J Theor Probab (2015) 28:1520–1555
DOI 10.1007/s10959-014-0550-3

The Entropic Erdős-Kac Limit Theorem

S. G. Bobkov · G. P. Chistyakov · H. Kösters

Received: 15 August 2013 / Revised: 19 February 2014 / Published online: 1 April 2014

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.

Keywords Entropy · Entropic distance · Erdős-Kac limit theorem

Mathematics Subject Classification (2010) 60F05 · 60E10 · 94A15

1 Introduction

Let $\{X_n\}_{n\geq 1}$ be independent identically distributed (i.i.d.) random variables with mean $\mathbb{E}X_1 = 0$ and variance $\mathbb{E}X_1^2 = 1$. Put

$$S_n := \sum_{k=1}^{n} X_k, \qquad \overline{S}_n := \max_{k=1,\dots,n} S_k, \qquad n \in \mathbb{N}.$$

S. G. Bobkov, G. P. Chistyakov and H. Kösters are supported by CRC 701. S. B. was also supported by DMS-1106530 and Simons Fellowship.

S. G. Bobkov
School of Mathematics, University of Minnesota, Minneapolis, MN, USA
e-mail: bobkov@math.umn.edu

G. P. Chistyakov · H. Kösters (✉)
Department of Mathematics, Bielefeld University, Bielefeld, Germany
e-mail: hkoesters@math.uni-bielefeld.de

G. P. Chistyakov
e-mail: chistyak@math.uni-bielefeld.de

Throughout, we denote by Z a standard normal random variable with its density $\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and use the symbol $\Rightarrow$ to denote convergence in distribution.

The classical central limit theorem states that

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

In 1986, Barron [4] established an entropic version of this result, the so-called *entropic central limit theorem*. To formulate it, first let us introduce some notation. Let Y be a random variable with density ψ, and let X be a random variable whose distribution is absolutely continuous with respect to that of Y. The relative entropy of X with respect to Y is defined by

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

where p is the density of X, $L(x) := x \log x$ for $x > 0$ and $L(x) := 0$ for $x = 0$. In case the distribution of X is not absolutely continuous with respect to that of Y, put $D(X \mid Y) := \infty$. Then, the entropic central limit theorem by Barron states that

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

if and only if $D(S_{n_0} \mid Z) < \infty$ for some $n_0 \in \mathbb{N}$. This result is motivated, inter alia, by the distinguished property of the standard normal distribution that it maximizes (Shannon) entropy.

Barron's result has sparked much further research on entropic limit theorems. For instance, there are several publications devoted to the rate of convergence, see Artstein et al. [3], Johnson and Barron [17], Johnson [16], and Bobkov et al. [8]. Entropic limit theorems have also been derived for certain non-normal limit distributions within the class of stable laws, cf. [9,16,19].

All these limit distributions arise in connection with sums of i.i.d. random variables and are therefore infinitely divisible. Our aim is to investigate a different situation, namely for the maxima of sums of i.i.d. summands, with a limit distribution that is not infinitely divisible. Here, the analog of the classical central limit theorem is given by the Erdős-Kac limit theorem [12], which states that

$$\overline{S}_n/\sqrt{n} \Rightarrow |Z| \qquad \text{as } n \to \infty. \tag{1.4}$$

The distribution of $|Z|$, which has density $\varphi_+(x) := \sqrt{\frac{2}{\pi}}\, e^{-x^2/2}\, \mathbf{1}_{(0,\infty)}(x)$, is commonly called the one-sided (or reflected) standard normal law. As explained below, this distribution plays a similar role to the normal distribution in that it maximizes entropy among all *positive* random variables with fixed second moment. It is therefore quite natural to ask whether the Erdős-Kac limit theorem [12] also admits an entropic formulation.

To state a corresponding assertion, we introduce more notation. Given a random variable X such that $\mathbb{P}(X > 0) > 0$, let $\widetilde{X}$ have the same distribution as X conditioned to be positive, i.e. $\mathbb{P}(\widetilde{X} \in A) = \mathbb{P}(X \in A | X > 0)$ for Borel sets A on the real line. Then, the relative entropy of X conditioned to be positive with respect to a positive random variable Y with density ψ is defined by

$$D_+(X \,|\, Y) := D(\widetilde{X} \,|\, Y). \tag{1.5}$$

In the sequel, Y will always be given by $|Z|$ or some scalar multiple of it. Our main result is as follows:

Theorem 1.1 *Suppose that $X_1, X_2, \ldots$ are i.i.d. random variables with a density, mean zero and variance one. Then,*

$$D_+(\overline{S}_n/\sqrt{n} \,|\, |Z|) \to 0 \qquad \text{as } n \to \infty \tag{1.6}$$

if and only if

$$D_+(X_1 \,|\, |Z|) < \infty. \tag{1.7}$$

In fact, the assumption that X_1 has a density is only for convenience and could be omitted. Note, however, that (1.7) implies that X_1 has a density on the positive half-line.

Let us recall that the relative entropy represents a rather strong measure of deviation of distributions. Indeed, by the Pinsker-Csiszár-Kullback inequality,

$$D(X \,|\, Z) \geq \tfrac{1}{2}\,(d_{\mathrm{TV}}(X, Z))^2, \tag{1.8}$$

where $d_{\mathrm{TV}}(X, Z)$ denotes the total variation distance between the distributions of X and Z (cf. [10,13,20,27]). Thus, (1.3) implies $d_{\mathrm{TV}}(S_n/\sqrt{n}, Z) \to 0$ as $n \to \infty$, and hence (1.1). Similarly, (1.6) implies $d_{\mathrm{TV}}(\overline{S}_n/\sqrt{n}, |Z|) \to 0$, and hence (1.4). This follows from (1.8) in combination with the well-known fact that, under our moment assumptions,

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

(cf. e.g. [14, pp. 414f]). In fact, for convergence in total variation distance, condition (1.7) is not needed, and we prove the following:

Theorem 1.2 *Suppose that $X_1, X_2, \ldots$ are i.i.d. random variables with a density, mean zero and variance one. Then,*

