infty}\frac{s^k}{k}\int_0^{\infty}(e^{itx}-1)\,dF_k(x)\right) \tag{8.1}$$

for any $t \in \mathbb{R}$. Also (see e.g. [Fe, p. 416]), we have

$$1 + \sum_{n=1}^{\infty} s^n \mathbb{P}(\overline{S}_n < 0) = \exp\left(\sum_{k=1}^{\infty}\frac{s^k}{k}\mathbb{P}(S_k < 0)\right) = \frac{1}{1-s}\exp\left(-\sum_{k=1}^{\infty}\frac{s^k}{k}\mathbb{P}(S_k \geq 0)\right)$$

for any $t \in \mathbb{R}$. Thus, Spitzer's formula (8.1) can be rewritten as

$$\sum_{n=0}^{\infty} s^n \mathbb{E}(e^{it\overline{S}_n^+}) = \left(1 + \sum_{n=1}^{\infty} s^n \mathbb{P}(\overline{S}_n < 0)\right)\exp\left(\sum_{k=1}^{\infty}\frac{s^k}{k}\int_{[0,\infty)} e^{itx}\,dF_k(x)\right) \tag{8.2}$$

for any $t \in \mathbb{R}$.

Let us note that the preceding results hold without any assumptions on moments or on densities. However, if the moment assumptions stated at the beginning of the Introduction are satisfied, then

$$\mathbb{P}(\overline{S}_n < 0) = \Theta(n^{-1/2}) \tag{8.3}$$

for $n \geq 1$ (see e.g. [Fe, pp. 414f]). Indeed, more precise information is available.

Expanding the right-hand side of Spitzer's formula (8.2) into a power series in s and comparing coefficients, we find that for any $n \geq 1$,

$$\mathbb{E}(e^{it\overline{S}_n^+}) = \overline{F}_n(0) + \overline{F}_{n-1}(0) \int_0^\infty e^{itx} p_{1,+}(x)\, dx$$
$$+ \sum_{m=2}^{n} \overline{F}_{n-m}(0) \sum_{l=1}^{\infty} \sum_{\substack{k_1,\dots,k_l \geq 1: \\ k_1+\dots+k_l=m}} \frac{1}{l!} \frac{1}{k_1 \cdots k_l} \int_0^\infty e^{itx} (p_{k_1,+} * \dots * p_{k_l,+})(x)\, dx\,,$$

where $\overline{F}_0(0) := 1$ and, for any $k \geq 1$, $p_{k,+}(x) := p_k(x)$ for $x > 0$ and $p_{k,+}(x) := 0$ for $x \leq 0$. Hence, by the uniqueness theorem for Fourier transforms, we have

$$\overline{p}_n(x) = \overline{F}_{n-1}(0)\, p_1(x) + \tilde{p}_n(x) \tag{8.4}$$

for almost all $x > 0$, where $\tilde{p}_n$ is a certain subprobability density on the positive half-line.

Now suppose that (1.7) holds. Then, using Lemma 2.1, we have $D(\overline{p}_n^* \,|\, \varphi_+) < \infty$ for all sufficiently large $n \in \mathbb{N}$. It is easy to see that this implies $D(\overline{p}_n \,|\, \varphi_+) < \infty$ for all sufficiently large $n \in \mathbb{N}$. Therefore, using (8.4), (8.3) and the remark (2.2) below Lemma 2.4, we may conclude that $D(p \,|\, \varphi_+) < \infty$, which entails (1.8) by Lemma 2.1. □

9. Proof of Theorem 1.3

Proof of Theorem 1.3. Fix $\varepsilon \in (0,1)$, and let $c \in (0,1)$ and $C \in (1,\infty)$ be such that

$$\int_{-\infty}^{c} \varphi_+(x)\, dx < \varepsilon \quad \text{and} \quad \int_{C}^{\infty} \varphi_+(x)\, dx < \varepsilon\,. \tag{9.1}$$

It then follows from the Erdős–Kac theorem (Theorem 1.1) that

$$\int_{-\infty}^{c} \overline{p}_n^*(x)\, dx < \varepsilon \quad \text{and} \quad \int_{C}^{\infty} \overline{p}_n^*(x)\, dx < \varepsilon \tag{9.2}$$

for all sufficiently large $n \in \mathbb{N}$. Furthermore, it follows from (9.1) and (9.2) that

$$d_{TV}(\overline{S}_n/\sqrt{n}, Z) \leq \int_{-\infty}^{+\infty} |\overline{p}_n^* - \varphi_+|\, dx \leq \int_c^C |\overline{p}_n^* - \varphi_+|\, dx + 4\varepsilon$$

for all sufficiently large $n \in \mathbb{N}$. Now, using Lemma 3.1 (a) and Proposition 4.2, we see that

$$\int_c^C |\overline{p}_n^* - \varphi_+|\; dx \leq \int_c^C |\overline{p}_n^* - \overline{q}_n^*|\; dx + \int_c^C |\overline{q}_n^* - \varphi_+|\; dx = o(1)$$

as $n \to \infty$. Therefore, $d_{TV}(\overline{S}_n/\sqrt{n}, Z) \leq 5\varepsilon$ for all sufficiently large $n \in \mathbb{N}$, and since $\varepsilon \in (0,1)$ is arbitrary, Theorem 1.3 is proved. □

