Explicit non-asymptotic bounds for the distance to the first-order Edgeworth expansion

Alexis Derumigny*, Lucas Girard†, Yannick Guyonvarch†,‡

January 15, 2021

Abstract

In this article, we study bounds on the uniform distance between the cumulative distribution function of a standardized sum of independent centered random variables with moments of order four and its first-order Edgeworth expansion. Existing bounds are sharpened in two frameworks: when the variables are independent but not identically distributed and in the case of independent and identically distributed random variables. Improvements of these bounds are derived if the third moment of the distribution is zero. We also provide adapted versions of these bounds under additional regularity constraints on the tail behavior of the characteristic function. We finally present an application of our results to the lack of validity of one-sided tests based on the normal approximation of the mean for a fixed sample size.

Keywords: Berry-Esseen bound, Edgeworth expansion, normal approximation, central limit theorem, non-asymptotic tests.

MCS 2020: Primary: 62E17; Secondary: 60F05, 62F03.

1 Introduction

As the number of observations n in a statistical experiment goes to infinity, many statistics of interest have the property to converge weakly to a $\mathcal{N}(0,1)$ distribution, once adequately centered and scaled, see e.g. van der Vaart (2000, Chapter 5) for a thorough introduction. Hence, when little is known on the distribution of a standardized statistic S_n for a fixed sample size $n > 0$, a classical approach to conduct inference on the parameters of the statistical model amounts to approximate the distribution of that statistic by its (tractable) Gaussian limit.

A natural and recurring theme in statistics and probability is thus to quantify how far the $\mathcal{N}(0,1)$ distribution lies from the unknown distribution of S_n for a given n. This article aims to present some refined results in one of the simplest and most studied cases: when S_n is a standardized sum of independent random variables. We consider independent but not necessarily identically distributed random variables to encompass a broader range of applications. For instance, certain bootstrap schemes such as the multiplier ones (see Chapter 9 in van der Vaart and Wellner (1996) or Chapter 10 in Kosorok (2006)) boil down to studying a sequence of (mutually) independent not necessarily identically distributed (*i.n.i.d.*) random variables conditionally on the initial sample.

*University of Twente, 5 Drienerlolaan, 7522 NB Enschede, Netherlands. a.f.f.derumigny@utwente.nl.
†CREST, 5, avenue Henry Le Chatelier, 91764 Palaiseau cedex, France. lucas.girard@ensae.fr
‡Télécom Paris, 19 Place Marguerite Perey, 91120 Palaiseau, France. yannick.guyonvarch@telecom-paris.fr.

More formally, let $(X_i)_{i=1,\dots,n}$ be a sequence of *i.n.i.d.* random variables satisfying for every $i \in \{1, ..., n\}$, $\mathbb{E}[X_i] = 0$ and $\gamma_i := \mathbb{E}[X_i^4] < +\infty$. We also define the standard deviation B_n of the sum of the X_i's, i.e. $B_n := \sqrt{\sum_{i=1}^n \mathbb{E}[X_i^2]}$, so that the standardized sum can be written as $S_n := \sum_{i=1}^n X_i / B_n$. Finally, we use the average individual standard deviation $\overline{B}_n := B_n/\sqrt{n}$ and the average standardized third moment $\lambda_{3,n} := \frac{1}{n}\sum_{j=1}^n \mathbb{E}[X_j^3]/\overline{B}_n^3$. The main results of this article are of the form

$$\underbrace{\sup_{x\in\mathbb{R}} \left| \mathbb{P}\left(S_n \leq x\right) - \Phi(x) - \frac{\lambda_{3,n}}{6\sqrt{n}}(1-x^2)\varphi(x) \right|}_{=: \Delta_{n,\mathrm{E}}} \leq \delta_n, \tag{1}$$

where Φ is the cumulative distribution function of a standard Gaussian random variable, φ is the density function associated with Φ, and δ_n is a positive sequence that depends on the first four moments of $(X_i)_{i=1,\dots,n}$ and tends to zero under some regularity conditions. In the following, we use the notation $G_n(x) := \Phi(x) + \frac{\lambda_{3,n}}{6\sqrt{n}}(1-x^2)\varphi(x)$.

The quantity $G_n(x)$ is usually called the one-term Edgeworth expansion of $\mathbb{P}\left(S_n \leq x\right)$, hence the letter E in the notation $\Delta_{n,\mathrm{E}}$. Controlling the uniform distance between $\mathbb{P}\left(S_n \leq x\right)$ and $G_n(x)$ has a long tradition in statistics and probability, see for instance Esseen (1945) and the books by Cramer (1962) and Bhattacharya and Ranga Rao (1976). As early as in the work of Esseen (1945), it was acknowledged that in independent and identically distributed (*i.i.d.*) cases, $\Delta_{n,\mathrm{E}}$ was of the order $n^{-1/2}$ in general and of the order n^{-1} if $(X_i)_{i=1,\dots,n}$ has a nonzero continuous component. These results were then extended in a wide variety of directions, often in connection with bootstrap procedures, see for instance Hall (1992) and Lahiri (2003) for the dependent case.

A one-term Edgeworth expansion can be seen as a refinement of the so-called Berry-Esseen inequality (Berry (1941), Esseen (1942)) which goal is to bound

$$\Delta_{n,\mathrm{B}} := \sup_{x\in\mathbb{R}} \left| \mathbb{P}\left(S_n \leq x\right) - \Phi(x) \right|.$$

The refinement stems from the fact that in $\Delta_{n,\mathrm{E}}$, the distance between $\mathbb{P}\left(S_n \leq x\right)$ and $\Phi(x)$ is adjusted for the presence of non-asymptotic skewness in the distribution of S_n. Contrary to the the literature on Edgeworth expansions, there is a substantial amount of work devoted to explicit constants in the Berry-Esseen inequality and its extensions, see e.g. Bentkus and Götze (1996), Bentkus (2003), Pinelis and Molzon (2016), Chernozhukov et al. (2017), Raič (2018), Raič (2019). The sharpest known result in the *i.n.i.d.* univariate framework is due to Shevtsova (2010), which shows that for every $n \in \mathbb{N}^*$, if $\mathbb{E}[|X_i|^3] < +\infty$ for every $i \in \{1, ..., n\}$, then $\Delta_{n,\mathrm{B}} \leq 0.56\, K_{3,n}/\sqrt{n}$ where $K_{p,n} := n^{-1}\sum_{i=1}^n \mathbb{E}[|X_i|^p]/(\overline{B}_n)^p$, $p \in \mathbb{N}^*$, denotes the average standardized p-th absolute moment. For $p \in \mathbb{N}^*$, $K_{p,n}$ measures the tail thickness of the distribution, with $K_{2,n}$ normalized to 1 and $K_{4,n}$ the kurtosis. An analogous result is given in Shevtsova (2011) under the *i.i.d.* assumption where 0.56 is replaced with 0.4748. A close lower bound is due to Esseen (1956): there exists a distribution such that $\Delta_{n,\mathrm{B}} = \frac{C_B}{\sqrt{n}}\left(n^{-1}\sum_{i=1}^n \mathbb{E}[|X_i|^3]/\overline{B}_n^3\right)$ with $C_B \approx 0.4098$. Another line of research uses Edgeworth expansions to get a bound on $\Delta_{n,\mathrm{B}}$ that contains higher-order terms, see Adell and Lekuona (2008), Boutsikas (2011) and Zhilova (2020).

Despite the breadth of those theoretical advances, there remain some limits to take full advantage of those results even in simple statistical applications, in particular when conducting inference on the expectation of a real random variable.[1] If we focus on Berry-Esseen inequalities, Example 1.1 shows that even the sharpest upper bound to date on $\Delta_{n,\mathrm{B}}$ can be quite uninformative when conducting inference on an expectation even for n larger than 80,000. It is therefore natural to wonder whether bounds derived from a one-term Edgeworth expansion could be tighter in moderately large samples (such as a few thousands). In the *i.i.d.* case and under some smoothness conditions, Senatov (2011) obtains such improved bounds. To our knowledge, the question is nevertheless still open in the *i.n.i.d.* setup, as well as in the general setup when no condition on the characteristic function is assumed. In particular, most articles that present results of the form of (1) do not provide a fully explicit value for δ_n, that is, δ_n is defined up to some "universal" but unknown constant.

In this article, we derive novel inequalities of the form (1) that aim to be relevant in practical applications. Such "user-friendly" bounds seek to achieve two goals. First, we provide explicit values for δ_n. Second, the bounds δ_n should be small enough to be informative even in small ($n \approx$ hundreds) to moderate ($n \approx$ thousands) sample sizes. We obtain these bounds in an *i.i.d.* setting and in a more general *i.n.i.d.* case only assuming finite fourth moments.

We give improved bounds on $\Delta_{n,\mathrm{E}}$ when we assume some regularity assumptions on the tail behavior of the characteristic function f_{S_n} of S_n. Such conditions are related to the continuity of the distribution of S_n with respect to Lebesgue's measure and the differentiability of the corresponding density. These are well-known conditions required for the Edgeworth expansion to be a good approximation of $\mathbb{P}(S_n \leq \cdot\,)$ with fast rate. Our main results are summed up in Table 1.

Setup	*General case*	*Under regularity assumptions on* $f_{S_n}(\cdot)$
i.n.i.d.	$\dfrac{0.399 K_{3,n}}{\sqrt{n}} + O(n^{-1})$ (Theorem 2.3)	$\dfrac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2}{n} + O(n^{-5/4} + n^{-p/2})$ (Corollary 3.2)
i.i.d.	$\dfrac{0.2(K_{3,n} + 1)}{\sqrt{n}} + O(n^{-1})$ (Theorem 2.3)	$\dfrac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2}{n} + O(n^{-5/4})$ (Corollary 3.4)

Table 1: Summary of the new bounds on $\Delta_{n,\mathrm{E}}$ under different scenarios. All the remainder terms are given with explicit expressions and are significantly reduced when there is no skewness. $p \geq 1$ is a constant depending on the smoothness of the characteristic function f_{S_n}.

In the rest of this section, we provide more details about the lack of information given by the Berry-Esseen inequality in Example 1.1 and introduce notation that is used in the rest of the paper. Section 2 presents our bounds on $\Delta_{n,\mathrm{E}}$ in the general *i.n.i.d.* case and in the

[1] In this article, we only give results for *standardized* sums of random variables, i.e. sums that are rescaled by their standard deviation. In practice, the variance is unknown and has to be replaced with some empirical counterpart, leading to what is usually called a *self-normalized* sum. This is an important question in practice that we leave aside for future research. There exist numerous results on self-normalized sums in the fields of Edgeworth expansions and Berry-Esseen inequalities (Hall (1987), de la Peña et al. (2009)) but the practical limitations that we point out in this work still prevail.

i.i.d. setting. In Section 3, we develop tighter bounds under regularity assumptions on the characteristic function of S_n. In Section 4, we apply our results to show that one-sided tests based on the normal approximation of a sample mean do not hold their nominal level in presence of non-asymptotic skewness. The main proofs are gathered in Section 5 and some useful lemmas are proved in Section 6.

Example 1.1 (The lack of information conveyed by the Berry-Esseen inequality for inference on an expectation)**.** *Let $(Y_i)_{i=1,\dots,n}$ an i.i.d. sequence of random variable with expectation μ, known variance σ^2 and finite fourth moment with $\kappa := \mathbb{E}\left[(Y_1-\mu)^4\right]/\sigma^4$ the kurtosis of the distribution of Y_1. We want to conduct a test with null hypothesis $H_0 : \mu = \mu_0$, for some real number μ_0, and alternative $H_1 : \mu > \mu_0$ with a type-one error at most $\alpha \in (0,1)$, and ideally equal to α. The classical approach to this problem amounts to comparing S_n (with $X_i = Y_i - \mu_0$) with the $1-\alpha$ quantile of the $\mathcal{N}(0,1)$ distribution, denoted $q_{\mathcal{N}(0,1)}(1-\alpha)$. Thanks to the Berry-Esseen inequality, we are able to quantify the mistake caused by using the $\mathcal{N}(0,1)$ approximation*

$$\Big|\mathbb{P}\big(S_n \leq q_{\mathcal{N}(0,1)}(1-\alpha)\big) - (1-\alpha)\Big| \leq \frac{0.4748\,\mathbb{E}\big[|Y_1-\mu_0|^3\big]}{\sqrt{n}\sigma^3} \leq \frac{0.4748\,\kappa^{3/4}}{\sqrt{n}}, \tag{2}$$

where the probability and expectation operators are to be understood under the null hypothesis H_0. Inequality (2) *is called* weakly informative *as long as $\frac{0.4748\kappa^{3/4}}{\sqrt{n}} \geq \alpha$, since in that case, we cannot exclude that $\mathbb{P}\left(S_n \leq q_{\mathcal{N}(0,1)}(1-\alpha)\right)$ is arbitrarily close to 1 and therefore arbitrarily conservative.*

We denote by $n_{\max}(\alpha)$ the largest weakly informative n for a given level α. We note that imposing $\kappa \leq 9$ allows for a wide array of distributions used in practice: any Gaussian, Gumbel, Laplace, Uniform, or Logistic distribution satisfies it, as well as any Student with at least 5 degrees of freedom, any Gamma or Weibull with shape parameter at least 1. Plugging $\kappa \leq 9$ into (2)*, we remark that for $\alpha = 0.1$, the bound is weakly informative for $n \leq 608$, namely $n_{\max}(0.1) = 608$. For $\alpha = 0.05$, we obtain $n_{\max}(\alpha) = 2,434$. Finally, the situation deteriorates strikingly for $\alpha = 0.01$ where the bound is weakly informative for $n \leq n_{\max}(\alpha) = 60,867$.*

Additional notation. $\vee$ (*resp.* $\wedge$) denotes the maximum (*resp.* minimum) operator. For a random variable V, we denote its probability distribution by P_V. For a distribution P, let f_P denote its characteristic function; similarly, for a random variable X, we denote by f_X its characteristic function. We recall that $f_{\mathcal{N}(0,1)}(t) = e^{-t^2/2}$. For two sequences (a_n), (b_n), we write $a_n = O(b_n)$ whenever there exists $C > 0$ such that $a_n \leq Cb_n$; $a_n = o(b_n)$ whenever $a_n/b_n \to 0$; and $a_n \asymp b_n$ whenever $a_n = O(b_n)$ and $b_n = O(a_n)$. We denote by χ_1 the constant $\chi_1 := \sup_{x>0} x^{-3}|\cos(x) - 1 + x^2/2| \approx 0.099162$ (Shevtsova, 2010), and by θ_1^* the unique root in $(0, 2\pi)$ of the equation $\theta^2 + 2\theta\sin(\theta) + 6(\cos(\theta) - 1) = 0$. We also define $t_1^* := \theta_1^*/(2\pi) \approx 0.635967$ (Shevtsova, 2010). For every $i \in \mathbb{N}^*$, we define the individual standard deviation $\sigma_i := \sqrt{\mathbb{E}[X_i^2]}$. Henceforth, we reason for a fixed but arbitrary sample size $n \in \mathbb{N}^*$.

2 Control of $\Delta_{n,\mathrm{E}}$ under moment conditions only

We start by introducing two versions of our basic assumptions on the distribution of the variables $(X_i)_{i=1,\dots,n}$.

Assumption 2.1 (Moment conditions in the *i.n.i.d.* framework)**.** *$(X_i)_{i=1,\dots,n}$ are independent and centered random variables such that for every $i \in \{1,\dots,n\}$, the fourth raw individual moment $\gamma_i := \mathbb{E}[X_i^4]$ is positive and finite.*

Assumption 2.2 (Moment conditions in the *i.i.d.* framework)**.** *$(X_i)_{i=1,\dots,n}$ are i.i.d. centered random variables such that the fourth raw moment $\gamma_n := \mathbb{E}[X_n^4]$ is positive and finite.*

Assumption 2.2 corresponds to the classical *i.i.d.* sampling with finite fourth moment while Assumption 2.1 is its generalization in the *i.n.i.d.* framework. Those two assumptions primarily ensure that enough moments of $(X_i)_{i=1,\dots,n}$ exist to build a non-asymptotic upper bound on $\Delta_{n,\mathrm{E}}$. In some applications, such as the bootstrap, it is required to consider an array of random variables $(X_{i,n})_{i=1,\dots,n}$ instead of a sequence. For example, Efron (1979)'s nonparametric bootstrap procedure consists in drawing n elements in the random sample $(X_{1,n}, ..., X_{n,n})$ with replacement. Conditional on $(X_{i,n})_{i=1,\dots,n}$, the n values drawn with replacement can be seen as a sequence of n *i.i.d.* random variables with distribution $\frac{1}{n}\sum_{i=1}^n \delta_{\{X_{i,n}\}}$, denoting by $\delta_{\{a\}}$ the Dirac measure at a given point $a \in \mathbb{R}$).

Our results encompass these situations directly. Nonetheless, we do not use the array terminology here as our results hold non-asymptotically, i.e. for any fixed sample size n. We can now state the main result of this section. It is proved in Sections 5.2 and 5.3.

Theorem 2.3 (Control of the one-term Edgeworth expansion with bounded moments of order four)**.** *If Assumption 2.1 or Assumption 2.2 holds, we have the bound*

$$\Delta_{n,E} \leq \frac{0.1995\, \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{0.031\, \widetilde{K}_{3,n}^2 + 0.195\, K_{4,n} + 0.054\, |\lambda_{3,n}| \widetilde{K}_{3,n} + 0.038\, \lambda_{3,n}^2}{n} + r_{1,n}\,, \qquad (3)$$

where $\widetilde{K}_{3,n} := K_{3,n} + \frac{1}{n}\sum_{i=1}^n \mathbb{E}|X_i|\sigma_i^2/\overline{B}_n^3$ and $r_{1,n}$ is a remainder term that depends only on $K_{3,n}$, $K_{4,n}$ and $\lambda_{3,n}$, as defined in Equation (22) *(resp. in Equation* (29) *under Assumption 2.2).*

When the fourth moment is bounded, i.e. $K_{4,n} = O(1)$, we get the rate $r_{1,n} = O(n^{-5/4})$.

If furthermore $\mathbb{E}[X_i^3] = 0$ for every $i = 1, \dots, n$, then

i. the average third raw moment is null, i.e. $\lambda_{3,n} = 0$,

ii. the first-order term in the Edgeworth expansion is null too so that $\Delta_{n,B} = \Delta_{n,E}$,

iii. the residual term $r_{1,n}$ converges to 0 at the faster rate $n^{-3/2}$ under Assumption 2.1 (resp. at the rate n^{-2} under Assumption 2.2).

Note that, in Theorem 2.3, it is possible to replace $\widetilde{K}_{3,n}$ by the simpler upper bound $2K_{3,n}$ (or by $K_{3,n} + 1$ under Assumption 2.2). This theorem displays a bound of order $n^{-1/2}$ on $\Delta_{n,\mathrm{E}}$. The rate $n^{-1/2}$ cannot be improved when only assuming moment conditions on $(X_i)_{i=1,\dots,n}$ (Esseen (1945), Cramer (1962)). The numerical constants that show up in the leading terms of the bound are quite small. Numerical evaluation of the remainder terms $r_{1,n}$ (using its explicit expression) shows they are small in practice too. Another nice aspect of those bounds is their dependence on $\lambda_{3,n}$. For many classes of distributions, $\lambda_{3,n}$ can in fact be exactly zero. This is the case if for every $i = 1, \dots, n$, X_i has a non-skewed distribution, such as any distribution that is symmetric around its expectation. More generally, $|\lambda_{3,n}|$ can be substantially smaller than $K_{3,n}$, decreasing the corresponding terms.