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

As already mentioned, both the centered and the one-sided normal distribution play a special role from the viewpoint of information theory. Let us recall that for a random variable X with density p, the entropy (also called Shannon entropy or differential entropy) is defined by

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

where L is as in (1.2). If $\mathbb{E}X^2 = \sigma^2$ is finite, then the entropy is well defined, and

$$h(\sigma Z) - h(X) = D(X \mid \sigma Z) \geq 0 \qquad \big(Z \sim N(0,1)\big),$$

with equality if and only if X and σZ have the same distribution. Thus, the centered normal distribution with second moment σ^2 maximizes entropy among all probability measures with the same second moment. Moreover, in

$$D(X \mid \tau Z) = -h(X) + \tfrac{1}{2}\log(2\pi\tau^2) + \tfrac{1}{2}\sigma^2/\tau^2 \qquad (\tau > 0)$$

the right-hand side is minimized for $\tau = \sigma$, so $D(X \mid \sigma Z)$ may be interpreted as a measure of deviation of the distribution of X from the class of all centered normal distributions. Similarly, for a positive random variable X with finite second moment $\mathbb{E}X^2 = \sigma^2$,

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

with equality if and only if X and $\sigma|Z|$ have the same distribution. Hence, 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, in

$$D_+(X \mid \tau|Z|) = -h(X) + \tfrac{1}{2}\log(\tfrac{1}{2}\pi\tau^2) + \tfrac{1}{2}\sigma^2/\tau^2 \qquad (\tau > 0)$$

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

In this respect, note that

$$\mathbb{E}\big(\max\{\overline{S}_n, 0\}/\sqrt{n}\big)^2 = 1 + o(1) \qquad \text{as } n \to \infty. \tag{1.10}$$

(For instance, this follows from Proposition 6.1 and Eq. (5.8) below.) Combining (1.9) and (1.10), it is easy to 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 to $|Z|$ in (1.6) is natural.

Finally, let us emphasize the following curious difference between the entropic central limit theorem and our Theorem 1.1. Even if X_1 itself has density, Barron's characterization uses the finiteness of $D(S_{n_0} \mid Z)$ for some $n_0 \in \mathbb{N}$ (which may be any natural number); see [4] for an example requiring $n_0 > 1$. In contrast to that, 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 the proof of (1.3) given in Barron [4], entropy convolution inequalities for sums of independent random variables play a major role. In our analysis for the maxima of sums, these inequalities still play a role in the proof of Theorem 1.1, but they have less far-reaching consequences. To control the density of the maximum, we use more classical methods and results based on Fourier analysis, see Nagaev [21,22] and Aleshkyavichene [2]. This approach does not only lead to proofs of entropic limit theorems (cf. [9]), but in principle, similarly as in Bobkov et al. [8], it should also lead to results on the (exact) rate of convergence. Apparently, such refined results cannot be obtained by using known information-theoretic tools.

A major ingredient in our proof will be the local limit theorem for maxima of sums of i.i.d. random variables from [2], see also [1,24,29] for related results. To obtain (1.6) under minimal conditions, we need to extend the result from [2] from bounded to unbounded densities (see Proposition 4.2).

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, 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 $\overline{S}_n$. (Let us emphasize that $\overline{F}_n$ always stands for the distribution function of $\overline{S}_n$ in our paper, and never for the function $1 - F_n$.) We write p_n^* and $\overline{p}_n^*$ for the densities of the rescaled random variables $S_n/\sqrt{n}$ and $\overline{S}_n/\sqrt{n}$.

For a real number x, set $x^+ := \max\{x, 0\}$ and $x^- := \max\{-x, 0\}$. Unless otherwise indicated, $\mathcal{O}$-bounds and o-bounds refer to the case where $n \to \infty$ and hold uniformly in x (in the region under consideration). Finally, $C_1, C_2, \ldots$ 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. Sections 3–7 are devoted to the proof of the sufficiency part of Theorem 1.1, while the necessity part of Theorem 1.1 is proved in Sect. 8. Section 9 contains the proof of Theorem 1.2.

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 p on the real line, set

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

where $L(x)$ is the function defined in the introduction. By abuse of terminology, we will call $D(p \,|\, \psi)$ *relative entropy* even when p is not a probability density on the positive half-line. Note that in the case that p is a probability density, we have $D(p \,|\, \psi) \geq 0$

by Jensen's inequality. If p is an arbitrary non-negative measurable function, this need not be true anymore, but we have at least $D(p \mid \psi) \geq \min\{L(x) : x \geq 0\} = -e^{-1}$.

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 p is a non-negative measurable function with $\int_0^\infty p(x)\,dx < \infty$. Then,*

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

Lemma 2.2 *Suppose that $\alpha_1, \ldots, \alpha_n$ are positive real numbers and $p_1, \ldots, p_n$ are non-negative measurable functions with $\int_0^\infty p_k(x)\,dx < \infty$, $k = 1, \ldots, n$. Then,*

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

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

$$D(p * q \mid \psi) \leq D(p \mid \psi) + e^{-1}.$$

Proof of Lemma 3 Since L is a convex function and q is a probability density on $(-\infty, 0)$, it follows from Jensen's inequality that

$$L\left(\int_{-\infty}^0 h(y)\,q(y)\,\mathrm{d}y \right) \leq \int_{-\infty}^0 L(h(y))\,q(y)\,\mathrm{d}y$$

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

$$\begin{aligned} D(p * q \mid \psi) &= \int_0^\infty L\left(\int_{-\infty}^0 \frac{p(x-y)}{\psi(x)}\,q(y)\,\mathrm{d}y \right) \psi(x)\,\mathrm{d}x \\ &\leq \int_0^\infty \int_{-\infty}^0 L\left(\frac{p(x-y)}{\psi(x)} \right) q(y)\,\mathrm{d}y\,\psi(x)\,\mathrm{d}x \\ &= \int_{-\infty}^0 \int_0^\infty p(x-y) \log\left(\frac{p(x-y)}{\psi(x)} \right) \mathrm{d}x\,q(y)\,\mathrm{d}y. \end{aligned}$$

