Asymptotically optimal Berry-Esseen-type bounds for distributions with an absolutely continuous part

Michael V. Boutsikas
Department of Statistics and Insurance Science, University of Piraeus, Greece

Abstract

Recursive and closed form upper bounds are offered for the *Kolmogorov* and the *total variation distance* between the standard normal distribution and the distribution of a standardized sum of n independent and identically distributed random variables. The method employed is a modification of the method of compositions along with Zolotarev's ideal metric. The approximation error in the CLT obtained vanishes at a rate $O(n^{-k/2+1})$, provided that the common distribution of the summands possesses an *absolutely continuous part*, and shares the same $k-1$ ($k \geq 3$) first moments with the standard normal distribution. Moreover, for the first time, these new uniform Berry-Esseen-type bounds are *asymptotically optimal*, that is, the ratio of the true distance to the respective bound converges to unity for a large class of distributions of the summands. Thus, apart from the correct rate, the proposed error estimates incorporate an optimal asymptotic constant (factor). Finally, three illustrative examples are presented along with numerical comparisons revealing that the new bounds are sharp enough even to be used in practical statistical applications.

Abbreviated Title: Optimal Berry-Esseen-type bounds.

Key words and phrases: Central limit theorem, Berry-Esseen theorem, Edgeworth expansions to the CLT, Rate of convergence, Kolmogorov and total variation distance.

AMS 2000 subject classification: Primary: 60F05
Secondary: 60G50; 60E15; 62E17; 62G20

1 Introduction

One of the most remarkable results in the field of probability and statistics is the Central Limit Theorem (CLT). Important role for the deeper investigation, as well as for the practical use of this celebrated asymptotic result, plays the establishment of the associated rates of convergence. These are usually expressed through asymptotic (Edgeworth) expansions or explicit (Berry-Esseen-type) bounds. The former provide better asymptotic accuracy for the approximation error, but they incorporate terms in the expansion which do not possess an explicit closed form (for example, are expressed via $o(n^{-p})$). Therefore, Edgeworth expansions are practically useless for fixed values of the number n of summands. On the other hand, Berry-Esseen-type bounds offer explicit approximation errors for every value of n, however usually these are very crude. The purpose of the present paper is to unify the advantages of the two aforementioned approaches. That is, to develop Berry-Esseen-type bounds which, for the first time, exhibit the same asymptotic accuracy as the Edgeworth expansions, under certain weak assumptions.

Let $X, X_1, X_2, ...$ be independent and identically distributed (i.i.d.) random variables (r.v.'s) with $\mathbb{E}X = 0, \mathbb{E}X^2 = 1$ and $\mathbb{E}|X|^3 < \infty$. Denote by F_n, the cumulative distribution function (c.d.f.) of the standardized sum $n^{-1/2} \sum_{i=1}^n X_i$ and by Φ the c.d.f. of the standard normal law. The classical uniform Berry-Esseen theorem (see Berry (1941) and Esseen (1942)) states that

$$\sup_x |F_n(x) - \Phi(x)| \leq C \frac{\mathbb{E}|X|^3}{\sqrt{n}}, \tag{1}$$

for some constant $C > 0$. The smallest known value of C fulfilling the above is 0.7056 (Shevtsova (2007)). If we further assume that X follows a non-lattice distribution then the one-term Edgeworth expansion of the c.d.f. F_n of $n^{-1/2} \sum_{i=1}^n X_i$ is (for example, see Petrov (1995)),

$$F_n(x) = \Phi(x) + \frac{\mathbb{E}X^3 (1 - x^2) e^{-x^2/2}}{3! \sqrt{2\pi n}} + o(\frac{1}{\sqrt{n}}), \tag{2}$$

uniformly in x. The above asymptotic estimate is better than the one implied by the Berry-Esseen theorem (note that, $\sup_x \left|1 - x^2\right| e^{-x^2/2} = 1$). However, contrary to (1), the asymptotic expansion (2) is useless in daily statistical practices where we are interested in estimates of the difference $F_n(x) - \Phi(x)$ for a fixed value of n. For example, relation (2) cannot be exploited when we want to determine the number n of summands needed for achieving a prescribed degree of approximation accuracy. A comparison between Berry-Esseen-type results and Edgeworth expansions can be found in Adell and Lekuona (2008); see also Seoh and Hallin (1997).