References

[Al] Aleshkyavichene, A. K. (1977): Local limit theorems for the density function of the maximum of sums of independent random variables. *Theory Prob. Appl.* **21**, no. 3, 449–469.

[ABBN] Artstein, S.; Ball, K. M.; Barthe, F.; Naor, A. (2004): On the rate of convergence in the entropic central limit theorem. *Probab. Theory Related Fields* **129**, no. 3, 381–390.

[Ba] Barron, A. R. (1986): Entropy and the central limit theorem. *Ann. Prob.* **14**, no. 1, 336–342.

[BR] Bhattacharya, R. N.; Rao, R. R. (1976): *Normal approximation and asymptotic expansions.* John Wiley & Sons, New York.

[BCG1] Bobkov, S. G.; Chistyakov, G. P.; Götze, F. (2011): Non-uniform bounds in local limit theorems in case of fractional moments I. *Math. Methods Statist.* **20**, no. 3, 171–191.

[BCG2] Bobkov, S. G.; Chistyakov, G. P.; Götze, F. (2011): Non-uniform bounds in local limit theorems in case of fractional moments II. *Math. Methods Statist.* **20**, no. 4, 269–287.

[BCG3] Bobkov, S. G.; Chistyakov, G. P.; Götze, F. (2011): Rate of convergence and Edgeworth-type expansion in the entropic central limit theorem. Accepted for publication in *Ann. Prob.* Preprint, arxiv:1104.3994.

[BCG4] Bobkov, S. G.; Chistyakov, G. P.; Götze, F. (2011): Convergence to stable laws in relative entropy. Accepted for publication in *J. Theoret. Probab.* Preprint, arxiv:1104.4360

[Cs] Csiszár, I. (1967): Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **2**, 299–318.

[DCT] Dembo, A.; Cover, T. M.; Thomas, J. A. (1991): Information-theoretic inequalities. *IEEE Trans. Inform. Theory* **37**, no. 6, 1501–1518.

[EK] Erdős, P.; Kac, M. (1946): On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52**, 292–302.

[FHT] Fedotov, A. A.: Harremoës, P.; Topsøe, F. (2003): Refinements of Pinsker's inequality. *IEEE Trans. Inform. Theory* **49**, no. 6, 1491–1498.

[Fe] Feller, W. (1971): *An introduction to probability theory and its applications, volume II,* 2nd edition. John Wiley & Sons, New York.

[IL] Ibragimov, I. A.; Linnik, Yu. V. (1965): *Independent and stationary sequences of random variables.* Izdat. "Nauka", Moscow.

[JB] Johnson, O.; Barron, A. (2004): Fisher information inequalities and the central limit theorem. *Probab. Theory Related Fields* **129**, no. 3, 391–409.

[Ku] Kullback, S. (1967): A lower bound for discrimination in terms of variation. *IEEE Trans. Inform. Theory* **T-13**, 126–127.

[Jo] Johnson, O. (2004): Information theory and the central limit theorem. Imperial College Press, London.

[N1] Nagaev, S. V. (1969): An estimate of the rate of convergence of the distribution of the maximum of the sums of independent random variables. *Siberian Math. J.* **10**, 443–458.

[N2] Nagaev, S. V. (1970): On the speed of convergence of the distribution of the maximum sums of independent random variables. *Theory Prob. Appl.* **15**, 309–314.

[N3] Nagaev, S. V. (1970): On the speed of convergence in a boundary problem I. *Theory Prob. Appl.* **15**, 163–186.

[Na] Naudziuniene, V. V. (1977): Nonuniform estimates of convergence rate in local limit theorems for densities of the maximum of sums of independent random variables. *Lith. Math. J.* **17**, no. 2, 244–258.

[Pe] Petrov, V. V. (1975): *Sums of independent random variables.* Springer-Verlag, New York.

[Pi] Pinsker, M. S. (1964): *Information and information stability of random variables and processes.* Holden-Day Inc., San Francisco, Calif.

[SM] Sirazhdinov, S. H.; Mamatov, M. (1962): On mean convergence for densities. *Theory Prob. Appl.* **7**, no. 4, 433–437.