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

$$\int_0^\infty p(x-y)\log\left(\frac{p(x-y)}{\psi(x)}\right)\mathrm{d}x \le \int_0^\infty p(x-y)\log\left(\frac{p(x-y)}{\psi(x-y)}\right)\mathrm{d}x$$
$$= \int_0^\infty L\left(\frac{p(u)}{\psi(u)}\right)\psi(u)\,\mathrm{d}u - \int_0^{-y} L\left(\frac{p(u)}{\psi(u)}\right)\psi(u)\,\mathrm{d}u \le D(p\,|\,\psi)+e^{-1}.$$

Combining these estimates, we get

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

and the lemma is proved. □

Lemma 2.4 *Suppose that p and q are non-negative measurable functions with $\alpha := \int_0^\infty p(x)\,dx < \infty$ and $\beta := \int_0^\infty q(x)\,dx < \infty$. Then,*

$$\begin{aligned} D(p\,|\,\psi)+D(q\,|\,\psi) &\le D(p+q\,|\,\psi)\\ &\le D(p\,|\,\psi)+D(q\,|\,\psi)+L(\alpha+\beta)-L(\alpha)-L(\beta). \end{aligned}$$

Proof Suppose w.l.o.g. that $\alpha, \beta > 0$. On the one hand, by Lemmas 2.1 and 2.2, we have

$$\begin{aligned} D(p+q\,|\,\psi) &= (\alpha+\beta)\,D\left(\tfrac{\alpha}{\alpha+\beta}\tfrac{p}{\alpha}+\tfrac{\beta}{\alpha+\beta}\tfrac{q}{\beta}\,|\,\psi\right)+L(\alpha+\beta)\\ &\le (\alpha+\beta)\left[\tfrac{\alpha}{\alpha+\beta}D\left(\tfrac{p}{\alpha}\,|\,\psi\right)+\tfrac{\beta}{\alpha+\beta}D\left(\tfrac{q}{\beta}\,|\,\psi\right)\right]+L(\alpha+\beta)\\ &= \alpha D\left(\tfrac{p}{\alpha}\,|\,\psi\right)+\beta D\left(\tfrac{q}{\beta}\,|\,\psi\right)+L(\alpha+\beta)\\ &= D(p\,|\,\psi)+D(q\,|\,\psi)-L(\alpha)-L(\beta)+L(\alpha+\beta). \end{aligned}$$

On the other hand, it is straightforward to check that $L(x+y) \ge L(x)+L(y)$ for any $x, y \ge 0$, whence $D(p+q\,|\,\psi) \ge D(p\,|\,\psi)+D(q\,|\,\psi)$. □

In particular, it follows from Lemmas 2.1 and 2.4 that for any non-negative measurable functions p, q with $\int_0^\infty p(x)\,\mathrm{d}x < \infty$, $\int_0^\infty q(x)\,\mathrm{d}x < \infty$ and any $\alpha, \beta > 0$, we have

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

Lemma 2.5 *Suppose that (p_n) and (q_n) are sequences of non-negative measurable functions such that $\int_0^\infty p_n(x)\,dx = 1+o(1)$ and $\int_0^\infty q_n(x)\,dx = o(1)$ as $n \to \infty$. Then,*

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

Proof This is an immediate consequence of Lemma 2.4. □

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 $n \in \{1, 2, 3, \dots\}$. 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 p, we have

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

3 Binomial Decomposition

In this section, we start with the proof of sufficiency in Theorem 1.1. 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 $\int_{-\infty}^{\infty} |h(x)|\, \mathrm{d}x < \infty$. 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 1.1. Let us remark that binomial decompositions are a well-known tool in the investigation of the classical central limit theorem, see e.g. [15,28]. In connection with entropic central limit theorems, they have recently been used in Bobkov et al. [8,9].

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

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

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

We now need the following formula due to Nagaev [23, Equation (0.8)]: Recall that f denotes the characteristic function of X_1 and $\overline{F}_n$ denotes the distribution function of $\overline{S}_n := \max\{S_1, \dots, S_n\}$. Then, 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

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

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

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

where

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

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

Using (3.2), we may write

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

where

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

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, $\overline{q}_n$ 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

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

$(j = 1, 2)$, 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)

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

Write $\overline{p}_n^*(x) := \sqrt{n}\,\overline{p}_n(\sqrt{n}\,x)$, $\overline{q}_n^*(x) := \sqrt{n}\,\overline{q}_n(\sqrt{n}\,x)$, $\overline{r}_n^*(x) := \sqrt{n}\,\overline{r}_n(\sqrt{n}\,x)$, etc. for the rescaled versions of the above densities. We then have the following result.

Lemma 3.1 *(a)* $\int_0^\infty |\overline{p}_n^*(x) - \overline{q}_n^*(x)|\,dx = \mathcal{O}(n^{-1/2})$.
(b) $\int_0^\infty x^2|\overline{p}_n^*(x) - \overline{q}_n^*(x)|\,dx = \mathcal{O}(n^{-1/2})$.
(c) If (1.7) *holds then* $D(\overline{p}_n^* \,|\, \varphi_+) = D((\overline{q}_n^*)^+ \,|\, \varphi_+) + o(1)$ *as* $n \to \infty$.

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

$$\|p\|_1 := \int_0^\infty |p(x)|\,\mathrm{d}x$$

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,

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

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

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

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

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

and we come to the conclusion that

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

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

We will now show that if (1.7) 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.