Traditionally, the standard way for establishing closed form error estimates in the CLT is through Fourier analytic methods (for example, see Petrov (1995)) while several other alternative methods were developed in the last decades. The most important among them are: (i) the approach based on the elegant theory of probability metrics developed by V.M. Zolotarev and his colleagues and students (for example, see Zolotarev (1997), Senatov (1998) or Rachev (1991) and the references therein) and (ii) the celebrated Stein's method, originally developed by Stein (1970) and extended by many authors in various distribution approximation directions, including the case of dependent summands (for example see Barbour and Hall (1984), Chen and Shao (2001), Barbour and Chen (2005) and the references therein). We also mention the approach through covariance kernels or $w-$functions (for example see Cacoullos, Papadatos and Papathanasiou (1997)) which can be considered that is related to Stein's method. But, despite the great interest and the vast literature concerning Berry-Esseen theorems, to the best of our knowledge, until now there do not seem to exist in the literature asymptotically optimal Berry-Esseen-type bounds. Note that, an upper bound for some distance can be considered asymptotically optimal when the ratio of the bound to the true distance converges to unity as $n \to \infty$. That is, the bound not only possesses the correct rate of convergence but, in addition, it incorporates an asymptotically optimal constant (factor). An attempt towards this goal was recently made by Adell and Lekuona (2008) who presented "near optimal" Berry-Esseen-type bounds for the uniform distance between F_n and Φ. Their bounds were obtained by substituting Esseen's smoothing inequality by an inequality involving the second modulus of continuity.

In the present paper, motivated by the comparison between Edgeworth expansions and Berry-Esseen bounds, we develop recursive and closed form upper bounds for the *uniform* (or *Kolmogorov*) and the *total variation distance* between the distribution of $n^{-1/2} \sum_{i=1}^n X_i$ and Φ. These bounds are proven to be *asymptotically optimal* for a large class of distributions of the summands. More specifically, suppose that $\mathbb{E}X^i = \mathbb{E}Z^i, i = 1, 2, ..., k-1$ for some integer $k \geq 3$, where Z is a r.v. following a standard normal law. Henceforth, the standard normal law will be denoted by $\mathcal{N}$ and the law of a r.v. X will be denoted by $\mathcal{L}X$. If $\mathcal{L}X$ has an absolutely continuous part, we derive an explicit bound $\mathbf{R}^*_{k,n}$ for the uniform distance ($\mathbf{R}^*_{k,n} = \mathbf{R}^*_{k,h,m}, n = hm$ in Theorem 12 and Corollary 15 below) such that

$$\sup_x |F_n(x) - \Phi(x)| \leq \mathbf{R}^*_{k,n} \sim \frac{\left\|\varphi^{(k-1)}\right\|_\infty}{n^{k/2-1}} \boldsymbol{\zeta}_k(\mathcal{L}X, \mathcal{N}) = \frac{\left\|\varphi^{(k-1)}\right\|_\infty}{n^{k/2-1}} \frac{\left|\mathbb{E}X^k - \mathbb{E}Z^k\right|}{k!}. \tag{3}$$

The quantity $\left\|\varphi^{(i)}\right\|_\infty$ denotes the L_∞-norm of the ith-order derivative of the probability density function (p.d.f.) φ of $\mathcal{N}$, and $\boldsymbol{\zeta}_k$ denotes the Zolotarev's ideal metric of order k. As usual, the notation $a_n \sim b_n$ implies that $a_n/b_n \to 1$ (or whenever $a_n = 0$ then $b_n = 0$ and vice versa). The last equality in (3) is valid if, in addition, $\mathcal{L}X$ and $\mathcal{N}$ are k-convex ordered (see Boutsikas and Vaggelatou (2002)). This stochastic ordering can usually be easily checked, and seems to be satisfied by many known distributions (see Remark 2 below). When $k = 3$ (that is, $\mathbb{E}X = 0, \mathbb{E}X^2 = 1$), a condition that is always valid after standardization of the summands, relation (3) reduces to

$$\sup_x |F_n(x) - \Phi(x)| \leq \mathbf{R}^*_{3,n} \sim \frac{\boldsymbol{\zeta}_3(\mathcal{L}X, \mathcal{N})}{\sqrt{2\pi n}} = \frac{\left|\mathbb{E}X^3\right|}{3!\sqrt{2\pi n}}.$$

The last equality is again valid when $\mathcal{L}X$ and $\mathcal{N}$ are 3-convex ordered (for example, when $\mathcal{L}X$ is the gamma distribution; see example (a) in Section 5). Remarkably, the above asymptotic form of the bound $\mathbf{R}_{3,n}$ is consistent to that implied by the Edgeworth expansions (see (2) above), and $\mathbf{R}_{3,n}$, or more generally $\mathbf{R}_{k,n}$, is indeed proven to be asymptotically optimal (see Corollary 15 below).