This theorem further gives a bound on $\Delta_{n,\mathrm{B}}$, even in the case where the third raw moments are not null. Indeed, using Theorem 2.3, the bound $(1 - x^2)\varphi(x)/6 \leq \varphi(0)/6 \leq 0.067$ for all $x \in \mathbb{R}$, and applying the triangle inequality, we remark that for every $n \in \mathbb{N}^*$

$$\Delta_{n,\mathrm{B}} \leq \frac{0.1995 \widetilde{K}_{3,n} + 0.067|\lambda_{3,n}|}{\sqrt{n}} + O(n^{-1}). \qquad (4)$$

Under Assumption 2.1, using the refined inequality $|\lambda_{3,n}| \leq 0.621 K_{3,n}$ (Pinelis, 2011, Theorem 1), we can derive a simpler bound that involves only $K_{3,n}$

$$\frac{0.1995 \widetilde{K}_{3,n} + 0.067 |\lambda_{3,n}|}{\sqrt{n}} \leq \frac{0.441 K_{3,n}}{\sqrt{n}}.$$

The bound $\Delta_{n,\mathrm{B}} \leq 0.441 K_{3,n}/\sqrt{n} + O(n^{-1})$ is already tighter than the sharpest known Berry-Esseen inequality in the *i.n.i.d.* framework, $\Delta_{n,\mathrm{B}} \leq 0.56 K_{3,n}/\sqrt{n}$, as soon as the remainder term $O(n^{-1})$ is smaller than the difference $0.12 K_{3,n}/\sqrt{n}$. Even in the *i.i.d.* case, this bound is still tighter than the sharpest known Berry-Esseen inequality, $\Delta_{n,\mathrm{B}} \leq 0.4748 K_{3,n}/\sqrt{n}$, up to a $O(n^{-1})$ term. We refer to Example 2.4 and Figure 1 for a numerical comparison, showing improvements for n of the order of a few thousands.

The most striking improvement is obtained whenever $\mathbb{E}[X_i^3] = 0$ for every integer i. In this case, Theorem 2.3 and the inequality $\widetilde{K}_{3,n} \leq 2K_{3,n}$ yield $\Delta_{n,\mathrm{B}} \leq 0.399 K_{3,n}/\sqrt{n} + O(n^{-1})$. Note that this result does not contradict Esseen (1956)'s lower bound $0.4098 K_{3,n}/\sqrt{n}$ as the distribution he constructs does not satisfy $\mathbb{E}[X_i^3] = 0$ for every i.

Under Assumption 2.2, $\widetilde{K}_{3,n} = K_{3,n} + 1$ and we can combine this with (4) and the inequality $|\lambda_{3,n}| \leq 0.621 K_{3,n}$ (Pinelis, 2011, Theorem 1), so that we obtain

$$\Delta_{n,\mathrm{B}} \leq \frac{0.1995(K_{3,n}+1) + 0.067 \times 0.621 K_{3,n}}{\sqrt{n}} + O(n^{-1}) \leq \frac{0.242 K_{3,n} + 0.1995}{\sqrt{n}} + O(n^{-1}).$$

One may find this result surprising, given that the numerical constant in front of $K_{3,n}$ in the leading term is smaller than the lower bound constant $C_B := 0.4098$ derived in Esseen (1956). The point is addressed in detail in Shevtsova (2012), where the author explains that the constant cannot be improved if one seeks a control of $\Delta_{n,\mathrm{B}}$ with leading term of the form $C K_{3,n}/\sqrt{n}$ for some $C > 0$. In contrast, our control of $\Delta_{n,\mathrm{B}}$ exhibits a leading term of the form $(C K_{3,n} + c)/\sqrt{n}$.

Example 2.4 (Implementation of our bounds on $\Delta_{n,\mathrm{B}}$)**.** *Theorem 2.3 provides new tools to control $\Delta_{n,B}$, and we compare them with existing results. To compute our bounds on $\Delta_{n,B}$ as well as previous ones, we need numerical values for $\widetilde{K}_{3,n}$, $\lambda_{3,n}$, and $K_{4,n}$ or upper bounds thereon. A bound on $K_{4,n}$ is in fact sufficient to control $\lambda_{3,n}$ and $K_{3,n}$: Pinelis (2011) ensures $|\lambda_{3,n}| \leq 0.621 K_{3,n}$, and a convexity argument yields $K_{3,n} \leq K_{4,n}^{3/4}$. Moreover, since the third standardized absolute moment has no particular statistical signification, it is not so intuitive to find a natural bound on $K_{3,n}$. On the contrary, the fourth standardized moment $K_{4,n}$ is well-known as the kurtosis of a distribution (the thickness of the tails compared to the central part of the distribution). As explained in Example 1.1, in the i.i.d. framework, imposing $K_{4,n} \leq 9$ is a reasonable assumption. For the sake of comparison, we also impose $K_{4,n} \leq 9$ in the i.n.i.d. case.*

We consider the improved bounds that rely on $\lambda_{3,n} = 0$ as well. The different bounds are plotted as a function of n in Figure 1:

- *Shevtsova (2010) i.n.i.d.:* $\frac{0.56}{\sqrt{n}} K_{3,n} \leq \frac{0.56}{\sqrt{n}} K_{4,n}^{3/4}$
- *Shevtsova (2011) i.i.d.:* $\frac{0.4748}{\sqrt{n}} K_{3,n} \leq \frac{0.4748}{\sqrt{n}} K_{4,n}^{3/4}$
- *Theorem 2.3 i.n.i.d.:* $\frac{0.441}{\sqrt{n}} K_{3,n} + O(n^{-1}) \leq \frac{0.441}{\sqrt{n}} K_{4,n}^{3/4} + O(n^{-1})$
- *Theorem 2.3 i.n.i.d. (unskewed):* $\frac{0.399}{\sqrt{n}} K_{3,n} + O(n^{-1}) \leq \frac{0.399}{\sqrt{n}} K_{4,n}^{3/4} + O(n^{-1})$
- *Theorem 2.3 i.i.d.:* $\frac{0.242 K_{3,n} + 0.1995}{\sqrt{n}} + O(n^{-1}) \leq \frac{0.242 K_{4,n}^{3/4} + 0.1995}{\sqrt{n}} + O(n^{-1})$

- *Theorem 2.3 i.i.d. (unskewed):* $\frac{0.1995 K_{3,n} + 0.1995}{\sqrt{n}} + O(n^{-1}) \leq \frac{0.1995 K_{4,n}^{3/4} + 0.1995}{\sqrt{n}} + O(n^{-1})$

As previously mentioned, our bound in the baseline i.n.i.d. case gets close to the best known Berry-Esseen bound in the i.i.d. setup (Shevtsova, 2011) when n is larger than $10,000$. When $\lambda_{3,n} = 0$, our bounds are smaller, notably in the i.i.d. scenario: in the latter case, the bound is smaller than 0.05 for $n \approx 1,000$, highlighting that taking (lack of) skewness into account matters in improving Berry-Esseen bounds. More generally, the results are considerably better in the i.i.d. framework.

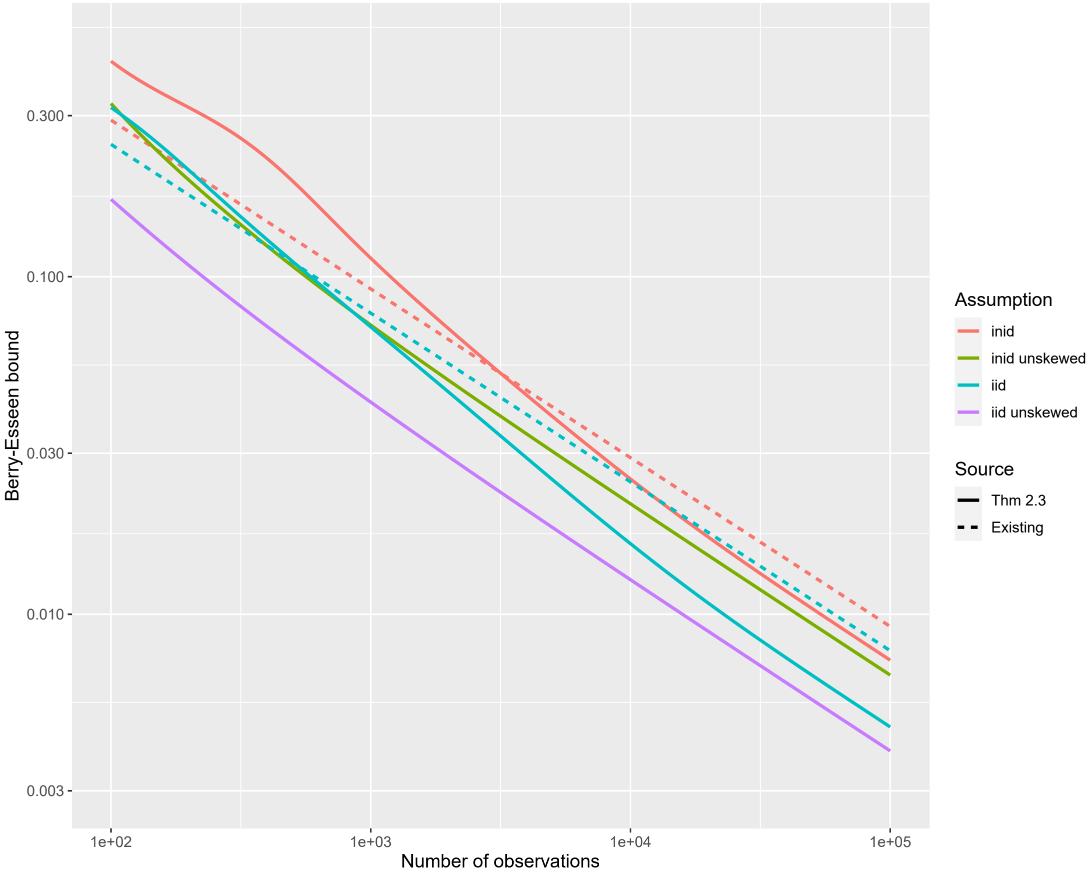

Figure 1: Comparison between existing and new Berry-Esseen bounds on $\Delta_{n,\mathrm{B}} := \sup_{x \in \mathbb{R}} |\mathbb{P}(S_n \leq x) - \Phi(x)|$ for different sample sizes under moment conditions only (log-log scale).

3 Improved bounds on $\Delta_{n,\mathrm{E}}$ under assumptions on the tail behavior of f_{S_n}

In this section, we derive tighter bounds on $\Delta_{n,\mathrm{E}}$ under additional regularity conditions on the tail behavior of the characteristic function of S_n. They follow from Theorem 3.1, which provides an alternative upper bound on $\Delta_{n,\mathrm{E}}$ that involves the tail behavior of f_{S_n}. This theorem is proved in Section 5.5.

Theorem 3.1. *If Assumption 2.1 or Assumption 2.2 holds, we have the bound*

$$\Delta_{n,E} \leq \frac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2}{n} + \frac{1.0253}{\pi} \int_{a_n}^{b_n} \frac{|f_{S_n}(t)|}{t} dt + r_{2,n}\,, \tag{5}$$

where $a_n := 2t_1^* \pi \sqrt{n} / \widetilde{K}_{3,n}$, $b_n := 16\pi^4 n^2 / \widetilde{K}_{3,n}^4$, *and* $r_{2,n}$ *is a remainder term that depends only on* $K_{3,n}$, $K_{4,n}$ *and* $\lambda_{3,n}$. *The term* $r_{2,n}$ *is defined in Equation* (34) *under Assumption 2.1 (resp. in Equation* (39) *under Assumption 2.2).*

When the fourth moment is bounded, i.e. $K_{4,n} = O(1)$, *we have* $r_{2,n} = O(n^{-5/4})$.

If furthermore $\mathbb{E}[X_i^3] = 0$ *for every* $i = 1, \dots, n$, *then*

i. the average third raw moment is null, i.e. $\lambda_{3,n} = 0$,

ii. the first-order term in the Edgeworth expansion is null too so that $\Delta_{n,B} = \Delta_{n,E}$,

iii. the residual term $r_{2,n}$ *converges to* 0 *at the faster rate* $n^{-3/2}$ *under Assumption 2.1 (resp. at the rate* n^{-2} *under Assumption 2.2).*

This theorem is obtained under the same conditions as Theorem 2.3. The first term contains quantities that were already present in the term of order $1/n$ in the bound of Theorem 2.3: $0.195 K_{4,n}$ and $0.038 \lambda_{3,n}^2$. On the contrary, the other terms are encompassed in the integral term and in the remainder. Indeed, a careful reading of the proofs shows that the leading term $0.1995\, \widetilde{K}_{3,n}/\sqrt{n}$ in the bound (3) comes from choosing a free tuning parameter T of the order of $\sqrt{n}$. Here, we make another choice for T such that this term is now negligible. The cost of this change of T is the introduction of the integral term involving f_{S_n}. The leading term of the bound thus depends on the tail behavior of f_{S_n}.

The next corollaries specify the magnitude of the bound according to the assumed regularity conditions on f_{S_n}. The first one considers a polynomial decrease in the *i.n.i.d.* framework.

Corollary 3.2. *If Assumption 2.1 holds and for all* $t \in (a_n, +\infty)$, $|f_{S_n}(t)| \leq C_0 t^{-p}$ *for some constants* $C_0, p > 0$, *then*

$$\Delta_{n,E} \leq \frac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2}{n} + \frac{1.0253\, C_0 a_n^{-p}}{\pi} + r_{3,n}$$

where $r_{3,n} := r_{2,n} - 1.0253\, C_0 b_n^{-p}/\pi$, *and* $a_n := 2t_1^* \pi \sqrt{n} / \widetilde{K}_{3,n}$.

Besides moment conditions, Corollary 3.2 requires a uniform control on the tail of f_{S_n} beyond the point $2t_1^* \pi \sqrt{n} / \widetilde{K}_{3,n}$. Whenever $\widetilde{K}_{3,n} = o(\sqrt{n})$, this condition is a tail control of the characteristic function of S_n above a point that tends to infinity, thus making this condition weaker to impose. When $p > 1$, f_{S_n} is absolutely integrable and thus P_{S_n} is a continuous distribution (Ushakov, 2011, Theorem 1.2.6). In this case, we give a characterization of the tail constraint on f_{S_n} in terms of smoothness of the underlying distribution function in Proposition 3.3.

Although Corollary 3.2 is valid for every positive p, it is only an improvement on the results of the previous section under the stricter condition $p > 1$. In particular when $p = 2$, a_n^{-p} is exactly of the order n^{-1}, we obtain

$$\Delta_{n,\mathrm{E}} \leq \frac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2 + 1.0253\, C_0 \pi^{-1}}{n} + o\big(n^{-1}\big),$$

and when $p > 2$

$$\Delta_{n,\mathrm{E}} \leq \frac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2}{n} + o\big(n^{-1}\big).$$

Combining these bounds on $\Delta_{n,\mathrm{E}}$ with the expression of the Edgeworth expansion translates into upper bounds on $\Delta_{n,\mathrm{B}}$ of the form

$$\Delta_{n,\mathrm{B}} \leq \frac{0.067\, |\lambda_{3,n}|}{\sqrt{n}} + O(n^{-1}) \leq \frac{0.042\, K_{4,n}^{3/4}}{\sqrt{n}} + O(n^{-1}).$$

In the regime when the $O(n^{-1})$ term is smaller than $0.042 K_{4,n}^{3/4}/\sqrt{n}$, the bound on $\Delta_{n,\mathrm{B}}$ becomes much better than $0.56 K_{3,n}/\sqrt{n}$ or $0.4748 K_{3,n}/\sqrt{n}$. This can happen even for sample sizes n of the order of a few thousands, assuming that $K_{3,n}$ and $K_{4,n}$ are reasonable. When $\mathbb{E}[X_i^3] = 0$ for every $i = 1, \dots, n$, we remark that $\Delta_{n,\mathrm{B}} = \Delta_{n,\mathrm{E}}$, meaning that we obtain a bound on $\Delta_{n,\mathrm{B}}$ of order n^{-1}.

We verify these rates through a numerical application in Example 3.5 for the specific choice $C_0 = 1$ and $p = 2$. These choices are satisfied for usual distributions such as the Laplace distribution (for which these values of C_0 and p are sharp) and for the Gaussian distribution.

More generally, a bound on the tail of the characteristic function is nearly equivalent to a regularity condition on the density. We detail this in the following proposition. The first part of this proposition is taken from (Ushakov, 2011, Theorem 2.5.4) (see also Ushakov and Ushakov (1999)). The second part is proved in Section 6.5.

Proposition 3.3. *Let $p \geq 1$ be an integer and Q be a probability measure that admits a density q with respect to Lebesgue's measure and f_Q its corresponding characteristic function.*

1. *If q is $(p-1)$ times differentiable and $q^{(p-1)}$ is a function with bounded variation, then*

$$|f_Q(t)| \leq \frac{\mathrm{Vari}[q^{(p-1)}]}{|t|^p},$$

   where $\mathrm{Vari}[\psi]$ *denotes the total variation of a function ψ.*

2. *If $t \mapsto |t|^{p-1} |f_Q(t)|$ is integrable on a neighborhood of $+\infty$, then q is $(p-1)$ times differentiable.*

It is sufficient that there exists $C > 0$ and $\beta > 1$ such that $|f_Q(t)| \leq C / \big(|t|^p \log(|t|)^\beta\big)$ to satisfy the integrability condition in the second part of Proposition 3.3. Proposition 3.3 shows that the tail condition on f_{S_n} in Corollary 3.2 is satisfied if P_{S_n} has a density g_{S_n} with respect to Lebesgue's measure that is $p-1$ times differentiable and such that its $(p-1)$-th derivative is of bounded variation with total variation $V_n := \mathrm{Vari}[g_{S_n}^{(p-1)}]$ uniformly bounded in n. In such cases, we can take $C_0 = 1 \wedge \sup_{n \in \mathbb{N}^*} V_n$.

Another possibility would be to impose $|f_{S_n}(t)| \leq \max_{1 \leq r \leq M} |\rho_r(t)|$ for every $|t| \geq a_n$ and for $(\rho_r)_{r=1,\dots,M}$ a family of known characteristic functions. Indeed, in a statistical framework, the characteristic function f_{S_n} is unknown, and this second suggestion boils down to a semiparametric assumption on P_{S_n}: f_{S_n} is assumed to be controlled in a neighborhood of $+\infty$ by the behavior of at least one of the M characteristic functions $(\rho_r)_{r=1,\dots,M}$, but f_{S_n} needs not be exactly one of those M characteristic functions. This semiparametric restriction becomes less and less stringent as n increases since we need to control f_{S_n} on a region that vanishes as n goes to infinity. Since

S_n is centered and of variance 1 by definition, the choice of possible ρ_r is naturally restricted to the set of characteristic functions that correspond to such standardized distributions.

We state a second corollary which deals with the *i.i.d.* framework. We need to define the following quantity $\kappa_n(\widetilde{K}_{3,n}) := \sup_{t:\, |t| \geq 2t_1^* \pi / \widetilde{K}_{3,n}} |f_{X_n/\sigma_n}(t)|$. Under Assumption 2.2, we remark that $\sup_{t:\, |t| \geq a_n} |f_{S_n}(t)| = \kappa_n(\widetilde{K}_{3,n})^n$.

Corollary 3.4. *If Assumption 2.2 holds and P_{X_n/σ_n} has an absolutely continuous component with respect to Lebesgue's measure, then $\kappa_n(\widetilde{K}_{3,n}) < 1$ and*

$$\Delta_{n,E} \leq \frac{0.195\, K_{4,n} + 0.038\, \lambda_{3,n}^2}{n} + \frac{1.0253\, \kappa_n(\widetilde{K}_{3,n})^n \log(c_n)}{\pi} + r_{2,n}\, ,$$

where $c_n := b_n/a_n = 8\pi^3 n^{3/2} / (t_1^* \widetilde{K}_{3,n}^3)$.