Note that $G^-_0 = 0$. For $k = 1, \dots, n-1$, write $G^-_{n-k}(\mathrm{d}x) = \overline{F}_{n-k}(0)\, s_{n-k}(x)\, \mathrm{d}x$, where $s_{n-k}(x) := \overline{p}_{n-k}(x)/\overline{F}_{n-k}(0)$ $(x < 0)$ is a probability density on $(-\infty, 0)$. Also, write $q_2 = \lambda_+ q_{2,+} + \lambda_- q_{2,-}$, where $\lambda_+, \lambda_- \geq 0$, $\lambda_+ + \lambda_- = 1$, 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 twofold application of Lemma 2.2 that

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

For the last step, note that $D(q_{2,-}^{*k} * s_{n-k} \,|\, \varphi_{n,+}) = 0$ and that $\sum_{k=1}^{n-1} \varrho^k \overline{F}_{n-k}(0) = \mathcal{O}(n^{-1/2})$ by (3.9). Using Lemma 2.3 with $f := q_{2,+}^{*j}$ and $g := q_{2,-}^{*(k-j)} * s_{n-k}$, we get

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

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 Dembo et al. [11]), we have

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

We therefore obtain

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

the implicit constants depending only on $q_{2,+}$. Here, the last step follows from (1.7), 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 only slightly changed by the addition of the error terms $\overline{r}_{n,1}^*, \overline{r}_{n,2}^*$ and $(\overline{q}_n^*)^-$. To begin with, it follows from (3.8) that

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

and therefore, since $\|\overline{p}_n^*\|_1 = 1 - \overline{F}_n(0) = 1 + \mathcal{O}(1/\sqrt{n})$ and $\|\overline{r}_{n,j}^*\|_1 = \mathcal{O}(1/\sqrt{n})$ $(j = 1, 2)$,

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

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 $k \geq 1$, 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

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

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

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

This completes the proof of (3.13).

Using (3.8), (3.11), (3.13) as well as Lemma 2.5, we now obtain

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

as $n \to \infty$, and Lemma 3.1 is proved. □

4 Proof of Sufficiency in Theorem 1.1

This section contains the main part of the proof of sufficiency in Theorem 1.1. It relies on two auxiliary results which do not depend on condition (1.7) and whose proof is postponed to the following sections.

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 1.1, 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) Uniformly in $x \in (0, \infty)$,

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

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

$$\overline{q}_n^*(x) = \varphi_+(x) + r_n(x) + \mathcal{O}\big((\log n) \wedge \tfrac{1}{\sqrt{nx}}\big) + \mathcal{O}\big(|\log x|\big) \qquad \text{as } n \to \infty.$$

Here, the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are defined as in the proof of Lemma 3.1. Moreover, by the statement that the $\mathcal{O}$-bounds and o-bounds hold uniformly in x, we mean that for sufficiently large $n \in \mathbb{N}$, the error term is bounded by ε_n/x in part (a), where $(\varepsilon_n)_{n\in\mathbb{N}}$ is a sequence of positive real numbers not depending on $x \in (0, \infty)$ such that $\lim_{n\to\infty} \varepsilon_n = 0$, and by $C_1\big((\log n) \wedge \frac{1}{\sqrt{nx}}\big) + C_2\big(|\log x|\big)$ in part (b), where C_1 and C_2 are positive constants not depending on $x \in (0, e^{-1})$. Similar conventions apply to the error terms in the proof of Proposition 4.2.

Note that Proposition 4.2 may be regarded as a local version of the Erdős-Kac theorem (1.4). 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 square integrable near the origin. This is the crucial point for our purposes.

It should be mentioned that the proof of Proposition 4.2 closely follows that in Aleshkyavichene [2], which is based on earlier work by Nagaev [21–23]. Indeed, in the special case where the 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 [2]) *If $X_1, X_2, \ldots$ have a bounded density $p(x)$, we have $\overline{p}_n^*(x) = \varphi_+(x) + \overline{F}_{n-1}(0)\sqrt{n}p(\sqrt{n}x) + o(1/x)$ as $n \to \infty$, uniformly in $x \in (0, \infty)$.*

Remark In Aleshkyavichene [2] 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.

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

Proof of Sufficiency in Theorem 1.1 Suppose that (1.7) 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 $D_+(\overline{S}_n/\sqrt{n} \,|\, |Z|) \geq 0$, it also follows from the preceding argument that

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

Thus, it remains to show that

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

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

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

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,

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

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.) To complete the proof, we will show that if $C \in (1,\infty)$ is sufficiently large and $c \in (0,1)$ is sufficiently small, then, for each $j \in \{1,2,3\}$, $E_j \leq \varepsilon_0$ for all sufficiently large $n \in \mathbb{N}$.

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

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

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

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

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

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

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 $y \in (0, \infty)$, we get

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

Using Proposition 4.2 (b), it follows that

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

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

$$\int_0^c |r_n(x)|^2 \,\mathrm{d}x \leq \|r_n\|_1 \|r_n\|_\infty = o(1)\,,$$

$$\int_0^c \left((\log n) \wedge \tfrac{1}{\sqrt{nx}} \right)^2 \mathrm{d}x = \frac{\log n}{\sqrt{n}} + \frac{1}{n}(-c^{-1} + \sqrt{n} \log n) = o(1),$$

$$\int_0^c |\log x|^2 \,\mathrm{d}x = \int_{\log(1/c)}^\infty y^2 e^{-y} \,\mathrm{d}y < \infty.$$

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

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

Using Proposition 4.2 (a), it follows that

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

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

This completes the proof of sufficiency in Theorem 1.1. □

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

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

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

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

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

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

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

(see e.g. [2, Equations (26) and (46)]), where

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

(see e.g. [2, 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. [2, 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 $j = 0, 1, 2$,

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

See e.g. [5, 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 $j = 0, 1, 2$,

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

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

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

Such a constant η exists because X_1 has a density, which implies that $|f(t)| < 1$ for all $t \neq 0$ as well as $\lim_{|t|\to\infty} |f(t)| = 0$ (by the Riemann-Lebesgue lemma).

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

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

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

$$\int_t^\infty e^{-\alpha x^2/2}\,\mathrm{d}x \le \sqrt{\tfrac{\pi}{2\alpha}} \wedge \left(\tfrac{1}{\alpha t}e^{-\alpha t^2/2}\right), \tag{5.13}$$

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

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

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

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

Finally, we need the observation that the Fourier transform $\hat{\varphi}_+(t)$ of the density $\varphi_+(x) := \sqrt{2/\pi}e^{-x^2/2}$ $(x > 0)$ satisfies

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

for all $n \in \mathbb{N}$ (see [2, page 452]). It follows from this that for any $x > 0$,

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

(see [2, page 452]).