It is also remarkable that, when $\mathcal{L}X$ share higher order moments with $\mathcal{N}$ (that is, $k > 3$), the vanishing rate of $\mathbf{R}_{k,n}$ is amplified. If, for example, $\mathcal{L}X$ is symmetric, then $\mathbb{E}X^3 = 0$ and therefore, after standardization of X, relation (3) is valid for some k at least equal to 4 (see example (b) in Section 5 regarding the uniform distribution).

Analogous results are proven for the *total variation distance* $\mathbf{d}_n$ between the law of the standardized sum of n X_i's and $\mathcal{N}$ (see Theorem 5 and Corollary 8).

Finally, it is worth pointing out a shortcoming of our bounds. They are applicable only when the distribution of X_i's is already relatively close to standard Normal. Fortunately, we can always overcome this situation as will be explained in Remark 3 below and practiced in the examples section.

For the proof of our main results we proceed to a refinement and an adaptation to the continuous case of the method employed by Boutsikas and Vaggelatou (2009) for compound Poisson approximations. This approach is based on the method of compositions and the theory of probability metrics, exploiting certain smoothing inequalities satisfied by Zolotarev's ideal metric. Hence, it can be considered as a modification or refinement of the method described by Senatov (1998, Section 4.1); see also Rachev (1991, Section 14.3) and the references therein.

The paper is organized as follows. In the Section 2 we present all necessary notations and auxiliary results needed for the exposition of our main outcomes. In the main results section (Section 3) we offer recursive and closed form upper bounds for the Kolmogorov and the total variation distance of interest, along with some results guaranteeing asymptotic optimality. The proofs of the main results, together with all necessary lemmas, are postponed to the next Section 4. The applicability and the performance of our main results is investigated by way of three illustrative applications in Section 5.

2 Preliminaries - Notations

(i) Probability metrics. The *uniform* or *Kolmogorov distance* and the *total variation distance* between the distributions $\mathcal{L}X$ and $\mathcal{L}Y$ of two r.v.'s X and Y are defined respectively by $\boldsymbol{\rho}(\mathcal{L}X, \mathcal{L}Y) = \sup_{x\in\mathbb{R}} |F_X(x) - F_Y(x)|$ and

$$\mathbf{d}\left(\mathcal{L}X, \mathcal{L}Y\right) = \sup_{A\in\mathcal{B}(\mathbb{R})} \left|P\left(X \in A\right) - P\left(Y \in A\right)\right| = \frac{1}{2}\int_{\mathbb{R}} |f_X(x) - f_Y(x)|dx \tag{4}$$

where, as usual, F_X denotes the c.d.f. of the r.v. X. The second equality above holds true when the r.v.'s X, Y possess Lebesgue densities f_X and f_Y respectively. Moreover, $\mathbf{d}(\mathcal{L}f(X), \mathcal{L}f(Y)) \le \mathbf{d}(\mathcal{L}X, \mathcal{L}Y)$ for *every* measurable f and thus if $\mathcal{L}X \approx \mathcal{L}Y$ with respect to $\mathbf{d}$, then, with the same accuracy, $\mathcal{L}f(X) \approx \mathcal{L}f(Y)$. The *Zolotarev's ideal metric* or *total variation distance of order* $s \in \mathbb{N}$ (see Zolotarev (1983)) is defined by

$$\boldsymbol{\zeta}_s(\mathcal{L}X, \mathcal{L}Y) = \tfrac{1}{(s-1)!}\int_{-\infty}^{\infty} \left|\mathbb{E}(X-t)_+^{s-1} - \mathbb{E}(Y-t)_+^{s-1}\right| dt,$$

(where $y_+^s \overset{\texttt{def}}{=} (\max\{y, 0\})^s$) which can be finite only when $\mathbb{E}X^j = \mathbb{E}Y^j$, $j = 1, 2, ..., s-1$. The latter condition assures also that $\boldsymbol{\zeta}_s(\mathcal{L}X, \mathcal{L}Y) \le \left(\mathbb{E}|X|^s + \mathbb{E}|Y|^s\right)/s!$. Throughout, for simplicity, we shall allow an abuse in the notation and write X instead of $\mathcal{L}X$ in $\boldsymbol{\rho}$,$\mathbf{d}$ or $\boldsymbol{\zeta}_s$. It is worth mentioning that the metrics $\mathbf{d}, \boldsymbol{\rho}$ and $\boldsymbol{\zeta}_s$ are *ideal metrics* of order $0, 0$ and s respectively and hence, they possess the *regularity,* the *homogeneity* and the *subadditivity* property for independent summands (for example, see Rachev (1991), p.264). Thus, for example,

$$\boldsymbol{\zeta}_s(c\sum_{i=1}^{n} X_i, c\sum_{i=1}^{n} Y_i) \le c^s \sum_{i=1}^{n} \boldsymbol{\zeta}_s(X_i, Y_i), \quad n, s \in \mathbb{N},\ c > 0, \tag{5}$$

for independent X_i's and Y_i's. For a comprehensive exposition of probability metrics and their properties see Zolotarev (1997), Senatov (1998) or Rachev (1991) and the references therein.