Note that for a given $s > 0$ and a variable Z, $\sup_{t:|t| \geq s} |f_Z(t)| = 1$ if and only if P_Z is a lattice distribution, i.e. concentrated on a set of the form $\{a + nh, n \in \mathbb{Z}\}$ (Ushakov, 2011, Theorem 1.1.3). Therefore, $\kappa_n(\widetilde{K}_{3,n}) < 1$ as soon as the distribution is not lattice, which is the case for any distribution with an absolute continuous component.

In Corollary 3.4, we derive an upper bound of order n^{-1} on $\Delta_{n,\mathrm{E}}$. As the first term on the right-hand side is independent of the behavior of the characteristic function, it is unchanged compared to Theorem 3.1 and Corollary 3.2. The second term in the bound, $(1.0253/\pi)\kappa_n(\widetilde{K}_{3,n})^n \log(c_n)$, corresponds to an upper bound on the integral term in Theorem 3.1. As a consequence, in the *i.i.d.* case with a distribution independent of n, $\kappa_n(\widetilde{K}_{3,n})$ is constant and this bound is exponentially better than the one derived in Corollary 3.2 and is asymptotically negligible in front of the first term of order $O(n^{-1})$.

In applications, we may want to compute explicitly this bound, and therefore we would need an explicit value for $\kappa_n(\widetilde{K}_{3,n})$. This value depends on the (unknown) law P_{X_n/σ_n} and if it is too close to one for a given n, $\kappa_n(\widetilde{K}_{3,n})^n$ will not be small. For instance, if $n = 1,000$ and $\kappa_n(\widetilde{K}_{3,n}) = 0.999$, we have $\kappa_n(\widetilde{K}_{3,n})^n \approx 0.37$. As a result, we must impose that $\kappa_n(\widetilde{K}_{3,n})$ be bounded away from one in order to use such a bound in practice, which amounts to placing a restriction on P_{X_n/σ_n}.

This is in fact more than a restriction on the tail of f_{X_n/σ_n} since a control of the tail of a characteristic function induces a control on its central part as well (Ushakov, 2011, Theorem 1.4.4). By convexity $\widetilde{K}_{3,n} \geq 1$ and in the most favorable case, we need a control of f_{X_n/σ_n} above the point $2t_1^*\pi$. As in the *i.n.i.d.* case, a possibility to do so is to impose that the characteristic function f_{X_n/σ_n} is controlled by some known family of characteristic functions $\rho_1, \dots, \rho_M$ beyond $2t_1^*\pi/\widetilde{K}_{3,n}$.

In Example 3.5, we derive a bound on $\Delta_{n,\mathrm{B}}$ from Corollary 3.4 which has the same flavor as the one obtained from Corollary 3.2. Similarly to what has been done previously, we derive a better result under the assumption that $\mathbb{E}[X_1^3] = 0$.

Example 3.5 (Implementation of our bounds on $\Delta_{n,\mathrm{B}}$)**.** *We compare the bounds on $\Delta_{n,B}$ obtained in Corollaries 3.2 and 3.4 to $0.56K_{3,n}/\sqrt{n}$ and $0.4748K_{3,n}/\sqrt{n}$. As in previous examples, we impose $K_{4,n} \leq 9$.*

- *Corollary 3.2 i.n.i.d., $p = 2$ and $C_0 = 1$:* $\Delta_{n,B} \leq \frac{0.042\, K_{4,n}^{3/4}}{\sqrt{n}} + \frac{0.195\, K_{4,n} + 0.024\, K_{4,n}^{3/2}}{n} + 1.0253\, \pi^{-1} a_n^{-2} + o\big(n^{-1}\big)$

- *Corollary 3.2 i.n.i.d. unskewed, $p = 2$ and $C_0 = 1$:* $\Delta_{n,B} \leq \frac{0.195\, K_{4,n}}{n} + 1.0253\, \pi^{-1} a_n^{-2} + o\big(n^{-1}\big)$

- *Corollary 3.4 i.i.d., $\kappa = 0.99$:* $\Delta_{n,B} \leq \frac{0.042\, K_{4,n}^{3/4}}{\sqrt{n}} + \frac{0.195\, K_{4,n} + 0.024\, K_{4,n}^{3/2}}{n} + \frac{1.0253\, \kappa^n \log(c_n)}{\pi} + r_{2,n}$,

- *Corollary 3.4 i.i.d. unskewed, $\kappa = 0.99$:* $\Delta_{n,B} \leq \frac{0.195\, K_{4,n}}{n} + \frac{1.0253\, \kappa^n \log(c_n)}{\pi} + r_{2,n}$,

As we assume some supplementary regularity conditions in this section, it should not be surprising that they leads to an improvement of the bounds. In Figure 2, we displayed the different bounds that we obtained, as a function of the sample size n, alongside with the existing bounds (that do not assume such regularity conditions). The new bounds take advantage of these regularity conditions and are therefore much tighter, especially in the unskewed case where the rate of convergence gets faster from $1/\sqrt{n}$ to $1/n$.

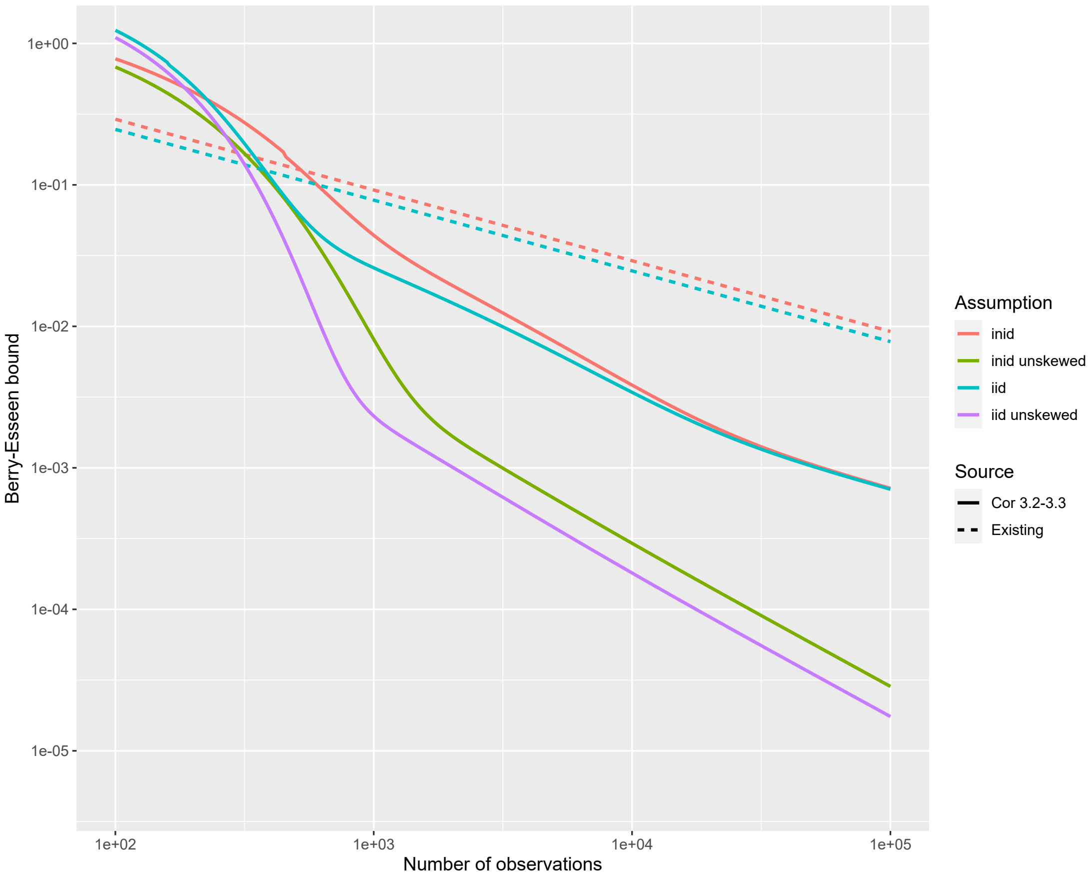

Figure 2: Comparison between existing and new Berry-Esseen bounds on $\Delta_{n,\mathrm{B}} := \sup_{x \in \mathbb{R}} |\mathbb{P}\left(S_n \leq x\right) - \Phi(x)|$ for different sample sizes under moment conditions and tail control of the characteristic function f_{S_n} (log-log scale).

4 Conclusion and statistical applications

In addition to Figures 1 and 2, we can illustrate the improvements provided by Theorem 2.3 and Corollary 3.4 in the *i.i.d.* case by resuming Example 1.1. This example considers the classical problem of conducting inference on a univariate expectation from *i.i.d.* observations. For usual confidence levels, Table 2 indicates the largest weakly informative sample size $n_{\max}(\alpha)$ (as defined in Example 1.1) under different assumptions. For these numerical applications, we keep the same constants as in the previous figures, i.e. $K_{4,n} = 9$ and $\kappa = 0.99$.

Sources	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
Example 1.1	608	2,434	60,867
Thm. 2.3	587	1,725	22,850
Thm. 2.3 unskewed	226	745	15,865
Cor. 3.4	417	575	3,134
Cor. 3.4 unskewed	351	427	624

Table 2: Largest weakly informative $n_{\max}(\alpha)$ from different sources and different levels α.

As seen in the previous examples, explicit values or bounds on some functionals of P_{S_n} – such as $\lambda_{3,n}$, $\widetilde{K}_{3,n}$, $K_{4,n}$, C_0, p, and κ – are required to obtain our non-asymptotic bounds on a standardized sample mean. The bound $K_{4,n} \leq 9$ encompasses a wide range of standard distributions and provides a bound on $\lambda_{3,n}$ and $\widetilde{K}_{3,n}$ too. Nonetheless, we may prefer to avoid such a restriction, either because it is unlikely to be satisfied, or, on the contrary, it might be overly conservative. An alternative would be to estimate the moments $\lambda_{3,n}$, $K_{3,n}$ (then $\widetilde{K}_{3,n} \leq 2K_{3,n}$), $K_{4,n}$ using the data. In the *i.i.d.* case, we suggest to estimate them by their empirical counterparts (via the so-called method of moments). We could then compute our bounds replacing the unknown needed quantities by their estimates. We acknowledge that this type of "plug-in" approach is only approximately valid.

Theorem 3.1, which underlies Corollaries 3.2 and 3.4, involves the integral $\int_{a_n}^{b_n} f_{S_n}(t)/t \, dt$, which depends on the a priori unknown the characteristic function of S_n. A nonparametric possibility would be to estimate f_{S_n}, and then to numerically integrate it. A classical estimator is the empirical characteristic function, see for instance (Ushakov, 2011, Chapter 3). However, this nonparametric estimator may be unstable since the uniform convergence of the empirical characteristic function only holds on compact sets. Therefore, it might be of interest to avoid such an estimation. Corollaries 3.2 and 3.4 offer this alternative under some regularity conditions on f_{S_n}. They require an upper control on the tail of the characteristic function. Such a bound can be given using expert knowledge of the regularity of the density of the S_n. The plot of the module of the empirical characteristic function can be used to back such an intuition.

We now examine some implications of our theoretical results for the non-asymptotic validity of statistical tests based on the Gaussian approximation of the distribution of a sample mean using *i.i.d.* data. As an example, we study the unilateral test of $H_0 : \mathbb{E}[X] \leq 0$ against $H_1 : \mathbb{E}[X] > 0$ with known variance for a confidence level $1 - \alpha$. Starting from Theorem 2.3 or Theorem 3.1, we set $x = q_{1-\alpha} := \Phi^{-1}(1 - \alpha)$ and deduce the following upper and lower bounds

$$\frac{\lambda_{3,n}}{6\sqrt{n}}(1 - q_{1-\alpha}^2)\varphi(q_{1-\alpha}) - \delta_n \leq \mathbb{P}\left(S_n \leq q_{1-\alpha}\right) - (1 - \alpha) \leq \frac{\lambda_{3,n}}{6\sqrt{n}}(1 - q_{1-\alpha}^2)\varphi(q_{1-\alpha}) + \delta_n, \quad (6)$$

where δ_n is the corresponding bound on $\Delta_{n,\mathrm{E}}$. Recall that either $\delta_n = O(n^{-1/2})$ under moments conditions only or $\delta_n = O(n^{-1})$ under supplementary conditions on the tail behavior of the characteristic function.

For the sake of conciseness, we consider henceforth the latter framework where the distribution of the variable has a continuously differentiable density. In that smooth case, the term $0.067\lambda_{3,n}/\sqrt{n}$ becomes dominant in the bound on $\Delta_{n,\mathrm{E}}$ (Corollary 3.2) and we can write up to a $O(n^{-1})$ term δ_n,

$$\mathbb{P}\left(S_n \le q_{1-\alpha}\right) \approx 1 - \alpha + \frac{\lambda_{3,n}}{6\sqrt{n}}(1 - q_{1-\alpha}^2)\varphi(q_{1-\alpha}). \tag{7}$$

As a consequence, as soon as the skewness $\lambda_{3,n}$ is higher than $\dfrac{6\sqrt{n}\delta_n}{(q_{1-\alpha}^2 - 1)\varphi(q_{1-\alpha})} > 0$ for $\alpha \le 0.15$ (so that $q_{1-\alpha} > 1$), the probability $\mathbb{P}\left(S_n \le q_{1-\alpha}\right)$ has to be smaller than $1 - \alpha$. In that case, the test does not reach its stated level and is said liberal. Conversely, when $\lambda_{3,n}$ is lower than $\dfrac{-6\sqrt{n}\delta_n}{(q_{1-\alpha}^2 - 1)\varphi(q_{1-\alpha})} < 0$, the probability $\mathbb{P}\left(S_n \le q_{1-\alpha}\right)$ has to be larger than $1 - \alpha$. In that case, the test is said conservative.

The distortion can also be seen in terms of p-values. In the unilateral test we consider, the p-value is $pval := \mathbb{P}_0(S_n > s_n)$ with $\mathbb{P}_0$ the probability distribution when $\mathbb{E}[X] = 0$ and s_n the observed value of S_n in the sample. In contrast, the approximated p-value is $\widetilde{pval} := 1 - \Phi(s_n)$. Analogous to (6), setting $x = s_n$ yields

$$\frac{\lambda_{3,n}}{6\sqrt{n}}(1 - s_n^2)\varphi(s_n) - \delta_n \le (1 - pval) - (1 - \widetilde{pval}) \le \frac{\lambda_{3,n}}{6\sqrt{n}}(1 - s_n^2)\varphi(s_n) + \delta_n.$$

Therefore,

$$pval - \frac{\lambda_{3,n}}{6\sqrt{n}}(s_n^2 - 1)\varphi(s_n) - \delta_n \le \widetilde{pval} \le pval - \frac{\lambda_{3,n}}{6\sqrt{n}}(s_n^2 - 1)\varphi(s_n) + \delta_n.$$

This bound indicates that for $|s_n| > 1$, the higher the skewness, the smaller the approximated p-value compared to the true one. This results in overconfidence in rejecting the null. The distortion towards lower approximated p-values occurs in a similar way whether $s_n < -1$ or $s_n > 1$ due to parity of the function $x \mapsto (x^2 - 1)\varphi(x)$. Conversely, negative skewness leads to under-rejection of the null hypothesis (Table 3).

Skewness	$s_n < -1$	$-1 < s_n < 1$	$s_n > 1$
Negative	$+$	$-$	$+$
Positive	$-$	$+$	$-$

Table 3: Sign of the distortion $\widetilde{pval} - pval$ in the continuous case.

More generally, for any n and point x, Equation (6) shows that $\mathbb{P}(S_n \le x)$ belongs to the interval $\left[\Phi(x) + \lambda_{3,n}(1 - x^2)\varphi(x)/(6\sqrt{n}) \pm \delta_n\right]$, which is not centered at $\Phi(x)$ whenever $\lambda_{3,n} \neq 0$. Figure 4 illustrates this distortion in terms of p-values for different sample sizes. The length of the interval does not depend on x and shrinks at speed δ_n. On the contrary, the location of the interval depends on x: it is all the more shifted away from the asymptotic approximation $\Phi(x)$ as $x \mapsto (1 - x^2)\varphi(x)$ is large in absolute value. That function has a global maximum at $x = 0$ and global minimum at the points $-x^*$ and $x^* \approx 1.732$. This implies that irrespective of n, the largest gaps between $\mathbb{P}(S_n \le x)$ and $\Phi(x)$ may be expected around $x = 0$ or $x = \pm x^*$. When $x = \pm 1$, the interval is exactly centered at $\Phi(x)$.

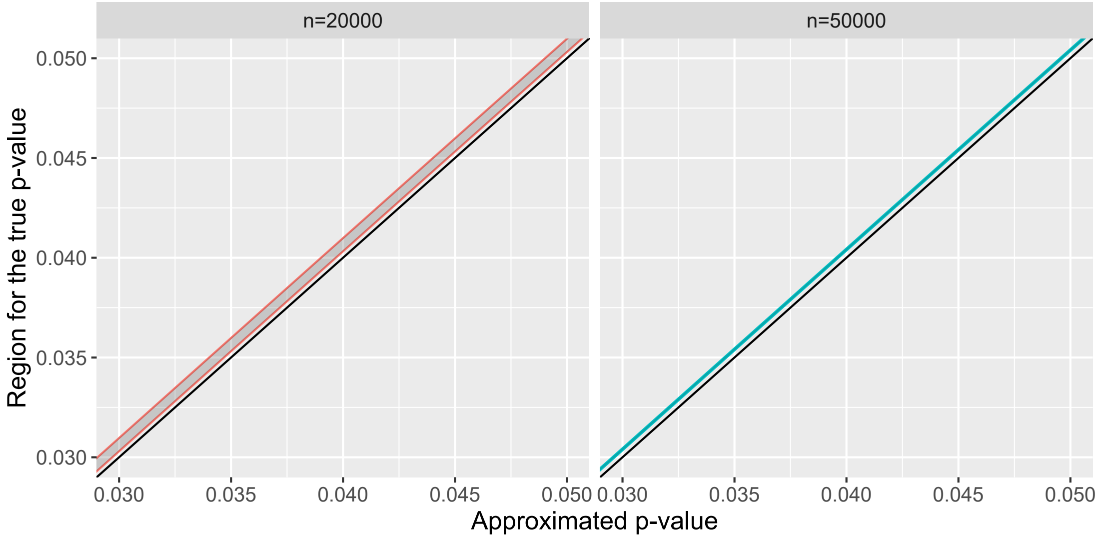

Figure 3: For every test statistic $s \in \mathbb{R}$, we represent possible pairs $(\widetilde{pval}(s), pval(s))$ (shaded region) of the true p-value $pval(s) := 1 - \mathbb{P}(S_n \leq s)$ and the approximated one $\widetilde{pval}(s) := 1 - \Phi(s)$. The boundaries of this region correspond to the curves $(1-\Phi(s), 1-\Phi(s) - \frac{\lambda_{3,n}}{6\sqrt{n}}(s^2-1)\varphi(s) \pm \delta_n)_{s \in \mathbb{R}}$ given by Corollary 3.4 (with $p = 4$) for two sample sizes. Here, we have chosen $K_{4,n} = 9$ and $\lambda_{3,n} = 0.6 \times 9^{3/4} \approx 3.11$. The black line corresponds to the first diagonal $(1 - \Phi(s), 1 - \Phi(s))$.

Example: for $n = 20000$ and an approximated p-value of 0.04, the true p-value must be in the interval $[0.0403, 0.0410]$.