6 Proof of Proposition 4.1

Recall the constant $\gamma > 0$ and the function $\hat{\varphi}_+(t)$ introduced in the last section. Proposition 4.1 will be deduced from the following result:

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

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

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, this 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 Aleshkyavichene [2], Naudziuniene [25], using (3.3) and (5.18), we have the following decomposition:

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

Denote the expressions in the square brackets by $D_1(t), \ldots, D_6(t)$. (Note that all these expressions implicitly depend on n.) We will show that for $j = 1, \ldots, 6$, 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 $\gamma, \delta_1, \delta_2, \delta_3, \ldots$ introduced in Sect. 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 $\beta \in \{0, 1, 2, \ldots\}$, 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

$$\left|h_\beta(w) - h_\beta(v)\right| = \left|\int_v^w h'_\beta(u)\,\mathrm{d}u\right| = \left|\int_v^w \left(\tfrac{\beta}{u} - \tfrac{t^2}{2n} + \tfrac{1}{2(n-u)}\right) h_\beta(u)\,\mathrm{d}u\right|$$
$$\le (w-v)\left(\tfrac{\beta}{v} + \tfrac{t^2}{2n} + \tfrac{1}{2(n-w)}\right)(w/n)^\beta\, e^{-\mathrm{v}t^2/2n}\,\tfrac{1}{\sqrt{n-w}}.$$

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

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

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,

$$\left|\frac{it^\alpha}{\sqrt{2\pi n}}\left(\sum_{k=3}^{n-1}(k/n)^\beta e^{-kt^2/2n}\frac{1}{\sqrt{n-k}} - \int_0^n (u/n)^\beta e^{-ut^2/2n}\frac{\mathrm{d}u}{\sqrt{n-u}}\right)\right|$$
$$\le \frac{|t|^\alpha}{\sqrt{2\pi n}}\sum_{k=3}^{n-2}\left|\int_k^{k+1}(k/n)^\beta e^{-kt^2/2n}\tfrac{1}{\sqrt{n-k}} - (u/n)^\beta e^{-ut^2/2n}\tfrac{1}{\sqrt{n-u}}\,\mathrm{d}u\right| + \mathcal{O}_\beta(n^{-1/2})$$
$$\le \frac{|t|^\alpha}{\sqrt{2\pi n}}\sum_{k=3}^{n-2}\left(\tfrac{((k+1)/n)^\beta}{(n-k-1)^{3/2}}e^{-kt^2/2n} + \tfrac{((k+1)/n)^\beta}{(n-k-1)^{1/2}}\tfrac{t^2}{2n}e^{-kt^2/2n} + \tfrac{((k+1)/n)^\beta}{(n-k-1)^{1/2}}\tfrac{\beta}{k}e^{-kt^2/2n}\right)$$

$$
\begin{aligned}
&+\mathcal{O}_\beta(n^{-1/2})\\
&=\mathcal{O}_\beta\Bigg(\sum_{k=3}^{n-2}\Big(\tfrac{1}{k^{1/2}(n-k-1)^{3/2}}+\tfrac{1}{k^{3/2}(n-k-1)^{1/2}}+\tfrac{1}{k^{3/2}(n-k-1)^{1/2}}\Big)\Bigg)+\mathcal{O}_\beta(n^{-1/2})\\
&=\mathcal{O}_\beta(n^{-1/2}).
\end{aligned}
$$

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

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

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

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

as well as

$$
\begin{aligned}
&\sum_{k=3}^{n-1}\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big[\Big(f^k(t/\sqrt{n})-e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n})\Big]\\
&=\sum_{k=3}^{n-1}\Big[\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big(f^k(t/\sqrt{n})-e^{-kt^2/2n}\Big)\overline{\varphi}_{n-k}(t/\sqrt{n})\\
&\quad+2\frac{\mathrm{d}}{\mathrm{d}t}\Big(f^k(t/\sqrt{n})-e^{-kt^2/2n}\Big)\overline{\varphi}'_{n-k}(t/\sqrt{n})/\sqrt{n}\\
&\quad+\Big(f^k(t/\sqrt{n})-e^{-kt^2/2n}\Big)\overline{\varphi}''_{n-k}(t/\sqrt{n})/n\Big]
\end{aligned}
$$

$$= \mathcal{O}\Bigg(\sum_{k=3}^{n-1} \Bigg[\frac{\delta_k\,(k/n)\,|t|e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{\delta_k\,(k/n)^{1/2}\,e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{\delta_k\,|\overline{b}_{n-k}|\,e^{-kt^2/4n}}{n} \Bigg] \Bigg)$$

$$= \mathcal{O}\Bigg(\sum_{k=3}^{n-1} \Bigg[\frac{\delta_k \sqrt{k/n}}{\sqrt{n(n-k)}} + \frac{\delta_k \sqrt{k/n}}{\sqrt{n(n-k)}} + \frac{\delta_k |\overline{b}_{n-k}|}{n} \Bigg] \Bigg) = o(1).$$

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

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

Since $\sum_{k=3}^{\infty} x e^{-kx}$ 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

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

as well as

$$\sum_{k=3}^{n-1} e^{-kt^2/2n} \left| \overline{\varphi}''_{n-k}(t/\sqrt{n})/n \right| \le \sum_{k=3}^{n-1} (1/n) e^{-kt^2/2n} |\overline{b}_{n-k}| \le \frac{1}{n} \sum_{k=3}^{n-1} |\overline{b}_{n-k}| = o(1).$$

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

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

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

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

(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 that 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 $T > 0$,

$$\int_{[-T,+T]^c} x^{2k}\, \mathbb{P}_X(\mathrm{d}x) \le \frac{T}{2} \int_{-2/T}^{+2/T} (-1)^k (f_X^{(2k)}(0) - f_X^{(2k)}(t))\, \mathrm{d}t.$$