5 Proof of the main results

In this section, we prove Theorems 2.3 and 3.1. All the proofs start with a so-called "smoothing inequality" which we present and prove in Subsection 5.1. The remaining subsections are devoted to proving the theorems themselves.

5.1 A smoothing inequality

The result given in Lemma 5.1 helps control the distance between the cumulative distribution function $\mathbb{P}(S_n \leq x)$ and $G_n(x) := \Phi(x) + \frac{\lambda_{3,n}}{6\sqrt{n}}(1 - x^2)\varphi(x)$ in terms of their respective Fourier transforms.

Lemma 5.1. *For every $t_0 \in (0, 1]$ and every $T > 0$, we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}(S_n \leq x) - G_n(x) \right| \leq \Omega_1(t_0, T, |\lambda_{3,n}|/\sqrt{n}) + \Omega_2(t_0, T) + \Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}), \tag{8}$$

where

$$\Omega_1(t_0, T, v) := 2 \int_0^{t_0} \left| \Psi(t) - \frac{i}{2\pi t} \right| e^{-(Tt)^2/2} \left(1 + \frac{v|Tt|^3}{6} \right) dt + \frac{1}{\pi} \int_{t_0}^{+\infty} \frac{e^{-(Tt)^2/2}}{t} \left(1 + \frac{v|Tt|^3}{6} \right) dt,$$

$$\Omega_2(t_0, T) \quad := 2 \int_{t_0}^{1} |\Psi(t)| \, |f_{S_n}(Tt)| dt,$$

$$\Omega_3(t_0, T, v) := 2 \int_0^{t_0} |\Psi(t)| \left| f_{S_n}(Tt) - e^{-(Tt)^2/2} \left(1 - \frac{vi(Tt)^3}{6} \right) \right| dt,$$

and $\Psi(t) := \frac{1}{2}\left(1 - |t| + i\left[(1-|t|)\cot(\pi t) + \frac{\text{sign}(t)}{\pi}\right]\right) \mathbb{1}\{|t| \leq 1\}$.

In the following, we resort to Equations (I.29) and (I.30) of Prawitz (1975) which claim that the function Ψ satisfies

$$\big|\Psi(t)\big| \leq \frac{1.0253}{2\pi|t|} \text{ and } \left|\Psi(t) - \frac{i}{2\pi t}\right| \leq \frac{1}{2}\left(1 - |t| + \frac{\pi^2}{18}t^2\right). \tag{9}$$

Proof of Lemma 5.1. Let us denote by "p.v. $\int$ " Cauchy's principal value, defined by

$$\text{p.v.} \int_{-a}^{a} f(u)du := \lim_{x \to 0,\, x>0} \int_{-a}^{-x} f(u)du + \int_{x}^{a} f(u)du,$$

where f is a measurable function on $[-a, a]\backslash\{0\}$ for a given $a > 0$. In the following, we use the following inequalities, which are due to Prawitz (1972), where F is the cumulative distribution function of S_n and $f = f_{S_n}$ its characteristic function,

$$\lim_{y \to x,\, y>x} F(y) \leq \frac{1}{2} + \text{p.v.} \int_{-U}^{+U} e^{-ixu} \frac{1}{U} \Psi\left(\frac{u}{U}\right) f(u)du,$$
$$\lim_{y \to x,\, y<x} F(y) \geq \frac{1}{2} + \text{p.v.} \int_{-U}^{+U} e^{-ixu} \frac{1}{U} \Psi\left(\frac{-u}{U}\right) f(u)du.$$

Therefore,

$$F(x) - G_n(x) \leq \frac{1}{2} + \text{p.v.} \int_{-U}^{+U} e^{-ixu} \frac{1}{U} \Psi\left(\frac{u}{U}\right) f(u)du - G_n(x) \tag{10}$$
$$F(x) - G_n(x) \geq \frac{1}{2} + \text{p.v.} \int_{-U}^{+U} e^{-ixu} \frac{1}{U} \Psi\left(\frac{-u}{U}\right) f(u)du - G_n(x). \tag{11}$$

Note that the Gil-Pelaez inversion formula (see Gil-Pelaez (1951)) is valid for any bounded-variation function. Formally, for every bounded-variation function $G(x) = \int_{-\infty}^{x} g(t)dt$, denoting the Fourier transform of a given function g by $\check{g} := \int_{-\infty}^{+\infty} e^{ixu} g(u)du$, we have

$$G(x) = \frac{1}{2} + \frac{i}{2\pi} \text{p.v.} \int_{-\infty}^{+\infty} e^{-ixu} \check{g}(u)du. \tag{12}$$

Therefore, applying Equation (12) to the function $G_n(x) := \Phi(x) + \lambda_{3,n}(1-x^2)\varphi(x)/(6\sqrt{n})$ whose (generalized) density has the Fourier transform $(1 - \lambda_{3,n} ix^3/(6\sqrt{n}))e^{-x^2/2}$, we get

$$G_n(x) = \frac{1}{2} + \frac{i}{2\pi} \text{p.v.} \int_{-\infty}^{+\infty} e^{-ixu} \left(1 - \frac{\lambda_{3,n}}{6\sqrt{n}} iu^3\right) e^{-u^2/2} \frac{du}{u}.$$

Combining this equality with the bounds (10) and (11) and using the triangular inequality, we get the claimed result (8).

□

5.2 Proof of Theorem 2.3 under Assumption 2.1

In this section, we state and prove a more general theorem (Theorem 5.4 below). We recover Theorem 2.3 when we set $\varepsilon = 0.1$ (in Supplement A, we compute $e_1(0.1)$ and obtain the upper bound $e_1(0.1) \leq 1.012$).

Theorem 5.2 (One-term Edgeworth expansion under Assumption 2.1)**.** *Under Assumption 2.1, for every $\varepsilon \in (0, 1/3)$ and every $n \geq 1$, we have the bound*

$$\Delta_{n,E} \leq \frac{0.1995\, \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{1}{n}\Bigg\{ 0.031\, \widetilde{K}_{3,n}^2 + 0.327\, K_{4,n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + 0.054\, |\lambda_{3,n}| \widetilde{K}_{3,n} + 0.037\, e_1(\varepsilon) |\lambda_{3,n}|^2 \Bigg\} + r_{1,n}, \qquad (13)$$

where $e_1(\varepsilon)$ is given in Equation (41) *and $r_{1,n}$ is given in Equation* (22)*. Note that $r_{1,n} = O(n^{-5/4})$ as soon as $K_{4,n} = O(1)$. If $\mathbb{E}[X_i^3] = 0$ for every $i = 1, \dots, n$, the upper bound reduces to*

$$\frac{0.1995 \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{1}{n}\left\{ 0.031 \widetilde{K}_{3,n}^2 + 0.327 K_{4,n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right)\right\} + r_{1,n}, \qquad (14)$$

with $r_{1,n} = O(n^{-3/2})$ when $K_{4,n} = O(1)$.

We follow the proof strategy initiated by Prawitz (1975) and complemented among others by Shevtsova (2012). The method of proof starts from Lemma 5.1. Note that in Prawitz (1975) or Shevtsova (2012), Lemma 5.1 is used with $\Phi(x)$ instead of $G_n(x)$ (i.e. these authors are only interested in the canonical Berry-Esseen inequality). It is shown in Section 6.1 how to upper bound the term $\Omega_1(t_0, T, |\lambda_{3,n}|/\sqrt{n})$. There remains to control $\Omega_2(t_0, T)$ and $\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n})$. The former is handled as in Shevtsova (2012) while the latter is tackled differently.

As in the classical Berry-Esseen setup, T is chosen of the order of $\sqrt{n}$. The parameter t_0 is chosen as in Prawitz (1975): a range of values is admissible, in particular t_0 must satisfy $1/(2\pi) \leq t_0 \leq 1/(6\pi\chi_1)$ where $\chi_1 := \sup_{x>0} x^{-3}|\cos(x) - 1 + x^2/2| \approx 0.099162$. Given the numerical evidence in Prawitz (1975), the choice $t_0 = 1/\pi$ is reasonable.

Let $\xi_n := \widetilde{K}_{3,n}/\sqrt{n}$. Based on the upper bound on Ω_1, we would like to pick T as large as possible. However, Lemma 5.3 below restricts the range of informative Ts to control Ω_2. More specifically, we show in Section 5.2.1 that a suitable choice for T is

$$T = \frac{2\pi}{\xi_n} = \frac{2\pi}{\widetilde{K}_{3,n}/\sqrt{n}} = \frac{2\pi\sqrt{n}}{\widetilde{K}_{3,n}}.$$

5.2.1 Bound on Ω_2

In this section, we control $\Omega_2(1/\pi, T) = 2\int_{1/\pi}^1 |\Psi(t)|\, |f_{S_n}(Tt)| dt$. To reach this goal, we use the following lemma which is a consequence of Lemma 2 in Shevtsova (2010) or Theorem 2.2 in Shevtsova (2012) (with $\delta = 1$):

Lemma 5.3. *Let $t_1^* = \theta_1^*/(2\pi)$ where θ_1^* is the unique root in $(0, 2\pi)$ of the equation $\theta^2 + 2\theta\sin(\theta) + 6(\cos(\theta) - 1) = 0$.*[2] *For T such that $t_1^*/\xi_n \leq Tt \leq 2\pi/\xi_n$ for every $t \in [t_1^*, 1]$, we have*

$$|f_{S_n}(Tt)| \leq e^{-(Tt)^2/2 + \chi_1 \xi_n |Tt|^3} \text{ for } t \in [1/\pi\,,\, t_1^*]$$
$$|f_{S_n}(Tt)| \leq e^{-(1-\cos(\xi_n Tt))/\xi_n^2} \text{ for } t \in [t_1^*\,,\, 1].$$

Therefore,

$$\int_{1/\pi}^1 |\Psi(t)|\, |f_{S_n}(Tt)| dt \leq \int_{1/\pi}^{t_1^*} |\Psi(t)| e^{-(Tt)^2/2 + \chi_1\xi_n|Tt|^3} dt + \int_{t_1^*}^1 |\Psi(t)| e^{-(1-\cos(\xi_n Tt))/\xi_n^2} dt.$$

[2] In Supplement A, we compute the approximation $t_1^* \approx 0.635966$.

Proof of Lemma 5.3: Applying Theorem 2.2 in Shevtsova (2012) with $\delta = 1$, we get for all $u \in \mathbb{R}$

$$|f_{S_n}(u)| \leq \exp\left(-\psi(u, \epsilon_n)\right),$$

where $\epsilon_n := n^{-1/2}\widetilde{K}_{3,n}$, and, for any real u, $\epsilon > 0$

$$\psi(u, \epsilon) := \begin{cases} t^2 - \chi_1 \epsilon |t|^3, & \text{for } |t| < \theta_1^* \epsilon^{-1}, \\ \dfrac{1 - \cos(\epsilon t))}{\epsilon^2}, & \text{for } \theta_1^* \epsilon^{-1} \leq |t| \leq 2\pi\epsilon^{-1}, \\ 0, & \text{for } |t| > 2\pi\epsilon^{-1}. \end{cases}$$

Multiplying by $|\Psi|$, integrating from $1/\pi$ to 1 and separating the two cases yields the claimed inequality. □

As already mentioned, we want to take T as large as possible to control Ω_1. However, the proof of Lemma 5.3 implies that choosing $T > 2\pi/\xi_n$ would resort in a useless bound on Ω_2. This is why we pick $T = 2\pi/\xi_n$. We can thus write $\chi_1 \xi_n |Tt|^3 = 2\pi\chi_1 (Tt)^2 |t|$ for $t \in [1/\pi, t_1^*]$ and $\cos(\xi_n Tt) = \cos(2\pi t)$ for $t \in [t_1^*, 1]$. Combining this with Lemma 5.3 yields

$$\begin{aligned} \int_{1/\pi}^{1} |\Psi(t)|\, |f_{S_n}(Tt)| dt &\leq \int_{1/\pi}^{t_1^*} |\Psi(t)| e^{-\frac{(Tt)^2}{2}(1 - 4\pi\chi_1|t|)} dt + \int_{t_1^*}^{1} |\Psi(t)| e^{-T^2(1-\cos(2\pi t))/(4\pi^2)} dt \\ &= \frac{I_{2,1}(T)}{2T^4} + \frac{I_{2,2}(T)}{2T^2}, \end{aligned}$$

where

$$\begin{aligned} I_{2,1}(T) &:= T^4 \int_{1/\pi}^{t_1^*} 2|\Psi(t)| e^{-\frac{(Tt)^2}{2}(1-4\pi\chi_1|t|)} dt, \\ I_{2,2}(T) &:= T^2 \int_{t_1^*}^{1} 2|\Psi(t)| e^{-T^2(1-\cos(2\pi t))/(4\pi^2)} dt. \end{aligned}$$

Note that the difference in the two exponents of T in the above definitions may seem surprising as these two integrals look similar. However they have very different behaviors since the first one decays much faster than the second one. In line with Section 6.1, we compute numerically these integrals using the R package `cubature` (Narasimhan et al., 2020) and optimize them using the `optimize` function with the `L-BFGS-B` method (see the code in Supplement A). This gives

$$\sup_{T \geq 0} I_{2,1}(T) \leq 67.0415, \quad \text{and} \quad \sup_{T \geq 0} I_{2,2}(T) \leq 1.2187.$$

Finally, we arrive at

$$\Omega_2(1/\pi, T) = 2 \int_{1/\pi}^{1} |\Psi(t)|\, |f_{S_n}(Tt)| dt \leq \frac{67.0415}{T^4} + \frac{1.2187}{T^2}. \tag{15}$$

5.2.2 Bound on Ω_3

In this section, we bound the third term of Equation (8), which is

$$\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) = 2 \int_0^{t_0} |\Psi(t)| \left| f_{S_n}(Tt) - e^{-(Tt)^2/2} \left(1 - \frac{\lambda_{3,n} i (Tt)^3}{6\sqrt{n}}\right) \right| dt.$$

Whenever $t_0 T = T/\pi \geq \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$ – or equivalently when $n \geq \varepsilon^2 K_{3,n}^4 / (4K_{4,n})$ – we can write

$$\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \leq I_{3,1}(T) + I_{3,2}(T) + I_{3,3}(T),$$

where

$$I_{3,1}(T) := \frac{2}{T}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\left(1 - \frac{iu^3\lambda_{3,n}}{6\sqrt{n}}\right)\right| du,$$

$$I_{3,2}(T) := \frac{2}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{t_0 T} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\right| du,$$

$$I_{3,3}(T) := \frac{2}{T}\frac{|\lambda_{3,n}|}{6\sqrt{n}}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{t_0 T} |\Psi(u/T)|\, e^{-u^2/2}|u|^3 du.$$

Note that when $n < \varepsilon^2 K_{3,n}^4/(4K_{4,n})$, we have the better inequality $\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) < I_{3,1}(T)$, so that our reasoning is still valid even in this case as our bounds are all positive.

The integrand of $I_{3,1}(T)$ can be controlled with the help of Lemma 6.2, which enables us to write

$$I_{3,1}(T) \leq \frac{K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) J_1\left(4, 0, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T\right)$$
$$+ \frac{e_{1,n}(\varepsilon)}{36}\frac{|\lambda_{3,n}|^2}{n} J_1\left(6, 0, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T\right) + I_{3,1,3} \tag{16}$$

where

$$I_{3,1,3}(T) := \frac{2}{T}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} |\Psi(u/T)| e^{-u^2/2} R_{1,n}(u,\varepsilon) du,$$

and J_1 is defined in Equation (54). Using Equation (54), we obtain the bounds $J_1(4, 0, +\infty, T) \leq 0.327$ and $J_1(6, 0, +\infty, T) \leq 1.306$. Besides by the first inequality in (9), we get

$$I_{3,1,3}(T) \leq \frac{1.0253}{\pi}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} u e^{-u^2/2} R_{1,n}(u,\varepsilon) du.$$

We finally get from Equation (16)

$$I_{3,1}(T) \leq \frac{0.327 K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306 e_{1,n}(\varepsilon)}{36}\frac{|\lambda_{3,n}|^2}{n}$$
$$+ \frac{1.0253}{\pi}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} u e^{-u^2/2} R_{1,n}(u,\varepsilon) du. \tag{17}$$

We now handle $I_{3,3}(T)$. We remark that the following set of inequalities is valid for $p = 4$

$$\begin{aligned}
I_{3,3}(T) &= \frac{|\lambda_{3,n}|}{3\sqrt{n}} J_1\big(p, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, t_0 T, T\big) \\
&\leq \frac{1.0253}{2\pi}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{2\sqrt{n}/\widetilde{K}_{3,n}} u^{p-1} e^{-u^2/2} du \\
&= \frac{1.0253}{2\pi}\int_{\varepsilon(n/K_{4,n})^{1/2}}^{2n/\widetilde{K}_{3,n}^2} (2v)^{(p-1)/2} e^{-v}\frac{dv}{\sqrt{2v}} \\
&= O\Big(\Gamma\big(p/2, \varepsilon(n/K_{4,n})^{1/2}\big) - \Gamma\big(p/2, 2n/\widetilde{K}_{3,n}^2\big)\Big) \\
&= O\Big(n^{p/4-1/2} e^{-\varepsilon\sqrt{n}/\sqrt{K_{4,n}}}\Big),
\end{aligned}$$

where we apply the change of variable $v = u^2/2$, and take advantage of the asymptotic expansion $\Gamma(a,x) = x^{a-1}e^{-x}(1 + O((a-1)/x))$ which is valid for every fixed a in the regime $x \to \infty$, see Equation (6.5.32) in Abramowitz and Stegun (1972).

In the following section, we show that $I_{3,2}(T)$ decays exponentially with n too, so that we obtain

$$\begin{aligned}
\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) &\leq I_{3,1}(T) + I_{3,2}(T) + I_{3,3}(T) \\
&\leq \frac{0.327 K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306 e_{1,n}(\varepsilon)}{36}\frac{|\lambda_{3,n}|^2}{n} + O(n^{-5/4}). \qquad (18)
\end{aligned}$$

5.2.3 Bound on $I_{3,2}(T)$.

Let t be a real in the interval $[\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, t_0 T]$. As in the proof of Lemma 2.7 in Shevtsova (2012) with $\delta = 1$, using the fact that for every $i = 1, \dots, n$

$$\max\left\{|f_{P_{X_i}}(t)|, \exp\left(-\frac{t^2\sigma_i^2}{2}\right)\right\} \leq \exp\left(-\frac{t^2\sigma_i^2}{2} + \frac{\chi_1 t^3(\mathbb{E}[|X_i|^3] + \mathbb{E}[|X_i|]\sigma_i^2)}{B_n^3}\right),$$

we get

$$\begin{aligned}
\left|f_{S_n}(t) - e^{-t^2/2}\right| &\leq \sum_{i=1}^n \left|f_{P_{X_i}}\left(\frac{t}{B_n}\right) - e^{-\frac{t^2\sigma_i^2}{2B_n^2}}\right| e^{\frac{t^2\sigma_i^2}{2B_n^2}} e^{-\frac{t^2}{2} + \frac{\chi_1|t|^3\sum_{l=1}^n\left(\mathbb{E}[|X_l|^3] + \mathbb{E}[|X_i|]\sigma_i^2\right)}{B_n^3}} \\
&= \sum_{i=1}^n \left|f_{P_{X_i}}\left(\frac{t}{B_n}\right) - e^{-\frac{t^2\sigma_i^2}{2B_n^2}}\right| e^{-\frac{t^2}{2} + \frac{\chi_1|t|^3\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2\sigma_i^2}{2B_n^2}}.
\end{aligned}$$

By Equation (43), we have $\max_{1\leq i\leq n}\sigma_i^2 \leq B_n^2 \times (K_{4,n}/n)^{1/2}$ so that we obtain

$$\left|f_{S_n}(t) - e^{-t^2/2}\right| \leq \sum_{i=1}^n \left|f_{P_{X_i}}\left(\frac{t}{B_n}\right) - e^{-\frac{t^2\sigma_i^2}{2B_n^2}}\right| e^{-\frac{t^2}{2} + \frac{\chi_1|t|^3\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2}\sqrt{\frac{K_{4,n}}{n}}}.$$

Applying Lemma 2.8 in Shevtsova (2012), we get that for every variable X such that $\mathbb{E}[|X|^3]$ is finite, $|f(t) - e^{-\sigma^2 t^2}| \leq \mathbb{E}[|X|^3] \times |t|^3/6$. Therefore,

$$\begin{aligned}
\left|f_{S_n}(t) - e^{-t^2/2}\right| &\leq \sum_{i=1}^n \frac{\mathbb{E}[|X_i|^3]}{6B_n^3}|t|^3 \exp\left(-\frac{t^2}{2} + \frac{\chi_1|t|^3\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2}\sqrt{\frac{K_{4,n}}{n}}\right) \\
&= \frac{K_{3,n}}{6\sqrt{n}}|t|^3 \exp\left(-\frac{t^2}{2} + \frac{\chi_1|t|^3\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2}\sqrt{\frac{K_{4,n}}{n}}\right). \qquad (19)
\end{aligned}$$

Recalling that $t_0 = 1/\pi$, and integrating the latter equation, we have

$$\begin{aligned}
I_{3,2}(T) &= \frac{2}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{t_0 T} |\Psi(u/T)| \left|f_{S_n}(u) - e^{-u^2/2}\right| du, \\
&\leq \frac{K_{3,n}}{3\sqrt{n}T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T/\pi} |\Psi(u/T)|\, u^3 \exp\left(-\frac{u^2}{2} + \frac{\chi_1|u|^3\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{u^2}{2}\sqrt{\frac{K_{4,n}}{n}}\right) du \\
&= \frac{K_{3,n}}{3\sqrt{n}} J_2\big(3, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, K_{4,n}, T, n\big), \qquad (20)
\end{aligned}$$

where J_2 is as defined in Equation (53).

5.2.4 Improved bound on $I_{3,2}(T)$ under the assumption that $\mathbb{E}[X_i^3] = 0$ for all i

When $\mathbb{E}[X_i^3] = 0$ for all $i = 1, \dots, n$, the bound on $I_{3,2}(T)$ can be further improved. The proof mostly follows the reasoning of Section 5.2.3, with suitable modifications.

First, using a Taylor expansion of order 3 of $f_{P_{X_i}}$ around 0 (with explicit Lagrange remainder) and the inequality $|e^{-x} - 1 + x| \leq x^2/2$, we can claim for every real t

$$\left| f_{P_{X_i}}(t) - e^{-t^2\sigma_i^2/2} \right| \leq \frac{t^4\gamma_i}{24} + \frac{\sigma_i^4 t^4}{8} \leq \frac{t^4\gamma_i}{6},$$

Reasoning as in the proof of Lemma 2.7 in Shevtsova (2012) with $\delta = 1$, we obtain

$$\begin{aligned}
\left| f_{S_n}(t) - e^{-t^2/2} \right| &\leq \sum_{i=1}^n \frac{t^4\gamma_i}{6B_n^4} \exp\left(-\frac{t^2}{2} + \frac{\chi_1 |t|^3 \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2}\sqrt{\frac{K_{4,n}}{n}} \right) \\
&\leq \frac{K_{4,n}}{6n} t^4 \exp\left(-\frac{t^2}{2} + \frac{\chi_1 |t|^3 \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2}\sqrt{\frac{K_{4,n}}{n}} \right).
\end{aligned}$$

Plugging this into the definition of $I_{3,2}(T)$, we can write

$$\begin{aligned}
I_{3,2}(T) &= \frac{2}{T} \int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{t_0 T} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2} \right| du, \\
&\leq \frac{K_{4,n}}{3nT} \int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T/\pi} |\Psi(u/T)|\, u^4 \exp\left(-\frac{u^2}{2} + \frac{\chi_1 |u|^3 \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2}\sqrt{\frac{K_{4,n}}{n}} \right) du \\
&\leq \frac{K_{4,n}}{3n} J_2\big(4, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, K_{4,n}, T, n\big). \qquad (21)
\end{aligned}$$

5.2.5 Conclusion: end of the proof of Theorem 5.2.

We start from Equation (8). The quantity Ω_1 is bounded by Lemma 6.1, the quantity Ω_2 in Equation (15) and the quantity Ω_3 in Equation (18). As a result,

$$\begin{aligned}
\sup_{x \in \mathbb{R}} |\mathbb{P}(S_n \leq x) - G_n(x)| &\leq \Omega_1(t_0, T, |\lambda_{3,n}|/\sqrt{n}) + \Omega_2(t_0, T) + \Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \\
&\leq \frac{1.2533}{T} + \frac{0.3334|\lambda_{3,n}|}{T\sqrt{n}} + \frac{14.1961}{T^4} + \frac{|\lambda_{3,n}| \exp(-T^2/(2\pi^2))}{3\pi\sqrt{n}} \\
&\quad + \frac{67.0415}{T^4} + \frac{1.2187}{T^2} + I_{3,1}(T) + I_{3,2}(T) + I_{3,3}(T) \\
&\leq \frac{1.2533\widetilde{K}_{3,n}}{2\pi\sqrt{n}} + \frac{0.3334|\lambda_{3,n}|\widetilde{K}_{3,n}}{2\pi n} + \frac{1.2187\widetilde{K}_{3,n}^2}{4\pi n} + \frac{0.327K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2} \right) \\
&\quad + \frac{1.306 e_{1,n}(\varepsilon)|\lambda_{3,n}|^2}{36n} + r_{1,n},
\end{aligned}$$

where

$$\begin{aligned}
r_{1,n} &:= \frac{(14.1961 + 67.0415)\widetilde{K}_{3,n}^4}{16\pi^4 n^2} + \frac{|\lambda_{3,n}| \exp(-2n^2/\widetilde{K}_{3,n}^4)}{3\pi\sqrt{n}} + I_{3,2}(T) + I_{3,3}(T) \\
&\quad + \frac{1.0253}{\pi} \int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} u e^{-u^2/2} R_{1,n}(u, \varepsilon) du. \qquad (22)
\end{aligned}$$

Given the definition of $R_{1,n}(u, \varepsilon)$, the last term in (22) can be written as a sum of elements depending on ε and on positive powers of $\frac{K_{4,n}}{n}$ and $\frac{|\lambda_{3,n}|}{\sqrt{n}}$. Since $K_{4,n}$ and $|\lambda_{3,n}|$ are bounded by assumption, we can see based on the definition of $R_{1,n}(u,\varepsilon)$ that the last (and dominant) term in the definition of $r_{1,n}$ comes from $U_{1,2,n}$, defined in Equation (42) and gives the rate $O(n^{-5/4})$.

When $\mathbb{E}[X_i^3] = 0$ for every $i = 1, \dots, n$, we have $\lambda_{3,n} = 0$ which removes the corresponding terms. The dominant term in $r_{1,n}$ which stems from Equation (42) is null whenever $\sum_{j=1}^n |\mathbb{E}[X_j^3]| = 0$. Under the new assumption, this term disappears and the next term becomes the dominant one in the remainder $r_{1,n}$. We finally obtain the bound $r_{1,n} = O(n^{-3/2})$.

5.3 Proof of Theorem 2.3 under Assumption 2.2

We present and prove a more general result, Theorem 5.4, and choose $\varepsilon = 0.1$ to recover Theorem 2.3 under Assumption 2.2 (in Supplement A, we compute $e_3(0.1)$ and obtain $e_3(0.1) \leq 1.012$).

Theorem 5.4 (One-term Edgeworth expansion under Assumption 2.2)**.** *Let Assumption 2.2 hold. For every $\varepsilon \in (0, 1/3)$ and every $n \geq 2$*

$$\Delta_{n,E} \leq \frac{0.1995\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{1}{n}\Bigg\{0.031\widetilde{K}_{3,n}^2 + 0.327K_{4,n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + 0.054|\lambda_{3,n}|\widetilde{K}_{3,n} + 0.037e_3(\varepsilon)|\lambda_{3,n}|^2\Bigg\} + r_{1,n}, \tag{23}$$

where $r_{1,n}$ is given in Equation (29) and $e_3(\varepsilon) = e^{\varepsilon^2/6+\varepsilon^2/(2(1-3\varepsilon))^2}$. We remark that $r_{1,n} = O(n^{-5/4})$ whenever $K_{4,n} = O(1)$. If $\mathbb{E}[X_n^3] = 0$ the upper bound reduces to

$$\frac{0.1995\widetilde{K}_{3,n}}{\sqrt{n}} + \frac{1}{n}\Bigg\{0.031\widetilde{K}_{3,n}^2 + 0.327K_{4,n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right)\Bigg\} + r_{1,n}, \tag{24}$$

with $r_{1,n} = O(n^{-2})$ when $K_{4,n} = O(1)$.

The overall method of proof is close to that in Section 5.2. In particular, the start of the proof is unchanged: we apply Lemma 5.1, choosing $t_0 = 1/\pi$, and Ω_1 is still controlled by Lemma 6.1. The steps leading to an upper bound on Ω_2 in Section 5.2.1 remain valid as well so that we can write

$$\Omega_2(1/\pi, T) \leq \frac{67.0415}{T^4} + \frac{1.2187}{T^2}, \tag{25}$$

as in Equation (15).

5.3.1 Improved bound on Ω_3

The control of Ω_3 can be refined under Assumption 2.2. We have

$$\begin{aligned}\Omega_3\big(1/\pi, T, \lambda_{3,n}/\sqrt{n}\big) &= 2\int_0^{1/\pi} |\Psi(t)| \left| f_{S_n}(Tt) - e^{-(Tt)^2/2}\left(1 - \frac{\lambda_{3,n} i (Tt)^3}{6\sqrt{n}}\right)\right| dt \\ &\leq I_{4,1}(T) + I_{4,2}(T) + I_{4,3}(T),\end{aligned} \tag{26}$$

where

$$I_{4,1}(T) := \frac{2}{T}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\left(1 - \frac{iu^3\lambda_{3,n}}{6\sqrt{n}}\right)\right| du,$$

$$I_{4,2}(T) := \frac{2}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T/\pi} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\right| du,$$

$$I_{4,3}(T) := \frac{|\lambda_{3,n}|}{6\sqrt{n}} \int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T/\pi} |\Psi(u/T)|\, e^{-u^2/2}|u|^3 du.$$

The integrand of $I_{4,1}(T)$ can be upper bounded thanks to Lemma 6.3. We obtain

$$\begin{aligned}
I_{4,1}(T) &\leq \frac{K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) J_1\left(4, 0, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T\right) \\
&\quad + \frac{e_{2,n}(\varepsilon)|\lambda_{3,n}|^2}{36n} J_1\left(6, 0, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T\right) \\
&\quad + \frac{2}{T}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} |\Psi(u/T)| e^{-u^2/2} R_{2,n}(u,\varepsilon) du \\
&\leq \frac{0.327\, K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306 e_{2,n}(\varepsilon)|\lambda_{3,n}|^2}{36n} \\
&\quad + \frac{1.0253}{\pi}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} u e^{-u^2/2} R_{2,n}(u,\varepsilon) du, \qquad (27)
\end{aligned}$$

using the expression of J_1 given in Equation (52) and the first inequality in (9). Moreover, we remark that

$$I_{4,3}(T) = \frac{|\lambda_{3,n}|}{3\sqrt{n}} J_1(3, \sqrt{2\varepsilon n}/K_{4,n}^{1/4}, T/\pi, T).$$

As in Section 6.4.1, we can prove that $I_{4,3}(T)$ decays exponentially with n.

We finally control the term $I_{4,2}(T)$. Under the *i.i.d.* assumption, we can prove that for every real t

$$\left| f_{S_n}(t) - e^{-t^2/2} \right| \leq \frac{K_{3,n}}{6\sqrt{n}} |t|^3 \exp\left(-\frac{t^2}{2} + \frac{\chi_1 |t|^3 \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2}{2n} \right),$$

following the method of Section 5.2.3. Multiplying by $|\Psi(t)|$ and integrating this inequality, we get

$$I_{4,2}(T) \leq \frac{K_{3,n}}{3\sqrt{n}} J_3(3, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, T, n), \qquad (28)$$

where

$$J_3(v, w, \widetilde{K}_{3,n}, T, n) := \frac{1}{T}\int_v^w |\Psi(u/T)| u^p \exp\left(-\frac{u^2}{2}\left(1 - \frac{2\chi_1 |u| \widetilde{K}_{3,n}}{\sqrt{n}} - \frac{1}{n} \right) \right) du.$$

Recalling that $T = 2\pi\sqrt{n}/\widetilde{K}_{3,n}$ we obtain

$$\begin{aligned}
&J_3(3, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, T, n) \\
&= \frac{1}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{2\sqrt{n}/\widetilde{K}_{3,n}} |\Psi(u/T)| u^3 \exp\left(-\frac{u^2}{2}\left(1 - \frac{2\chi_1 |u| \widetilde{K}_{3,n}}{\sqrt{n}} - \frac{1}{n} \right) \right) du \\
&\leq \frac{1}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{2\sqrt{n}/\widetilde{K}_{3,n}} |\Psi(u/T)| u^3 \exp\left(-\frac{u^2}{2}\left(1 - 4\chi_1 - \frac{1}{n} \right) \right) du.
\end{aligned}$$

Note that $1 - 4\chi_1 - 1/n > 0.1$ as soon as $n \geq 2$. As in Section 6.4.1, we can prove that the latter term decays exponentially with n. The term $I_{4,2}(T)$ is thus negligible. If $\mathbb{E}[X_n^3] = 0$, the bound (28) can be further improved to

$$I_{4,2}(T) \leq \frac{K_{4,n}}{3n} J_3(3, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, T, n).$$

This can be recovered using the same techniques as in Section 5.2.4, ensuring again that $I_{4,2}(T)$ decays exponentially fast to zero with n.

5.3.2 Conclusion: end of the proof of Theorem 5.4

To conclude, we first use Equation (8), and manage all the terms separately. Ω_1 is bounded in Lemma 6.1, Ω_2 is bounded in Equation (25) and Ω_3 is bounded in Equation (27). We can claim that

$$\begin{aligned}
\sup_{x\in\mathbb{R}} \left|\mathbb{P}\left(S_n \leq x\right) - G_n(x)\right| &\leq \Omega_1(t_0, T, |\lambda_{3,n}|/\sqrt{n}) + \Omega_2(t_0, T) + \Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \\
&\leq \frac{1.2533}{T} + \frac{0.3334|\lambda_{3,n}|}{T\sqrt{n}} + \frac{14.1961}{T^4} + \frac{|\lambda_{3,n}|\exp(-T^2/(2\pi^2))}{3\pi\sqrt{n}} \\
&\qquad + \frac{67.0415}{T^4} + \frac{1.2187}{T^2} + I_{4,1}(T) + I_{4,2}(T) + I_{4,3}(T) \\
&\leq \frac{1.2533\,\widetilde{K}_{3,n}}{2\pi\sqrt{n}} + \frac{0.3334|\lambda_{3,n}|\,\widetilde{K}_{3,n}}{2\pi n} + \frac{1.2187\,\widetilde{K}_{3,n}^2}{4\pi^2 n} \\
&\quad + \frac{0.327\,K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306 e_3(\varepsilon)|\lambda_{3,n}|^2}{36n} + r_{1,n},
\end{aligned}$$

where

$$\begin{aligned}
r_{1,n} := &\frac{(14.1961 + 67.0415)\widetilde{K}_{3,n}^4}{2^4\pi^4 n^2} + \frac{|\lambda_{3,n}|\exp(-2n^2/\widetilde{K}_{3,n}^4)}{3\pi\sqrt{n}} + I_{4,2}(T) + I_{4,3}(T) \\
&+ \frac{1.306(e_{2,n}(\varepsilon) - e_3(\varepsilon))|\lambda_{3,n}|^2}{36n} + \frac{1.0253}{\pi}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} u e^{-u^2/2} R_{2,n}(u,\varepsilon)\,du \qquad (29)
\end{aligned}$$

The last term of Equation (29) is of order $n^{-3/2}$ given the definition of $R_{2,n}(u, \varepsilon)$. The quantity $(e_{2,n}(\varepsilon) - e_3(\varepsilon))/n$ is of the order $n^{-5/4}$ and is therefore dominant. Grouping terms together yields Equation (23).

When $\mathbb{E}[X_n^3] = 0$, we have $\lambda_{3,n} = 0$ which removes the corresponding terms. Furthermore, the leading terms in $r_{1,n}$ are $O(n^{-2})$ when $\lambda_{3,n} = 0$ given the definition of $R_{2,n}(u, \varepsilon)$.

5.4 Proof of Theorem 3.1 under Assumption 2.1

We use Theorem 5.5, proved below, with the choice $\varepsilon = 0.1$. We let $t_1^* := \theta_1^*/(2\pi)$ where θ_1^* is the unique root in $(0, 2\pi)$ of the equation $\theta^2 + 2\theta\sin(\theta) + 6(\cos(\theta) - 1) = 0$. Numerical approximations conducted in Supplement A allow us to write $t_1^* \approx 0.635$.

Theorem 5.5 (Alternative one-term Edgeworth expansion under Assumption 2.1)**.** *Let Assumption 2.1 hold. For every $\varepsilon \in (0, 1/3)$ and every $n \geq 2$*

$$\begin{aligned}
\Delta_{n,E} \leq &\frac{1}{n}\left\{0.327 K_{4,n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + 0.037 e_{1,n}(\varepsilon)\lambda_{3,n}^2\right\} \\
&+ \frac{1.0253}{\pi}\int_{a_n}^{b_n} \frac{|f_{S_n}(t)|}{t}\,dt + r_{2,n}, \qquad (30)
\end{aligned}$$

where $a_n = 2t_1^*\pi/\widetilde{K}_{3,n}$, $b_n = 4\pi^2 n/(t_1^*\widetilde{K}_{3,n}^2)$ and $r_{2,n}$ is given in Equation (34). When $K_{4,n} = O(1)$, we have $r_{2,n} = O(n^{-5/4})$. If $\mathbb{E}[X_i^3] = 0$ for every $i = 1, \dots, n$, the upper bound becomes

$$\frac{0.327 K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.0253}{\pi}\int_{a_n}^{b_n} \frac{|f_{S_n}(t)|}{t}\,dt + r_{2,n}, \qquad (31)$$

where $r_{2,n} = O(n^{-3/2})$ when $K_{4,n} = O(1)$.

As before, we start the proof of Theorem 5.5 by applying Lemma 5.1 (with $t_0 = 1/\pi$) and use Lemma 6.1 to control Ω_1. There remains to bound Ω_2 and Ω_3. We choose $T = 16\pi^4 n^2/\widetilde{K}_{3,n}^4$.