(See e.g. Lemma 5.1 in Kallenberg [18] for the special case $k = 0$; the general case is similar.) Applying this inequality with $X = \overline{S}_n/\sqrt{n}$ and $T = C$, we get

$$\int_C^\infty x^2\,\overline{p}_n^*(x)\,\mathrm{d}x \le \frac{C}{2}\int_{-2/C}^{+2/C}(-1)\left(f''_{\overline{S}_n/\sqrt{n}}(0)-f''_{\overline{S}_n/\sqrt{n}}(t)\right)\mathrm{d}t$$
$$\le 2\sup_{|t|\le 2/C}\left|f''_{\overline{S}_n/\sqrt{n}}(0)-f''_{\overline{S}_n/\sqrt{n}}(t)\right|.$$

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

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

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

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

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

7 Proof of Proposition 4.2

Write $p = (1-\varrho)q_1 + \varrho q_2$ as in (3.1), and let g_1 and g_2 be the Fourier transforms of q_1 and q_2, respectively. Then,

$$f^k(t) = \sum_{j=0}^k \binom{k}{j}(1-\varrho)^j\varrho^{k-j}g_1^j(t)g_2^{k-j}(t).$$

For $k \ge 3$, put

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

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 and the fact that $\varrho < \frac{1}{2}$, it is easy to see for $k \geq 3$ and $t \in \mathbb{R}$,

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

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

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

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

$$|\tilde{f}_k(t)| \leq \sum_{j=3}^{k} \binom{k}{j} (1-\varrho)^j \varrho^{k-j} \left| g_1^j(t) g_2^{k-j}(t) \right| \leq C_0 \eta^{k-2} |g_1(t)|^2, \tag{7.2}$$

$$|\tilde{f}_k'(t)| \leq \sum_{j=3}^{k} \binom{k}{j} (1-\varrho)^j \varrho^{k-j} \left| \frac{\mathrm{d}}{\mathrm{d}t} \left[g_1^j(t) g_2^{k-j}(t) \right] \right| \leq C_0 k \eta^{k-3} |g_1(t)|^2. \tag{7.3}$$

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 and q_1 and q_2 being probability measures with finite moments.

Recalling (3.2) and (3.6) and 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

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

The functions r_n are the signed densities occurring in Proposition 4.2. It is easy to see that $\|r_n\|_1 = \mathcal{O}(1/\sqrt{n})$ and $\|r_n\|_\infty = \mathcal{O}(1)$. 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

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

so that the asserted properties of the densities r_n follow from the estimate

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

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 Aleshkyavichene [2], using Fourier inversion and (5.19), we obtain the representation, for $x > 0$,

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

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

Convention: We always assume that $n \ge 4$ 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 $\gamma, \delta_1, \delta_2, \delta_3, \dots$ introduced in Sect. 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 $n \ge 4$ and $x \in (0,\infty)$. The proof is very similar to that of Theorem 1 in Aleshkyavichene [2].

On the Integral I_1 Using integration by parts, we get

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

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

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

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

On the Integral I_2 By [2, Equation (24)], we have $I_2 = \mathcal{O}(1/(\sqrt{n}x))$.

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

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

Then, it follows via integration by parts that

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

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

$$\begin{aligned} |I_{3,k,1}| \le &\int_{(-\gamma\sqrt{n},\gamma\sqrt{n})}\Big|\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big|\Big|\overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n})\Big|\,\mathrm{d}t \\ &+ \int_{(-\gamma\sqrt{n},\gamma\sqrt{n})}\left|\frac{\mathrm{d}}{\mathrm{d}t}\Big[\tilde{f}_k(t/\sqrt{n}) - e^{-kt^2/2n}\Big]\right|\Big|\overline{\varphi}_{n-k}(t/\sqrt{n})\Big|\,\mathrm{d}t \end{aligned}$$

$$= \mathcal{O}\left(\int\limits_{(-\gamma\sqrt{n},\gamma\sqrt{n})} \left[\frac{\delta_k\, e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{2^{-k}}{\sqrt{n(n-k)}} \right] \mathrm{d}t \right)$$

$$+ \mathcal{O}\left(\int\limits_{(-\gamma\sqrt{n},\gamma\sqrt{n})} \left[\frac{\delta_k (k/n)^{1/2} |t|\, e^{-kt^2/4n}}{\sqrt{n(n-k)}} + \frac{2^{-k}}{\sqrt{n(n-k)}} \right] \mathrm{d}t \right)$$

$$= \mathcal{O}\left(\frac{\delta_k}{\sqrt{k(n-k)}} + \frac{2^{-k}}{\sqrt{n-k}} \right).$$

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

$$|I_{3,k,2}| \leq \int\limits_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} |\tilde{f}_k(t/\sqrt{n})||\overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n})|\,\mathrm{d}t$$

$$+ \int\limits_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} |\tilde{f}'_k(t/\sqrt{n})(1/\sqrt{n})||\overline{\varphi}_{n-k}(t/\sqrt{n})|\,\mathrm{d}t$$

$$+ \int\limits_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} |e^{-kt^2/2n}||\overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n})|\,\mathrm{d}t$$

$$+ \int\limits_{(-\gamma\sqrt{n},\gamma\sqrt{n})^c} |e^{-kt^2/2n}(kt/n)||\overline{\varphi}_{n-k}(t/\sqrt{n})|\,\mathrm{d}t$$

$$= \mathcal{O}\left(\frac{\eta^{k-2} + k\eta^{k-3} + \frac{1}{k\gamma} e^{-k\gamma^2/2} + e^{-k\gamma^2/2}}{\sqrt{n-k}} \right).$$

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 [2, Equation (47)] that $I_4 = o(1/x)$.