5.4.1 Bound on Ω_3

We decompose this term in five parts

$$\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \leq I_{5,1}(T) + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + I_{5,5}(T),$$

where

$$\begin{aligned}
I_{5,1}(T) &:= \frac{2}{T}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\left(1 - \frac{\lambda_{3,n} i u^3}{6\sqrt{n}}\right)\right| du, \\
I_{5,2}(T) &:= E_{1,n}\frac{|\lambda_{3,n}|}{3T\sqrt{n}}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T^{1/4}/\pi} |\Psi(u/T)| u^3 e^{-u^2/2} du, \\
I_{5,3}(T) &:= E_{1,n}\frac{2}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T^{1/4}/\pi} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\right| du, \\
I_{5,4}(T) &:= E_{2,n}\frac{|\lambda_{3,n}|}{3T\sqrt{n}}\int_{T^{1/4}/\pi}^{T/\pi} |\Psi(u/T)|\, |u|^3 e^{-u^2/2} du, \\
I_{5,5}(T) &:= E_{2,n}\frac{2}{T}\int_{T^{1/4}/\pi}^{T/\pi} |\Psi(u/T)| \left| f_{S_n}(u) - e^{-u^2/2}\right| du,
\end{aligned}$$

where $E_{1,n} := \mathbb{1}_{\{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4} < T^{1/4}/\pi\}}$ and $E_{2,n} := \mathbb{1}_{\{T^{1/4} < T\}}$. Note that if $T^{1/4} > T$ or $\sqrt{2\varepsilon}(n/K_{4,n})^{1/4} > T^{1/4}/\pi$, our bounds are still valid and can even be improved in the sense that the corresponding integrals can be removed.

Remarking that $I_{5,1}(T) = I_{3,1}(T)$, we can bound this term using Equation (17). We now turn to $I_{5,2}(T)$ and $I_{5,3}(T)$. Assume that $\sqrt{2\varepsilon}(n/K_{4,n})^{1/4} < T^{1/4}/\pi$, as the bound is trivially proved in the other case. We remark that $I_{5,2}(T)$ (*resp.* $I_{5,3}(T)$) can be bounded exactly as $I_{3,3}(T)$ in Section 5.2.2 (*resp.* as $I_{3,2}(T)$). Consequently

$$\begin{aligned}
I_{5,2}(T) &\leq \frac{|\lambda_{3,n}|}{3\sqrt{n}}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T^{1/4}/\pi} \frac{1.0253}{2\pi} u^2 e^{-u^2/2} du \\
&= \frac{1.0253\,|\lambda_{3,n}|}{3\pi\sqrt{2}\sqrt{n}}\Big(\Gamma\big(3/2\,,\,\varepsilon(n/K_{4,n})^{1/2}\big) - \Gamma\big(3/2\,,\,T^{1/2}/2\pi^2\big)\Big),
\end{aligned}$$

and

$$\begin{aligned}
I_{5,3}(T) &\leq \frac{2}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T^{1/4}/\pi} |\Psi(u/T)| \frac{K_{3,n}}{6\sqrt{n}} |t|^3 \exp\left(-\frac{t^2}{2} + \frac{\chi_1 |t|^3 \widetilde{K}_{3,n}}{\sqrt{n}} + \frac{t^2\sqrt{K_{4,n}}}{2\sqrt{n}}\right) du \\
&= \frac{K_{3,n}}{3\sqrt{n}} J_2\big(3, \sqrt{2\varepsilon}/(nK_{4,n})^{1/4}, T^{1/4}/\pi, \widetilde{K}_{3,n}, K_{4,n}, T, n\big),
\end{aligned}$$

and these terms decrease to zero exponentially fast when n goes to infinity.

Note that whenever $\mathbb{E}[X_i^3] = 0$ for all $i = 1, \dots, n$, the improvements detailed in Section 5.2.4 can be used as well, resulting in a tighter bound on $I_{5,2}$ where the factor $K_{3,n}/\sqrt{n}$ is replaced by $K_{4,n}/n$ and the first argument of J_2 becomes 4 instead of 3.

We have to deal with $I_{5,4}(T)$ and $I_{5,5}(T)$. We assume that $T^{1/4} < T$, otherwise there is nothing to prove. Obviously $I_{5,4}(T)$ can be bounded in a similar vein as $I_{5,2}(T)$

$$I_{5,4}(T) = \frac{1.0253\,|\lambda_{3,n}|}{3\pi\sqrt{2}\sqrt{n}}\Big(\Gamma\big(3/2\,,\,T^{1/2}/2\pi^2\big) - \Gamma\big(3/2\,,\,T^2/2\pi^2\big)\Big),$$

and it converges exponentially fast to zero.

To control $I_{5,5}(T)$, we write

$$I_{5,5}(T) \leq J_3(T) + J_4(T) + J_5(T),$$

where

$$J_3(T) := \frac{2}{T}\int_{T^{1/4}/\pi}^{t_1^* T^{1/4}} |\Psi(u/T)|\,|f_{S_n}(u)|\,du = \frac{2}{T^{3/4}}\int_{1/\pi}^{t_1^*} |\Psi(v/T^{3/4})| \left| f_{S_n}(T^{1/4}v) \right| dv,$$

$$J_4(T) := \mathbb{1}_{\{t_1^* T^{1/4} < T/\pi\}} \frac{2}{T}\int_{t_1^* T^{1/4}}^{T/\pi} |\Psi(u/T)|\,|f_{S_n}(u)|\,du,$$

$$J_5(T) := \frac{2}{T}\int_{T^{1/4}/\pi}^{T/\pi} |\Psi(u/T)| e^{-u^2/2} du.$$

By Lemma 5.3 and our choice of T, we know $|f_{S_n}(T^{1/4}v)|$ can be upper bounded by $e^{-\frac{T^{1/2}v^2}{2}(1-4\pi\chi_1|v|)}$ when $v \in [1/\pi, t_1^*]$. We get (using $1 - 4\pi\chi_1 t_1^* > 0$)

$$\begin{aligned}
J_3(T) \leq & \frac{2}{T^{3/4}}\int_{1/\pi}^{t_1^*} |\Psi(v/T^{3/4})| e^{-\frac{T^{1/2}v^2}{2}(1-4\pi\chi_1|v|)} dv \\
\leq & \frac{1.0253}{\pi}\int_{1/\pi}^{t_1^*} v^{-1} e^{-\frac{T^{1/2}v^2}{2}(1-4\pi\chi_1 t_1^*)} dv \\
= & \frac{1.0253}{2\pi}\int_{T^{1/2}(1-4\pi\chi_1 t_1^*)/(2\pi^2)}^{t_1^{*2}T^{1/2}(1-4\pi\chi_1 t_1^*)/2} u^{-1} e^{-u} du \\
= & \frac{1.0253}{2\pi}\left(\Gamma\left(T^{1/2}(1-4\pi\chi_1 t_1^*)/(2\pi^2), 0\right) - \Gamma\left(t_1^{*2}T^{1/2}(1-4\pi\chi_1 t_1^*)/2, 0\right)\right).
\end{aligned}$$

As a result, we conclude that $J_3(T)$ decreases to zero exponentially fast with n.

To control $J_4(T)$, we use the properties of $u \mapsto \Psi(u)$ to write

$$J_4(T) \leq \mathbb{1}_{\{t_1^* T^{1/4} < T/\pi\}} \frac{1.0253}{\pi}\int_{t_1^* T^{1/4}}^{T/\pi} u^{-1}|f_{S_n}(u)|du.$$

To upper bound $J_5(T)$, we can reason as for $J_3(T)$ to conclude that this term converges to zero exponentially fast.

As a result, we conclude

$$\begin{aligned}
\Omega_3(1/\pi, T, \lambda_{3,n}/\sqrt{n}) \leq & I_{5,1}(T) + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + I_{5,5}(T) \\
\leq & \frac{0.327K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306e_{1,n}(\varepsilon)}{36}\frac{|\lambda_{3,n}|^2}{n} \\
& + \mathbb{1}_{\{t_1^* T^{1/4} < T/\pi\}} \frac{1.0253}{\pi}\int_{t_1^* T^{1/4}}^{T/\pi} u^{-1}|f_{S_n}(u)|du + O(n^{-5/4}). \qquad (32)
\end{aligned}$$

5.4.2 Conclusion: end of the proof of Theorem 5.5

We can write

$$\Omega_2(1/\pi, T) \leq \frac{1.0253}{\pi}\int_{T/\pi}^{T} u^{-1}|f_{S_n}(u)|du. \qquad (33)$$

To sum up, we first use Equation (8), and manage all the terms separately. Ω_1 is bounded in Lemma 6.1, Ω_2 is bounded in Equation (33) and Ω_3 is bounded in Equation (32). Given the definitions of a_n and b_n, we conclude

$$\begin{aligned}
\sup_{x\in\mathbb{R}} |\mathbb{P}(S_n \le x) - G_n(x)| &\le \Omega_1(t_0, T, |\lambda_{3,n}|/\sqrt{n}) + \Omega_2(t_0, T) + \Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \\
&\le \frac{1.2533}{T} + \frac{0.3334|\lambda_{3,n}|}{T\sqrt{n}} + \frac{14.1961}{T^4} + \frac{|\lambda_{3,n}|\exp(-T^2/(2\pi^2))}{3\pi\sqrt{n}} \\
&\qquad + \frac{1.0253}{\pi}\int_{T/\pi}^{T} u^{-1}|f_{S_n}(u)|du + I_{5,1}(T) + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + I_{5,5}(T) \\
&\le \frac{0.327\,K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306 e_{1,n}(\varepsilon)|\lambda_{3,n}|^2}{36n} \\
&\qquad + \frac{1.0253}{\pi}\int_{a_n}^{b_n} u^{-1}|f_{S_n}(u)|du + r_{2,n},
\end{aligned}$$

where

$$\begin{aligned}
r_{2,n} :=& \frac{1.2533\widetilde{K}_{3,n}^4}{16\pi^4 n^2} + \frac{0.3334|\lambda_{3,n}|\widetilde{K}_{3,n}^4}{16\pi^4 n^{5/2}} + \frac{14.1961\widetilde{K}_{3,n}^{16}}{16^4\pi^{16}n^8} + \frac{|\lambda_{3,n}|\exp(-128\pi^6 n^4/\widetilde{K}_{3,n}^8)}{3\pi\sqrt{n}} \\
&+ I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + J_3(T) + J_5(T) \\
&+ \frac{1.0253}{\pi}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} ue^{-u^2/2}R_{1,n}(u,\varepsilon)du. \qquad (34)
\end{aligned}$$

All terms but the last one in the definition of $r_{2,n}$ are at most of order n^{-2}. As explained in Section 5.2, the last term is of order $O(n^{-5/4})$ so that $r_{2,n} = O(n^{-5/4})$.

When $\mathbb{E}[X_i^3] = 0$ for every $i = 1, \dots, n$, we have $\lambda_{3,n} = 0$, which removes the corresponding terms. Under the new assumption, the dominant term in $r_{2,n}$ has the rate $n^{-3/2}$ which implies $r_{2,n} = O(n^{-3/2})$.

5.5 Proof of Theorem 3.1 under Assumption 2.2

We use Theorem 5.6, proved below, with the choice $\varepsilon = 0.1$.

Theorem 5.6 (Alternative one-term Edgeworth expansion under Assumption 2.2)**.** *Let Assumption 2.2 hold. For every $\varepsilon \in (0, 1/3)$ and every $n \ge 1$*

$$\begin{aligned}
\Delta_{n,E} \le& \frac{1}{n}\left\{0.327K_{4,n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + 0.037e_3(\varepsilon)\lambda_{3,n}^2\right\} \\
&+ \frac{1.0253}{\pi}\int_{a_n}^{b_n}\frac{|f_{S_n}(t)|}{t}dt + r_{2,n}, \qquad (35)
\end{aligned}$$

where $a_n = 2t_1^*\pi/\widetilde{K}_{3,n}$, $b_n = 4\pi^2 n/(t_1^*\widetilde{K}_{3,n}^2)$ and $r_{2,n}$ is given in Equation (39). We have $r_{2,n} = O(n^{-5/4})$ as soon as $K_{4,n} = O(1)$. If $\mathbb{E}[X_n^3] = 0$, the upper bound becomes

$$\frac{0.327K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.0253}{\pi}\int_{a_n}^{b_n}\frac{|f_{S_n}(t)|}{t}dt + r_{2,n}, \qquad (36)$$

where $r_{2,n} = O(n^{-2})$ when $K_{4,n} = O(1)$.

The proof of Theorem 5.6 is very similar to that of Theorem 5.5. We start by applying Lemma 5.1 (with $t_0 = 1/\pi$) and use Lemma 6.1 to control Ω_1. There remains to bound Ω_2 and Ω_3. We choose $T = 16\pi^4 n^2/\widetilde{K}_{3,n}^4$ again.

5.5.1 Bound on Ω_3

As in the proof of Theorem 5.5, we write

$$\Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \leq I_{5,1}(T) + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + I_{5,5}(T),$$

where the terms in the upper bound are defined in Section 5.4.1.

The terms $I_{5,2}(T)$, $I_{5,3}(T)$ and $I_{5,4}(T)$ can be controlled similarly as in Section 5.4.1. We have $I_{5,1}(T) = I_{4,1}(T)$ so that we can use the upper bound in (27).

We upper bound $I_{5,5}(T)$ as in Section 5.4.1

$$I_{5,5}(T) \leq J_3(T) + J_4(T) + J_5(T).$$

The proof that $J_3(T)$ and $J_5(T)$ decrease exponentially fast to zero with n is still valid.

We finally obtain

$$\begin{aligned}\Omega_3(1/\pi, T, \lambda_{3,n}/\sqrt{n}) \leq & I_{5,1}(T) + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + I_{5,5}(T) \\ \leq & \frac{0.327K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306e_3(\varepsilon)}{36}\frac{|\lambda_{3,n}|^2}{n} \\ & + \mathbb{1}_{\{t_1^* T^{1/4} < T/\pi\}} \frac{1.0253}{\pi} \int_{t_1^* T^{1/4}}^{T/\pi} u^{-1}|f_{S_n}(u)|du. \end{aligned} \tag{37}$$

5.5.2 Conclusion: end of the proof of Theorem 5.6

We have

$$\Omega_2(1/\pi, T) \leq \frac{1.0253}{\pi} \int_{T/\pi}^{T} u^{-1}|f_{S_n}(u)|du. \tag{38}$$

To sum up, we first use Equation (8), and manage all the terms separately. Ω_1 is bounded in Lemma 6.1, Ω_2 is bounded in Equation (38) and Ω_3 is bounded in Equation (37). Using the definitions of a_n and b_n, we conclude

$$\begin{aligned}\sup_{x\in\mathbb{R}} |\mathbb{P}(S_n \leq x) - G_n(x)| \leq & \ \Omega_1(t_0, T, |\lambda_{3,n}|/\sqrt{n}) + \Omega_2(t_0, T) + \Omega_3(t_0, T, \lambda_{3,n}/\sqrt{n}) \\ \leq & \frac{1.2533}{T} + \frac{0.3334|\lambda_{3,n}|}{T\sqrt{n}} + \frac{14.1961}{T^4} + \frac{|\lambda_{3,n}|\exp(-T^2/(2\pi^2))}{3\pi\sqrt{n}} \\ & + \frac{2}{T}\int_{T/\pi}^{T} |\Psi(u/T)|\,|f_{P_{X_1/\sigma_{1,n}}}(u/\sqrt{n})|^n du \\ & + I_{5,1}(T) + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + I_{5,5}(T) \\ \leq & \frac{0.327\,K_{4,n}}{n}\left(\frac{1}{12} + \frac{1}{4(1-3\varepsilon)^2}\right) + \frac{1.306e_3(\varepsilon)|\lambda_{3,n}|^2}{36n} \\ & + \frac{1.0253}{\pi}\int_{a_n}^{b_n} u^{-1}|f_{S_n}(u)|du + r_{2,n}, \end{aligned}$$

where

$$\begin{aligned} r_{2,n} := & \frac{1.2533\widetilde{K}_{3,n}^4}{16\pi^4 n^2} + \frac{0.3334|\lambda_{3,n}|\widetilde{K}_{3,n}^4}{16\pi^4 n^{5/2}} + \frac{14.1961\widetilde{K}_{3,n}^{16}}{16^4\pi^{16}n^8} + \frac{|\lambda_{3,n}|\exp(-128\pi^6 n^4/\widetilde{K}_{3,n}^8)}{3\pi\sqrt{n}} \\ & + I_{5,2}(T) + I_{5,3}(T) + I_{5,4}(T) + J_3(T) + J_5(T) \\ & + \frac{1.306(e_{2,n}(\varepsilon) - e_3(\varepsilon))|\lambda_{3,n}|^2}{36n} + \frac{1.0253}{\pi}\int_0^{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}} ue^{-u^2/2}R_{2,n}(u,\varepsilon)du. \end{aligned} \tag{39}$$

All terms but the last one in the definition of $r_{2,n}$ are at most of order n^{-2}. As explained in Section 5.3, the last term is of order $O(n^{-5/4})$ so that $r_{2,n} = O(n^{-5/4})$.

When $\mathbb{E}[X_n^3] = 0$, we have $\lambda_{3,n} = 0$ which removes the corresponding terms. Under the new assumption, the dominant term in $r_{2,n}$ has the rate n^{-2} which implies $r_{2,n} = O(n^{-2})$.

6 Technical lemmas

6.1 Control of the term Ω_1

Lemma 6.1. *For every $T > 0$ we have*

$$\Omega_1(1/\pi, T, |\lambda_{3,n}|/\sqrt{n}) \leq \frac{1.2533}{T} + \frac{0.3334|\lambda_{3,n}|}{T\sqrt{n}} + \frac{14.1961}{T^4} + \frac{|\lambda_{3,n}| \exp(-T^2/(2\pi^2))}{3\pi\sqrt{n}}. \tag{40}$$

Proof. With the choice $t_0 = 1/\pi$, the function $\Omega_1(1/\pi, T, v)$ can be decomposed as

$$\Omega_1(t_0, T, v) := \frac{I_{1,1}(T)}{T} + v \times \frac{I_{1,2}(T)}{T} + \frac{I_{1,3}(T)}{T^4} + v \times \frac{I_{1,4}(T)}{T}$$

where

$$\begin{aligned}
I_{1,1}(T) &:= T \int_0^{1/\pi} \left| 2\Psi(t) - \frac{i}{\pi t} \right| e^{-(Tt)^2/2} dt, \\
I_{1,2}(T) &:= T^4 \int_0^{1/\pi} \left| 2\Psi(t) - \frac{i}{\pi t} \right| e^{-(Tt)^2/2} \frac{t^3}{6} dt, \\
I_{1,3}(T) &:= T^4 \frac{1}{\pi} \int_{1/\pi}^{+\infty} \frac{e^{-(Tt)^2/2}}{t} dt = \frac{T^4}{2\pi} \Gamma\left(0\,, \frac{T^2}{2\pi^2}\right), \\
I_{1,4}(T) &:= T^4 \frac{1}{\pi} \int_{1/\pi}^{+\infty} e^{-(Tt)^2/2} \frac{t^3}{6} dt = T^4 \frac{1}{\pi} \int_{T^2/(2\pi^2)}^{+\infty} e^{-u} \frac{u}{3T^4} dt = \frac{1}{3\pi} e^{-T^2/(2\pi^2)}.
\end{aligned}$$

On the last two lines, we used the change of variable $u = (tT)^2/2$ and the incomplete Gamma function $\Gamma(a, x) := \int_x^{+\infty} u^{a-1} e^{-u} du$ which can be computed numberically using the package `expint` (Goulet, 2016). We estimate numerically the first two integrals using the R package `cubature` (Narasimhan et al., 2020) and optimize using the `optimize` function with the `L-BFGS-B` method (see the code in Supplement A), we find the following upper bounds:

$$\sup_{T \geq 0} I_{1,1}(T) \leq 1.2533, \quad \sup_{T \geq 0} I_{1,2}(T) \leq 0.3334, \quad \text{and} \quad \sup_{T \geq 0} I_{1,3}(T) \leq 14.1961,$$

which finishes the proof.

□

Note that the first term on the right-hand side of (40) is of leading order as soon as $|\lambda_{3,n}|/\sqrt{n} = o(1)$ and $T = T(n) = o(1)$. Our approach is related to the one used in Shevtsova (2012), except that we do not upper bound Ω_1 analytically, which allows us to get a sharper control on this term. To further highlight the gains from using numerical approximations instead of direct analytical upper bounds, we remark that from $\left|\Psi(t) - \frac{i}{2\pi t}\right| \leq \frac{1}{2}\left(1 - |t| + \frac{\pi^2 t^2}{18}\right)$ and some straightforward integration steps, we get

$$\begin{aligned}I_{1,1}(T) &\leq T\int_0^{1/\pi}\left(1-|t|+\frac{\pi^2t^2}{18}\right)e^{-(Tt)^2/2}dt\\&=\sqrt{2\pi}\left(\Phi(T/\pi)-\frac{1}{2}\right)+\frac{1}{T}\left(e^{-(T/\pi)^2/2}-1\right)+\frac{\pi^{5/2}}{9\sqrt{2}T^2}\mathbb{E}_{U\sim\mathcal{N}(0,1)}[U^2\mathbb{1}\left\{0\leq U\leq T/\pi\right\}]\\&\leq\sqrt{2\pi}+\frac{1}{T}\left(e^{-T^2/(2\pi^2)}-1\right)+\frac{\pi^{5/2}}{9\sqrt{2}T^2},\end{aligned}$$

whose main term is approximately twice as large as the numerical bound 1.2533 that we obtained before.

6.2 Control of the residual term in an Edgeworth expansion under Assumption 2.1

For $\varepsilon\in(0,1/3)$ and $t\geq 0$, let us define the following quantities:

$$\begin{aligned}R_{1,n}(t,\varepsilon) &:= \frac{U_{1,1,n}(t)+U_{1,2,n}(t)}{2(1-3\varepsilon)^2}\\&\quad+e_1(\varepsilon)\left(\frac{t^8K_{4,n}^2}{2n^2}\left(\frac{1}{24}+\frac{P_{1,n}(\varepsilon)}{2(1-3\varepsilon)^2}\right)^2+\frac{|t|^7|\lambda_{3,n}|K_{4,n}}{6n^{3/2}}\left(\frac{1}{24}+\frac{P_{1,n}(\varepsilon)}{2(1-3\varepsilon)^2}\right)\right),\\P_{1,n}(\varepsilon) &:= \frac{144+48\varepsilon+4\varepsilon^2+\left\{96\sqrt{2\varepsilon}+32\varepsilon+16\sqrt{2}\varepsilon^{3/2}\right\}\mathbb{1}\left\{\exists i\in\{1,...,n\}:\mathbb{E}[X_i^3]\neq 0\right\}}{576},\\e_1(\varepsilon) &:= \exp\left(\varepsilon^2\left(\frac{1}{6}+\frac{2P_{1,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)\right),\end{aligned}\tag{41}$$

$$\begin{aligned}U_{1,1,n}(t) &:= \frac{t^6}{24}\left(\frac{K_{4,n}}{n}\right)^{3/2}+\frac{t^8}{24^2}\left(\frac{K_{4,n}}{n}\right)^2,\\U_{1,2,n}(t) &:= \left(\frac{|t|^5}{6}\left(\frac{K_{4,n}}{n}\right)^{5/4}+\frac{t^6}{36}\left(\frac{K_{4,n}}{n}\right)^{3/2}+\frac{|t|^7}{72}\left(\frac{K_{4,n}}{n}\right)^{7/4}\right)\mathbb{1}\left\{\exists i\in\{1,...,n\}:\mathbb{E}[X_i^3]\neq 0\right\}.\end{aligned}\tag{42}$$

We want to show the following lemma:

Lemma 6.2. *Under Assumption 2.1, for every $\varepsilon\in(0,1/3)$ and t such that $|t|\leq\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$, we have*

$$\left|f_{S_n}(t)-e^{-\frac{t^2}{2}}\left(1-\frac{it^3\lambda_{3,n}}{6\sqrt{n}}\right)\right|\leq e^{-t^2/2}\left\{\frac{t^4K_{4,n}}{8n}\left(\frac{1}{3}+\frac{1}{(1-3\varepsilon)^2}\right)+\frac{e_1(\varepsilon)|t|^6|\lambda_{3,n}|^2}{72n}+R_{1,n}(t,\varepsilon)\right\}.$$

Proof of Lemma 6.2: Remember that $\gamma_j:=\mathbb{E}[X_j^4]$, $\sigma_j:=\sqrt{\mathbb{E}[X_j^2]}$, $B_n:=\sqrt{\sum_{i=1}^n\mathbb{E}[X_i^2]}$ and $K_{4,n}:=n^{-1}\sum_{i=1}^n\mathbb{E}[X_i^4]\,/\left(n^{-1}B_n^2\right)^2$. Applying Cauchy-Schwartz inequality, we get

$$\max_{1\leq j\leq n}\sigma_j^2\leq\max_{1\leq j\leq n}\gamma_j^{1/2}\leq\left(\sum_{j=1}^n\gamma_j\right)^{1/2}=B_n^2(K_{4,n}/n)^{1/2},\tag{43}$$

$$\max_{1\leq j\leq n}\mathbb{E}[|X_j|^3]\leq\max_{1\leq j\leq n}\gamma_j^{3/4}\leq\left(\sum_{j=1}^n\gamma_j\right)^{3/4}=B_n^3(K_{4,n}/n)^{3/4},\tag{44}$$

and

$$\max_{1\leq j\leq n} \gamma_j \leq \sum_{j=1}^{n} \gamma_j = B_n^4 K_{4,n}/n. \tag{45}$$

Combining (43), (44) and (45), we observe that for every $\varepsilon \in (0,1)$ and t such that $|t| \leq \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$,

$$\max_{1\leq j\leq n} \left\{ \frac{\sigma_j^2 t^2}{2B_n^2} + \frac{\mathbb{E}[|X_j|^3] \times |t|^3}{6B_n^3} + \frac{\gamma_j t^4}{24B_n^4} \right\} \leq 3\varepsilon. \tag{46}$$

As we assume that X_j has a moment of order four for every $j = 1, \dots, n$, the characteristic functions $(f_{P_{X_j}})_{j=1,\dots,n}$ are four times differentiable on $\mathbb{R}$. Applying a Taylor-Lagrange expansion, we get the existence of a complex number $\theta_{1,j,n}(t)$ such that $|\theta_{1,j,n}(t)| \leq 1$ and

$$U_{j,n}(t) := f_{P_{X_j}}(t/B_n) - 1 = -\frac{\sigma_j^2 t^2}{2B_n^2} - \frac{i\mathbb{E}[X_j^3]\, t^3}{6B_n^3} + \frac{\theta_{1,j,n}(t)\gamma_j t^4}{24B_n^4},$$

for every $t \in \mathbb{R}$ and $j = 1, \dots, n$. Let log stand for the principal branch of the complex logarithm function. For every $\varepsilon \in (0, 1/3)$ and t such that $|t| \leq \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$, Equation (46) shows that $|U_{j,n}(t)| \leq 3\varepsilon < 1$, so that we can use another Taylor-Lagrange expansion. This ensures existence of a complex number $\theta_{2,j,n}(t)$ such that $|\theta_{2,j,n}(t)| \leq 1$ and

$$\log(f_{P_{X_j}}(t/B_n)) = \log(1 + U_{j,n}(t)) = U_{j,n}(t) - \frac{U_{j,n}(t)^2}{2(1 + \theta_{2,j,n}(t)U_{j,n}(t))^2}.$$

Summing over $j = 1, \dots, n$ and exponentiating, we can claim that under the same conditions on t and ε,

$$f_{S_n}(t) = \exp\left(-\frac{t^2}{2} - \frac{it^3\lambda_{3,n}}{6\sqrt{n}} + t^4 \sum_{j=1}^{n} \frac{\theta_{1,j,n}(t)\gamma_j}{24B_n^4} - \sum_{j=1}^{n} \frac{U_{j,n}(t)^2}{2(1 + \theta_{2,j,n}(t)U_{j,n}(t))^2} \right).$$

A third Taylor-Lagrange expansion guarantees existence of a complex number $\theta_{3,n}(t)$ with modulus at most $\exp\left(\frac{t^4 K_{4,n}}{24n} + \sum_{j=1}^{n} \frac{|U_{j,n}(t)|^2}{2|1+\theta_{2,j,n}(t)U_{j,n}(t)|^2} \right)$ such that

$$\begin{aligned} f_{S_n}(t) = e^{-t^2/2} \Bigg(1 &- \frac{it^3\lambda_{3,n}}{6\sqrt{n}} + t^4 \sum_{j=1}^{n} \frac{\theta_{1,j,n}(t)\gamma_j}{24B_n^4} - \sum_{j=1}^{n} \frac{U_{j,n}(t)^2}{2(1 + \theta_{2,j,n}(t)U_{j,n}(t))^2} \\ &+ \frac{\theta_{3,n}(t)}{2} \left(-\frac{it^3\lambda_{3,n}}{6\sqrt{n}} + t^4 \sum_{j=1}^{n} \frac{\theta_{1,j,n}(t)\gamma_j}{24B_n^4} - \sum_{j=1}^{n} \frac{U_{j,n}(t)^2}{2(1 + \theta_{2,j,n}(t)U_{j,n}(t))^2} \right)^2 \Bigg). \end{aligned}$$

Using the triangle inequality and its reverse version, as well as the restriction on $|t| \leq \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$, we can write

$$\begin{aligned} &\left| f_{S_n}(t) - e^{-t^2/2}\left(1 - \frac{it^3\lambda_{3,n}}{6\sqrt{n}}\right) \right| \leq e^{-t^2/2} \times \Bigg(\frac{t^4 K_{4,n}}{24n} + \frac{1}{2(1-3\varepsilon)^2} \sum_{j=1}^{n} |U_{j,n}(t)|^2 \\ &+ \frac{1}{2} \exp\left(\frac{\varepsilon^2}{6} + \frac{1}{2(1-3\varepsilon)^2} \sum_{j=1}^{n} |U_{j,n}(t)|^2 \right) \times \left(\frac{|t|^3|\lambda_{3,n}|}{6\sqrt{n}} + \frac{t^4 K_{4,n}}{24n} + \frac{1}{2(1-3\varepsilon)^2} \sum_{j=1}^{n} |U_{j,n}(t)|^2 \right)^2 \Bigg). \end{aligned} \tag{47}$$

We now control $\sum_{j=1}^n |U_{j,n}(t)|^2$. We first expand the squares, giving the decomposition

$$\sum_{j=1}^n |U_{j,n}(t)|^2 = \frac{t^4 \sum_{j=1}^n \sigma_j^4}{4B_n^4} + \frac{t^6 \sum_{j=1}^n |\mathbb{E}[X_j^3]|^2}{36B_n^6} + \frac{t^8 \sum_{j=1}^n \gamma_j^2}{24^2 B_n^8} + \frac{|t|^5 \sum_{j=1}^n \sigma_j^2 |\mathbb{E}[X_j^3]|}{6B_n^5} + \frac{t^6 \sum_{j=1}^n \sigma_j^2 \gamma_j}{24B_n^6} + \frac{|t|^7 \sum_{j=1}^n |\mathbb{E}[X_j^3]| \gamma_j}{72B_n^7}. \tag{48}$$

Using Equations (43)-(45), we can bound the right-hand side of Equation (48) in the following manner

$$\frac{t^4 \sum_{j=1}^n \sigma_j^4}{4B_n^4} \leq \frac{t^4 K_{4,n}}{4n},$$

$$\frac{t^6 \sum_{j=1}^n \sigma_j^2 \gamma_j}{24B_n^6} + \frac{t^8 \sum_{j=1}^n \gamma_j^2}{24^2 B_n^8} \leq \frac{t^6}{24} \left(\frac{K_{4,n}}{n}\right)^{3/2} + \frac{t^8}{24^2} \left(\frac{K_{4,n}}{n}\right)^2 =: U_{1,1,n}(t),$$

and

$$\begin{aligned}
&\frac{|t|^5 \sum_{j=1}^n \sigma_j^2 |\mathbb{E}[X_j^3]|}{6B_n^5} + \frac{t^6 \sum_{j=1}^n |\mathbb{E}[X_j^3]|^2}{36B_n^6} + \frac{|t|^7 \sum_{j=1}^n |\mathbb{E}[X_j^3]| \gamma_j}{72B_n^7} \\
&\leq \left(\frac{|t|^5}{6} \left(\frac{K_{4,n}}{n}\right)^{5/4} + \frac{t^6}{36} \left(\frac{K_{4,n}}{n}\right)^{3/2} + \frac{|t|^7}{72} \left(\frac{K_{4,n}}{n}\right)^{7/4} \right) \mathbb{1}\left\{\exists i \in \{1, ..., n\} : \mathbb{E}[X_i^3] \neq 0\right\} \\
&=: U_{1,2,n}(t).
\end{aligned} \tag{49}$$

Moreover, we have $\sum_{j=1}^n U_{j,n}(t)^2 \leq \frac{t^4 K_{4,n}}{n} P_{1,n}(\varepsilon)$ under our conditions on ε and t. Combining Equation (47), the decomposition (48) and the previous three bounds, and grouping similar terms together, we conclude that for every $\varepsilon \in (0, 1/3)$ and t such that $|t| \leq \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$,

$$\begin{aligned}
&\left| f_{S_n}(t) - e^{-\frac{t^2}{2}} \left(1 - \frac{it^3 \lambda_{3,n}}{6\sqrt{n}}\right) \right| \\
&\leq e^{-t^2/2} \left\{ \frac{t^4 K_{4,n}}{8n} \left(\frac{1}{3} + \frac{1}{(1-3\varepsilon)^2}\right) + \frac{e_{1,n}(\varepsilon)|t|^6 |\lambda_{3,n}|^2}{72n} + \frac{U_{1,1,n}(t) + U_{1,2,n}(t)}{2(1-3\varepsilon)^2} \right. \\
&\left. \quad + e_1(\varepsilon) \left(\frac{t^8 K_{4,n}^2}{2n^2} \left(\frac{1}{24} + \frac{P_{1,n}(\varepsilon)}{2(1-3\varepsilon)^2}\right)^2 + \frac{|t|^7 |\lambda_{3,n}| K_{4,n}}{6n^{3/2}} \left(\frac{1}{24} + \frac{P_{1,n}(\varepsilon)}{2(1-3\varepsilon)^2}\right) \right) \right\},
\end{aligned}$$

where $e_1(\varepsilon) := \exp\left(\varepsilon^2 \left(\frac{1}{6} + \frac{2P_{1,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)\right)$. Combining this with the definition of $R_{1,n}(t, \varepsilon)$ finishes the proof. □

6.3 Control of the residual term in an Edgeworth expansion under Assumption 2.2

Lemma 6.2 can be improved in the *i.i.d.* framework. To do so, we introduce analogues of $R_{1,n}(t,\varepsilon)$, $P_{1,n}(\varepsilon)$, $e_{2,n}(\varepsilon)$ and $U_{1,2,n}(t)$ defined by

$$\begin{aligned}
R_{2,n}(t,\varepsilon) &:= \frac{U_{2,2,n}(t)}{2(1-3\varepsilon)^2} + e_{2,n}(\varepsilon)\left(\frac{t^8}{8n^2}\left(\frac{K_{4,n}}{12} + \frac{P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)^2 + \frac{|t|^7|\lambda_{3,n}|}{12n^{3/2}}\left(\frac{K_{4,n}}{12} + \frac{P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)\right),\\
P_{2,n}(\varepsilon) &:= \frac{1}{4} + \frac{P_{3,n}(\varepsilon)}{576},\\
P_{3,n}(\varepsilon) &:= \frac{96\sqrt{2\varepsilon}|\lambda_{3,n}|}{(K_{4,n}^{1/4}n^{1/4})} + 48\varepsilon\left(\frac{K_{4,n}}{n}\right)^{1/2} + \frac{32\varepsilon\lambda_{3,n}^2}{(K_{4,n}n)^{1/2}} + \frac{16\sqrt{2}K_{4,n}^{1/4}|\lambda_{3,n}|\varepsilon^{3/2}}{n^{3/4}} + \frac{4\varepsilon^2 K_{4,n}}{n},\\
e_{2,n}(\varepsilon) &:= \exp\left(\varepsilon^2\left(\frac{1}{6} + \frac{2P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)\right),\\
U_{2,2,n}(t) &:= \frac{|t|^5|\lambda_{3,n}|}{6n^{3/2}} + \frac{t^6 K_{4,n}}{24n^2} + \frac{t^6\lambda_{3,n}^2}{36n^2} + \frac{|t|^7 K_{4,n}|\lambda_{3,n}|}{72n^{5/2}} + \frac{t^8 K_{4,n}^2}{576n^3}.
\end{aligned}$$

Note that

$$e_{2,n}(\varepsilon) = e^{\varepsilon^2/6} \times \exp\left(\frac{2\varepsilon^2 P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right) = e_3(\varepsilon)\exp\left(\frac{2\varepsilon^2 P_{3,n}(\varepsilon)}{576(1-3\varepsilon)^2}\right),$$

where $e_3(\varepsilon) := e^{\varepsilon^2/6+\varepsilon^2/(2(1-3\varepsilon))^2}$.