For the convenience of the reader, let us briefly sketch the argument from Aleshkyavichene [2]. For $k = 3, \dots, n-1$, let

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

Using integration by parts, we get

$$|I_{4,k}| \leq \frac{1}{x} \int\limits_{\mathbb{R}} e^{-kt^2/2n} \Big(\tfrac{k}{n} |t| \left| \overline{\varphi}_{n-k}(t/\sqrt{n}) - (-\overline{a}_{n-k})\, it/\sqrt{n} \right|$$
$$+ \left| \overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n}) - (-\overline{a}_{n-k})\, i/\sqrt{n} \right| \Big)\, \mathrm{d}t.$$

We now split the integral at $\pm A (A > 2)$ and use the bounds (5.4) and (5.5) in the region $(-A, +A)$ and the bounds (5.2) and (5.3) in the region $(-A, +A)^c$. In combination with the Gaussian tail estimates (5.13) and (5.15), we obtain

$$\begin{aligned}
&\int\limits_{\mathbb{R}} e^{-kt^2/2n} \Big(\tfrac{k}{n} |t| \left| \overline{\varphi}_{n-k}(t/\sqrt{n}) - (-\overline{a}_{n-k})\, it/\sqrt{n} \right| \\
&\qquad + \left| \overline{\varphi}'_{n-k}(t/\sqrt{n})(1/\sqrt{n}) - (-\overline{a}_{n-k})\, i/\sqrt{n} \right| \Big)\, \mathrm{d}t \\
&\leq \int\limits_{(-A,A)} |\overline{b}_{n-k}| e^{-kt^2/2n} \Big(\tfrac{k}{n} |t|\, t^2/n + |t|/n \Big)\, \mathrm{d}t \\
&\qquad + \int\limits_{(-A,A)^c} 2|\overline{a}_{n-k}| e^{-kt^2/2n} \Big(\tfrac{k}{n} |t|\, |t|/\sqrt{n} + 1/\sqrt{n} \Big)\, \mathrm{d}t \\
&= \mathcal{O}\left(A \frac{|\overline{b}_{n-k}|}{\sqrt{n}\sqrt{k}} \right) \\
&\qquad + \mathcal{O}\left(\frac{|\overline{a}_{n-k}|}{\sqrt{n}} \left[\sqrt{\frac{n\pi}{2k}} \wedge \frac{n}{kA} e^{-kA^2/2n} + \sqrt{\frac{n\pi}{2k}} \wedge \frac{n}{kA} \left(\frac{k}{n} A^2 + 1 \right) e^{-kA^2/2n} \right] \right),
\end{aligned}$$

with implicit constants not depending on n or A. Note that the term in the square brackets is bounded by $\sqrt{2\pi n/k}$ for $k \leq n/A$ and by $(A+2)e^{-A/2}$ for $k \geq n/A$. Thus, using (5.8), it follows that

$$\begin{aligned}
|I_4| &= \mathcal{O}\left(\frac{A}{x} \sum_{k=3}^{n-1} \frac{|\overline{b}_{n-k}|}{\sqrt{n}\sqrt{k}} \right) \\
&\qquad + \mathcal{O}\left(\frac{1}{x} \sum_{k \leq n/A} \frac{1}{\sqrt{k(n-k)}} \right) + \mathcal{O}\left(\frac{(A+2)e^{-A/2}}{x} \sum_{k \geq n/A} \frac{1}{\sqrt{n(n-k)}} \right) \\
&= o(x^{-1} A) + \mathcal{O}(x^{-1} A^{-1/2}) + \mathcal{O}(x^{-1} (A+2) e^{-A/2}).
\end{aligned}$$

Letting $A \equiv A_n \to \infty$ sufficiently slowly as $n \to \infty$, we conclude that $|I_4| = o(1/x)$.

On the Integral I_5 It is shown in [2, Equation (48)] that $I_5 = o(1/x)$.

Clearly, combining the estimates for $I_1, \ldots, I_5$, 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

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

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

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

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

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

$$= \mathcal{O}\Big(\sum_{1\le k\le n/2} \frac{1}{k}\Big) + \mathcal{O}\Big(\sum_{n/2\le k\le n-1} \frac{1}{\sqrt{n(n-k)}}\Big) + \mathcal{O}(1) = \mathcal{O}(\log n).$$

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

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

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

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

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

whence

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

Using (7.7) for $k \leq nx^2$ and (7.8) for $k \geq nx^2$ and recalling that $x \in (0, e^{-1})$, it follows that

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

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

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

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

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

$$|I_{4,k}| \leq \frac{1}{x} \int_{\mathbb{R}} 2|\overline{a}_{n-k}| e^{-kt^2/2n} \Big(\tfrac{k}{n}|t|\,|t|/\sqrt{n} + 1/\sqrt{n}\Big)\, \mathrm{d}t = \mathcal{O}\left(\frac{1}{x\sqrt{k(n-k)}}\right).$$

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

$$|I_{4,k}| \leq \int_{\mathbb{R}} 2|\overline{a}_{n-k}| e^{-kt^2/2n} \Big(|t|/\sqrt{n}\Big)\, \mathrm{d}t = \mathcal{O}\Big(\frac{n}{k} \frac{1}{\sqrt{n(n-k)}}\Big).$$

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 $k = 3, \dots, n-1$, we can estimate the integral

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

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

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

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

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

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

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

8 Proof of Necessity in Theorem 1.1

Let us quote some well-known results from the literature: Suppose that $|s| < 1$. Recall that $\overline{S}_n^+ = \max\{\overline{S}_n, 0\}$. By Spitzer's formula (see e.g. [14, p. 618]), we have

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

for any $t \in \mathbb{R}$. Also (see e.g. [14, p. 416 and p. 428]), we have

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

Thus, Spitzer's formula (8.1) can be rewritten as

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

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 there exist positive constants $c_0 < C_0$ such that

$$c_0 n^{-1/2} \leq \mathbb{P}(\overline{S}_n < 0) \leq C_0 n^{-1/2} \tag{8.3}$$

for all $n \geq 1$ (see e.g. [14, 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 $n \geq 1$,

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

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 (1.6) 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 (1.7) by Lemma 2.1. □