Lemma 6.3. *Under Assumption 2.2, for every $\varepsilon \in (0,1/3)$ and t such that $|t| \le \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$,*

$$\left|f_{S_n}(t) - e^{-\frac{t^2}{2}}\left(1 - \frac{it^3\lambda_{3,n}}{6\sqrt{n}}\right)\right| \le e^{-t^2/2}\left\{\frac{t^4 K_{4,n}}{8n}\left(\frac{1}{3} + \frac{1}{(1-3\varepsilon)^2}\right) + \frac{e_{2,n}(\varepsilon)|t|^6|\lambda_{3,n}|^2}{72n} + R_{2,n}(t,\varepsilon)\right\}.$$

Proof of Lemma 6.3: This proof is very similar to that of Lemma 6.2. We note that $B_n = \sigma\sqrt{n}$. As before, using two Taylor-Lagrange expansions successively, we can write that for every $\varepsilon \in (0,1/3)$ and t such that $|t| \le \sqrt{2\varepsilon n}/K_{4,n}^{1/4}$

$$\log(f_{P_{X_1}}(t/B_n)) = U_{1,n}(t) - \frac{U_{1,n}(t)^2}{2(1+\theta_{2,n}(t)U_{1,n}(t))^2},$$

where

$$U_{1,n}(t) := -\frac{t^2}{2n} - \frac{i\lambda_{3,n}t^3}{6n^{3/2}} + \frac{\theta_{1,n}(t)K_{4,n}t^4}{24n^2},$$

and $\theta_{1,n}(t)$ and $\theta_{2,n}(t)$ are two complex numbers with modulus bounded by 1. Using a third Taylor-Lagrange expansion, we can write that for some complex $\theta_{3,n}(t)$ with modulus bounded by $\exp\left(\frac{K_{4,n}t^4}{24n} + \frac{n|U_{1,n}(t)|^2}{2(1-3\varepsilon)^2}\right)$, the following holds

$$\begin{aligned}
f_{S_n}(t) = e^{-\frac{t^2}{2}}\Bigg(1 &- \frac{it^3\lambda_{3,n}}{6\sqrt{n}} + \frac{t^4 K_{4,n}\theta_{1,n}(t)}{24n} - \frac{nU_{1,n}(t)^2}{2(1+\theta_{2,n}(t)U_{1,n}(t))^2}\\
&+ \frac{\theta_{3,n}(t)}{2}\left(-\frac{it^3\lambda_{3,n}}{6\sqrt{n}} + \frac{t^4 K_{4,n}\theta_{1,n}(t)}{24n} - \frac{nU_{1,n}(t)^2}{2(1+\theta_{2,n}(t)U_{1,n}(t))^2}\right)^2\Bigg).
\end{aligned}$$

Using the triangle inequality and its reverse version plus the condition $|t| \le \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}$, we obtain

$$\begin{aligned}
\left|f_{S_n}(t) - e^{-t^2/2}\left(1 - \frac{it^3\lambda_{3,n}}{6\sqrt{n}}\right)\right| &\le e^{-t^2/2}\Bigg\{\frac{t^4 K_{4,n}}{24n} + \frac{nU_{1,n}(t)^2}{2(1-3\varepsilon)^2}\\
&+ \frac{1}{2}\exp\left(\frac{\varepsilon^2}{6} + \frac{n|U_{1,n}(t)|^2}{2(1-3\varepsilon)^2}\right) \times \left(\frac{|t|^3|\lambda_{3,n}|}{6\sqrt{n}} + \frac{t^4 K_{4,n}}{24n} + \frac{nU_{1,n}(t)^2}{2(1-3\varepsilon)^2}\right)^2\Bigg\}. \quad (50)
\end{aligned}$$

We can decompose $nU_{1,n}(t)^2$ as

$$
\begin{aligned}
nU_{1,n}(t)^2 &= \frac{t^4}{4n} + \underbrace{\frac{|t|^5|\lambda_{3,n}|}{6n^{3/2}} + \frac{t^6K_{4,n}}{24n^2} + \frac{t^6\lambda_{3,n}^2}{36n^2} + \frac{|t|^7K_{4,n}|\lambda_{3,n}|}{72n^{5/2}} + \frac{t^8K_{4,n}^2}{576n^3}}_{=U_{2,2,n}(t)} \\
&\leq \frac{t^4P_{2,n}(\varepsilon)}{n},
\end{aligned} \tag{51}
$$

Combining Equations (50) and (51) and grouping terms, we conclude that for every $\varepsilon \in (0, 1/3)$ and t such that $|t| \leq \sqrt{2\varepsilon}(n/K)^{1/4}$,

$$
\begin{aligned}
\left| f_{S_n}(t) - e^{-\frac{t^2}{2}} \left(1 - \frac{it^3\lambda_{3,n}}{6\sqrt{n}}\right)\right| &\leq e^{-t^2/2}\Bigg\{ \frac{t^4K_{4,n}}{8n}\left(\frac{1}{3} + \frac{1}{(1-3\varepsilon)^2}\right) + \frac{e_{2,n}(\varepsilon)t^6\lambda_{3,n}^2}{72n} + \frac{U_{2,2,n}(t)}{2(1-3\varepsilon)^2} \\
&\quad + e_{2,n}(\varepsilon)\left(\frac{t^8}{8n^2}\left(\frac{K_{4,n}}{12} + \frac{P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)^2 + \frac{|t|^7|\lambda_{3,n}|}{12n^{3/2}}\left(\frac{K_{4,n}}{12} + \frac{P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)\right)\Bigg\},
\end{aligned}
$$

where $e_{2,n}(\varepsilon) := \exp\left(\varepsilon^2\left(\frac{1}{6} + \frac{2P_{2,n}(\varepsilon)}{(1-3\varepsilon)^2}\right)\right)$. □

6.4 Two bounds on incomplete Gamma-like integrals

For every $p \geq 1$, $l < m$ and $T > 0$, we define

$$
J_1(p, l, m, T) := \frac{1}{T}\int_l^m |\Psi(u/T)|\, u^p e^{-u^2/2} du \tag{52}
$$

$$
J_2(p, l, m, \widetilde{K}_{3,n}, K_{4,n}, T, n) := \frac{1}{T}\int_v^w |\Psi(u/T)| u^p \exp\left(-\frac{u^2}{2}\left(1 - \frac{2\chi_1|u|\widetilde{K}_{3,n}}{\sqrt{n}} - \sqrt{\frac{K_{4,n}}{n}}\right)\right) du. \tag{53}
$$

We have

$$
\begin{aligned}
J_1(p, l, m, T) &\leq \frac{1.0253}{2\pi}\int_l^m u^{p-1}e^{-u^2/2}du \\
&\leq \frac{1.0253}{2\pi}\int_0^{+\infty} u^{p-1}e^{-u^2/2}du \\
&= \frac{1.0253}{2\pi}\int_0^{+\infty} \sqrt{2v}^{p-1} e^{-v}\frac{dv}{\sqrt{2v}} = \frac{1.0253 \times 2^{p/2-2}\Gamma(p/2)}{\pi}.
\end{aligned} \tag{54}
$$

To obtain (54), we resort to the first inequality in (9) and the change of variable $v = u^2/2$, and we let $\Gamma(a)$ stand for $\Gamma(a, 0)$.

6.4.1 Exponential decay of the term J_2.

Using the first inequality in (9), we get

$$
\begin{aligned}
&J_2(p, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, K_{4,n}, T, n) \\
&\leq \frac{1}{T}\int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{T/\pi} \frac{1.0253\,T}{2\pi u} u^p \exp\left(-\frac{u^2}{2}\left(1 - \frac{2\chi_1|u|\widetilde{K}_{3,n}}{\sqrt{n}} - u^2\sqrt{\frac{K_{4,n}}{n}}\right)\right) du.
\end{aligned}
$$

We now use our choice of T which leads to

$$\begin{aligned}
&J_2(p, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, K_{4,n}, T, n) \\
&\quad\leq \frac{1.0253}{2\pi} \int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{2\sqrt{n}/\widetilde{K}_{3,n}} u^{p-1} \exp\bigg(- \frac{u^2}{2}\bigg(1 - \frac{2\chi_1 |u| \widetilde{K}_{3,n}}{\sqrt{n}} - u^2 \sqrt{\frac{K_{4,n}}{n}}\bigg)\bigg) du, \\
&\quad\leq \frac{1.0253}{2\pi} \int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{2\sqrt{n}/\widetilde{K}_{3,n}} u^{p-1} \exp\bigg(- \frac{u^2}{2}\bigg(1 - 4\chi_1 - \Big(\frac{K_{4,n}}{n}\Big)^{1/2}\bigg)\bigg) du.
\end{aligned}$$

Note that $1 - 4\chi_1 - (K_{4,n}/n)^{1/2} > 1/4$ when $n > K_{4,n}/(0.75 - 4\chi_1)^2$, i.e. as soon as $n \geq 8\, K_{4,n}$. When this is the case, we have

$$\begin{aligned}
&J_2(p, \sqrt{2\varepsilon}(n/K_{4,n})^{1/4}, T/\pi, \widetilde{K}_{3,n}, K_{4,n}, T, n) \\
&\quad\leq \frac{1.0253}{2\pi} \int_{\sqrt{2\varepsilon}(n/K_{4,n})^{1/4}}^{2\sqrt{n}/\widetilde{K}_{3,n}} u^{p-1} \exp\Big(- \frac{u^2}{8}\Big) du \\
&\quad\leq \frac{1.0253}{2\pi} \int_{\varepsilon(n/K_{4,n})^{1/2}/4}^{n/(2\widetilde{K}_{3,n}^2)} 8^{(p-1)/2} v^{(p-1)/2} e^{-v} \sqrt{\frac{2}{v}} dv \\
&\quad= O\bigg(\Gamma\bigg(p/2, \frac{\varepsilon}{4}\Big(\frac{n}{K_{4,n}}\Big)^{1/2}\bigg) - \Gamma\bigg(p/2, \frac{n}{2\widetilde{K}_{3,n}^2}\bigg)\bigg) \\
&\quad= O\bigg(\bigg(\frac{\varepsilon}{4}\Big(\frac{n}{K_{4,n}}\Big)^{1/2}\bigg)^{\frac{p}{2}-1} e^{-\varepsilon(n/K_{4,n})^{1/2}/4}\Big(1 + O(n^{-1/2})\Big) - \bigg(\frac{n}{2\widetilde{K}_{3,n}^2}\bigg)^{\frac{p}{2}-1} e^{-n/(2\widetilde{K}_{3,n}^2)}\big(1 + O(n^{-1})\big)\bigg) \\
&\quad= O(n^{p/4-1/2} e^{-\varepsilon\sqrt{n}/4\sqrt{K_{4,n}}}),
\end{aligned}$$

where we use the change of variable $v = u^2/8$.

6.5 Proof of Proposition 3.3

The assumed integrability condition implies that f_Q is absolutely integrable, and therefore we can apply the inversion formula (Ushakov, 2011, Theorem 1.2.6) so that for any $x \in \mathbb{R}$,

$$q(x) = \int_{-\infty}^{+\infty} r(x,t) dt.$$

where $r(x,t) := \dfrac{1}{2\pi} e^{-itx} f_P(t)$. Note that r is infinitely differentiable with respect to x, and that

$$\left| \frac{\partial r(x,t)}{\partial x^{p-1}} \right| = \left| \frac{1}{2\pi} (-it)^{p-1} e^{-itx} f_P(t) \right| = \frac{1}{2\pi} |t|^{p-1} \big| f_P(t) \big|,$$

which is integrable with respect to t, by assumption. This concludes the proof that q is $(p-1)$ times differentiable, as r is measurable.

References

Abramowitz, M. and I. A. Stegun (1972): *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, vol. 55, National Bureau of Standards, Applied Mathematics Series.

Adell, J. A. and A. Lekuona (2008): "Shortening the distance between Edgeworth and Berry–Esseen in the classical case," *Journal of Statistical Planning and Inference*, 138, 1167 – 1178.

Bentkus, V. (2003): "On the dependence of the Berry–Esseen bound on dimension," *Journal of Statistical Planning and Inference*, 113, 385 – 402.

Bentkus, V. and F. Götze (1996): "The Berry-Esseen bound for student's statistic," *Ann. Probab.*, 24, 491–503.

Berry, A. (1941): "The Accuracy of the Gaussian Approximation to the Sum of Independent Variates," *Transactions of the American Mathematical Society*, 49, 122–136.

Bhattacharya, R. N. and R. Ranga Rao (1976): *Normal Approximation and Asymptotic Expansions*, New York: Wiley.

Boutsikas, M. V. (2011): "Asymptotically optimal Berry–Esseen-type bounds for distributions with an absolutely continuous part," *Journal of Statistical Planning and Inference*, 141, 1250 – 1268.

Chernozhukov, V., D. Chetverikov, and K. Kato (2017): "Central limit theorems and bootstrap in high dimensions," *Ann. Probab.*, 45, 2309–2352.

Cramer, H. (1962): *Random Variables and Probability Distributions*, Cambridge University Press, 2 ed.

de la Peña, V. H., T. Leung Lai, and Q.-M. Shao (2009): *Self-Normalized Processes: Limit Theory and Statistical Applications*, Probability and Its Aplications, Springer-Verlag, Berlin Heidelberg.

Efron, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *Ann. Statist.*, 7, 1–26.

Esseen, C.-G. (1942): "On the Liapunoff limit of error in the theory of probability," *Arkiv för Matematik, Astronomi och Fysik.*

——— (1945): "Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law," *Acta Math.*, 77, 1–125.

Esseen, C. G. (1956): "A moment inequality with an application to the central limit theorem," *Scandinavian Actuarial Journal*, 1956, 160–170.

Gil-Pelaez, J. (1951): "Note on the inversion theorem," *Biometrika*, 38, 481–482.

Goulet, V. (2016): *expint: Exponential Integral and Incomplete Gamma Function*, r package.

Hall, P. (1987): "Edgeworth Expansion for Student's t Statistic Under Minimal Moment Conditions," *Ann. Probab.*, 15, 920–931.

——— (1992): *The bootstrap and Edgeworth expansion*, Springer series in statistics, Springer-Verlag.

Kosorok, M. (2006): *Introduction to Empirical Processes and Semiparametric Inference*, Springer Verlag New York.

Lahiri, S. N. (2003): *Resampling methods for dependent data*, Springer Science & Business Media.

Narasimhan, B., S. G. Johnson, T. Hahn, A. Bouvier, and K. Kiêu (2020): *cubature: Adaptive Multivariate Integration over Hypercubes*, r package version 2.0.4.1.

Pinelis, I. (2011): "Relations between the first four moments," .

Pinelis, I. and R. Molzon (2016): "Optimal-order bounds on the rate of convergence to normality in the multivariate delta method," *Electronic Journal of Statistics*, 10, 1001–1063.

Prawitz, H. (1972): "Limits for a distribution, if the characteristic function is given in a finite domain," *Scandinavian Actuarial Journal*, 1972, 138–154.

——— (1975): "On the remainder in the central limit theorem," *Scandinavian Actuarial Journal*, 1975, 145–156.

Raič, M. (2018): "A multivariate central limit theorem for Lipschitz and smooth test functions," *arXiv preprint arXiv:1812.08268*.

Raič, M. (2019): "A multivariate Berry–Esseen theorem with explicit constants," *Bernoulli*, 25, 2824–2853.

Senatov, V. V. (2011): "On the real accuracy of approximation in the central limit theorem," *Siberian Mathematical Journal*, 52, 19 – 38.

Shevtsova, I. (2010): "Refinement of estimates for the rate of convergence in Lyapunov's theorem," *Dokl. Akad. Nauk*, 435, 26–28.

——— (2011): "On the absolute constants in the Berry-Esseen type inequalities for identically distributed summands," *arXiv preprint arXiv:1111.6554*.

——— (2012): "Moment-type estimates with asymptotically optimal structure for the accuracy of the normal approximation," *Annales Mathematicae et Informaticae*, 39, 241–307.

Ushakov, N. G. (2011): *Selected Topics in Characteristic Functions*, Berlin, Boston: De Gruyter.

Ushakov, N. G. and V. G. Ushakov (1999): "Some inequalities for characteristic functions of densities with bounded variation," *Preprint series. Statistical Research Report http://urn.nb. no/URN: NBN: no-23420*.

van der Vaart, A. (2000): *Asymptotics Statistics*, Cambridge University Press.

van der Vaart, A. and J. Wellner (1996): *Weak Convergence of Empirical Processes: with Applications to Statistics*, Springer-Verlag New York.

Zhilova, M. (2020): "New Edgeworth-type expansions with finite sample guarantees," *arXiv preprint arXiv:2006.03959*.

A Numerical computations

```
library("cubature")
library("expint")
```

A.1 Optimization of the components of Ω_1

A.1.1 Bound on $I_{1,1}$

```
# First term of Omega_1
toIntegrate_1_1 <- function(t, capitalT) {
  indic = as.numeric(t <= 1)
  module = sqrt((1 - t)^2 * indic +
                ( ( (1-t) * cos(pi * t) / sin(pi * t) + 1/pi) * indic
                 - 1/(pi * t) )^2 )
  expo = exp( - (t * capitalT)^2 / 2)

  return (module * expo)
}

# Function to optimize
Integral_1_1 <- function (capitalT){
  integral = -hcubature(f = toIntegrate_1_1, lower = 0, upper = 1/pi,
                        capitalT = abs(capitalT))$integral
  return(as.numeric(capitalT * integral))
}

# Computation of the maximum of I_{1,1}
result_1_1 =
  optim(f = Integral_1_1, par = 1,
        method = "L-BFGS-B", control = list(trace = 1))
```

```
## iter   10 value -1.236798
## iter   20 value -1.252322
## iter   30 value -1.253254
## final  value -1.253279
## converged
```

```
print(- result_1_1$value)
```

```
## [1] 1.253279
```

A.1.2 Bound on $I_{1,2}$

```r
# Second term of Omega_1
toIntegrate_1_2 <- function(t, capitalT) {
  indic = as.numeric(t <= 1)
  module = sqrt((1 - t)^2 * indic +
                  ( ( (1-t) * cos(pi * t) / sin(pi * t) + 1/pi) * indic
                   - 1/(pi * t) )^2 )
  expo = exp( - (t * capitalT)^2 / 2) * t^3 / 6

  return (module * expo)
}

# Function to optimize
Integral_1_2 <- function (capitalT){
  integral = -hcubature(f = toIntegrate_1_2, lower = 0, upper = 1/pi,
                        capitalT = abs(capitalT))$integral
  return(as.numeric(capitalT^4 * integral))
}

# Computation of the maximum of I_{1,2}
result_1_2 =
  optim(f = Integral_1_2, par = 1,
        method = "L-BFGS-B", control = list(trace = 1))
```

```
## iter   10 value -0.332670
## iter   20 value -0.333293
## final  value -0.333311
## converged
```

```r
print(- result_1_2$value)
```

```
## [1] 0.3333106
```

A.1.3 Bound on $I_{1,3}$

```r
Integral_1_3 <- function (capitalT){
  value = - capitalT^4 * gammainc(0, capitalT^2 / (2 * pi^2) ) / (2*pi)
  return(value)
}

# Computation of the maximum of I_{1,3}
result_1_3 =
  optim(f = Integral_1_3, par = 1,
        method = "L-BFGS-B", control = list(trace = 1))
```

```
## final  value -14.196144
## converged
```

```r
print(- result_1_3$value)
```

```
## [1] 14.19614
```

A.2 Optimization of the components of Ω_2

A.2.1 Computation of t_1^* and χ_1

```r
t0 = 1/pi

result_t1 =
  optim(
    fn = function(theta){return(abs(
      theta^2 + 2 * theta * sin(theta) + 6 * (cos(theta) - 1 ) ))},
    lower = 3.8, upper = 4.2, par = 4,
    method = "L-BFGS-B", control = list(trace = 1))
```

```
## final  value 0.000005
## stopped after 3 iterations
```

```r
t1star = result_t1$par / (2*pi)
print(t1star)
```

```
## [1] 0.6359664
```

```r
result_chi1 =
  optim(f = function(x){return( - abs(cos(x)-1 + x^2/2) / x^3)},
        par = 1, lower = 10^(-5),
        method = "L-BFGS-B")
chi1 = - result_chi1$value
print(chi1)
```

```
## [1] 0.09916191
```

A.2.2 Bound on $I_{2,1}$

```r
toIntegrate_2_1 <- function(t, capitalT) {
  indic = as.numeric(t <= 1)
  module = sqrt((1 - t)^2 * indic +
                  ( ( (1-t) * cos(pi * t) / sin(pi * t) + 1/pi)
                * indic )^2 )
  expo = exp( - 0.5 * (t * capitalT)^2 * ( 1 - 4*pi*chi1*abs(t) ) )
  return (module * expo)
}

# Function to optimize
```

```r
Integral_2_1 <- function (capitalT){
  integral = - hcubature(f = toIntegrate_2_1, lower = t0, upper = t1star,
                         capitalT = abs(capitalT))$integral
  return(as.numeric(capitalT^4 * integral))
}

# Computation of the maximum of I_{2,1}
result_2_1 =
  optim(f = Integral_2_1, par = 1, lower = 0,
        method = "L-BFGS-B", control = list(trace = 1))
```

```
## final  value -67.041372
## converged
```

```r
print(- result_2_1$value)
```

```
## [1] 67.04137
```

A.2.3 Bound on $I_{2,2}$

```r
toIntegrate_2_2 <- function(t, capitalT) {
  indic = as.numeric(t <= 1)
  module = sqrt((1 - t)^2 * indic +
                 ( ( (1-t) * cos(pi * t) / sin(pi * t) + 1/pi)
               * indic )^2 )
  expo = exp( - capitalT^2 * ( 1 - cos(2*pi*t) ) / (4 * pi^2) )
  return (module * expo)
}

# Function to optimize
Integral_2_2 <- function (capitalT){
  integral = - hcubature(f = toIntegrate_2_2, lower = t1star, upper = 1,
                         capitalT = abs(capitalT))$integral
  return(as.numeric(capitalT^2 * integral))
}

# Computation of the maximum of I_{2,2}
result_2_2 =
  optim(f = Integral_2_2, par = 1, lower = 0,
        method = "L-BFGS-B", control = list(trace = 1))
```

```
## final  value -1.218606
## converged
```

```r
print(- result_2_2$value)
```

```
## [1] 1.218606
```