9 Proof of Theorem 1.2

Fix $\varepsilon \in (0, 1)$, and let $c \in (0, 1)$ and $C \in (1, \infty)$ be such that

$$\int_c^C \varphi_+(x)\, \mathrm{d}x > 1 - \varepsilon. \tag{9.1}$$

Then, using Lemma 3.1 (a) and Proposition 4.2 (a), we have

$$\int_c^C \left|\overline{p}_n^* - \varphi_+\right| \mathrm{d}x \leq \int_c^C \left|\overline{p}_n^* - \overline{q}_n^*\right| \mathrm{d}x + \int_c^C \left|\overline{q}_n^* - \varphi_+\right| \mathrm{d}x = o(1) \tag{9.2}$$

as $n \to \infty$, which implies that

$$\int_c^C \overline{p}_n^*(x)\, \mathrm{d}x > 1 - \varepsilon \tag{9.3}$$

for all sufficiently large $n \in \mathbb{N}$. It follows from (9.1)–(9.3) that

$$d_{\mathrm{TV}}(\overline{S}_n/\sqrt{n}, |Z|) \leq \int_{\mathbb{R}} |\overline{p}_n^* - \varphi_+|\, \mathrm{d}x \leq \int_{(c,C)} |\overline{p}_n^* - \varphi_+|\, \mathrm{d}x + \int_{(c,C)^c} (\overline{p}_n^* + \varphi_+)\, \mathrm{d}x < 2\varepsilon$$

for all sufficiently large $n \in \mathbb{N}$. Since $\varepsilon \in (0, 1)$ is arbitrary, Theorem 1.2 is proved. □

Acknowledgments We thank the referee for her or his helpful suggestions.

References

1. Aleshkyavichene, A.K.: Local theorems for the maximum of sums of independent identically distributed random variables. Litovsk. Mat. Sb. **13**(4), 5–21 (1973)

2. Aleshkyavichene, A.K.: Local limit theorems for the density function of the maximum of sums of independent random variables. Theory Probab. Appl. **21**(3), 449–469 (1977)
3. Artstein, S., Ball, K.M., Barthe, F., Naor, A.: On the rate of convergence in the entropic central limit theorem. Probab. Theory Relat. Fields **129**(3), 381–390 (2004)
4. Barron, A.R.: Entropy and the central limit theorem. Ann. Probab. **14**(1), 336–342 (1986)
5. Bhattacharya, R.N., Rao, R.R.: Normal Approximation and Asymptotic Expansions. Wiley, New York (1976)
6. Bobkov, S.G., Chistyakov, G.P., Götze, F.: Non-uniform bounds in local limit theorems in case of fractional moments I. Math. Methods Stat. **20**(3), 171–191 (2011)
7. Bobkov, S.G., Chistyakov, G.P., Götze, F.: Non-uniform bounds in local limit theorems in case of fractional moments II. Math. Methods Stat. **20**(4), 269–287 (2011)
8. Bobkov, S.G., Chistyakov, G.P., Götze, F.: Rate of convergence and Edgeworth-type expansion in the entropic central limit theorem. Ann. Probab. **41**(4), 2479–2512 (2011)
9. Bobkov, S.G., Chistyakov, G.P., Götze, F.: Convergence to stable laws in relative entropy. J. Theor. Prob. **26**(3), 803–818 (2013)
10. Csiszár, I.: Information-type measures of difference of probability distributions and indirect observations. Stud. Sci. Math. Hungar. **2**, 299–318 (1967)
11. Dembo, A., Cover, T.M., Thomas, J.A.: Information-theoretic inequalities. IEEE Trans. Inf. Theory **37**(6), 1501–1518 (1991)
12. Erdős, P., Kac, M.: On certain limit theorems of the theory of probability. Bull. Am. Math. Soc. **52**, 292–302 (1946)
13. Fedotov, A.A., Harremoës, P., Topsøe, F.: Refinements of Pinsker's inequality. IEEE Trans. Inf. Theory **49**(6), 1491–1498 (2003)
14. Feller, W.: An Introduction to Probability Theory and its Applications, vol. II, 2nd edn. Wiley, New York (1971)
15. Ibragimov, I.A., Linnik, Y.V.: Independent and stationary sequences of random variables. Izdat. "Nauka", Moscow (1965)
16. Johnson, O.: Information Theory and the Central Limit Theorem. Imperial College Press, London (2004)
17. Johnson, O., Barron, A.: Fisher information inequalities and the central limit theorem. Probab. Theory Relat. Fields **129**(3), 391–409 (2004)
18. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2002)
19. Kontoyiannis, I., Harremoës, P., Johnson, O.: Entropy and the law of small numbers. IEEE Trans. Inf. Theory **51**(2), 466–472 (2005)
20. Kullback, S.: A lower bound for discrimination in terms of variation. IEEE Trans. Inf. Theory **T–13**, 126–127 (1967)
21. Nagaev, S.V.: An estimate of the rate of convergence of the distribution of the maximum of the sums of independent random variables. Sib. Math. J. **10**, 443–458 (1969)
22. Nagaev, S.V.: On the speed of convergence of the distribution of the maximum sums of independent random variables. Theory Probab. Appl. **15**, 309–314 (1970)
23. Nagaev, S.V.: On the speed of convergence in a boundary problem I. Theory Probab. Appl. **15**, 163–186 (1970)
24. Nagaev, S.V., Eppel, M.S.: On a local limit theorem for the maximum of sums of independent random variables. Theory Probab. Appl. **21**(2), 384–385 (1976)
25. Naudziuniene, V.V.: Nonuniform estimates of convergence rate in local limit theorems for densities of the maximum of sums of independent random variables. Lith. Math. J. **17**(2), 244–258 (1977)
26. Petrov, V.V.: Sums of Independent Random Variables. Springer, New York (1975)
27. Pinsker, M.S.: Information and Information Stability of Random Variables and Processes. Holden-Day Inc., San Francisco (1964)
28. Sirazhdinov, S.H., Mamatov, M.: On mean convergence for densities. Theory Probab. Appl. **7**(4), 433–437 (1962)
29. Wachtel, V.: Local limit theorem for the maximum of asymptotically stable random walks. Probab. Theory Relat. Fields **152**(3), 407–424 (2012)