(ii) Derivative norms of the standard normal density. Denote by φ the standard normal density and by

$$\left\|\varphi^{(k)}\right\| = \int_{-\infty}^{\infty} \left|\varphi^{(k)}(x)\right| dx, \quad \text{and} \quad \left\|\varphi^{(k)}\right\|_\infty = \sup_x \left|\varphi^{(k)}(x)\right|, \quad k = 0, 1, ...,$$

the L_1 and the L_∞ norm respectively of $\varphi^{(k)}$, the kth derivative of φ. It is well known that

$$\varphi^{(k)}(x) = (-1)^k \frac{e^{-x^2/2}}{\sqrt{2\pi}} \mathcal{H}_k(x), \tag{6}$$

where

$$\mathcal{H}_k(x) = (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2} = k! \sum_{i=0}^{[k/2]} \frac{(-1)^i x^{k-2i}}{i!(k-2i)!2^i}, \tag{7}$$

denotes the Chebyshev-Hermite polynomial of degree k (e.g. see Petrov (1995)). In particular, $\mathcal{H}_0(x) = 1$, $\mathcal{H}_1(x) = x$, $\mathcal{H}_2(x) = x^2 - 1$, $\mathcal{H}_3(x) = x^3 - 3x$, $\mathcal{H}_4(x) = x^4 - 6x^2 + 3$, etcetera. It is easy to verify that,

$$\left\|\varphi^{(k-1)}\right\|_\infty = \sup_x \left|\int_{-\infty}^{x} \varphi^{(k)}(t)dt\right| \le \int_{-\infty}^{\infty} \left|\varphi^{(k)}(t)\right| dt = \left\|\varphi^{(k)}\right\|.$$

If φ_{μ,σ^2} denotes the density of the normal distribution with mean μ, variance σ^2, then

$$\left\|\varphi^{(k)}_{\mu,\sigma^2}\right\| = \frac{1}{\sigma^k}\left\|\varphi^{(k)}\right\|, \quad \left\|\varphi^{(k)}_{\mu,\sigma^2}\right\|_\infty = \frac{1}{\sigma^{k+1}}\left\|\varphi^{(k)}\right\|_\infty . \tag{8}$$

(iii) s-convex ordering. A r.v. X is said to be s-convex smaller than a r.v. Y (denoted by $X \leq_{s-cx} Y$, or $\mathcal{L}X \leq_{s-cx} \mathcal{L}Y$, $s \in \mathbb{N}$) if $\mathbb{E}\phi(X) \leq \mathbb{E}\phi(Y)$ for all regular s-convex functions ϕ, (that is, all $\phi : \mathbb{R} \to \mathbb{R}$ such that their sth derivative, $\phi^{(s)}$, exists, is nonnegative and continuous) for which the expectations exist. Denuit, Lefèvre and Shaked (1998) offered the following characterization of the s-convex orders.

Proposition 1 *If X, Y are two real valued r.v.'s such that $\mathbb{E}|X|^{s-1}, \mathbb{E}|Y|^{s-1} < \infty$, then*

$$X \leq_{s-cx} Y \Leftrightarrow \left\{ \begin{array}{c} \mathbb{E}X^k = \mathbb{E}Y^k,\ k = 1, 2, ..., s-1, \ \text{ and} \\ \mathbb{E}(X-t)_+^{s-1} \leq \mathbb{E}(Y-t)_+^{s-1} \ \text{ for all } t \in \mathbb{R}. \end{array} \right.$$

Moreover, if $X \leq_{s-cx} Y$ then $\mathbb{E}X^s \leq \mathbb{E}Y^s$ (in this case, $\mathbb{E}X^s = \mathbb{E}Y^s$ implies that $X =_d Y$). Denuit, Lefèvre and Shaked (1998) generalized some well known properties of the usual convex orders (see for example, Shaked and Shanthikumar (1994)) in the s-convex orders. In particular, they proved that the s-convex orders are closed under mixtures, convolutions, compounding and they are preserved under limits. It is remarkable that, when $X \leq_{s-cx} Y$, the distance $\boldsymbol{\zeta}_s(X, Y)$ takes on the following simple form (see Boutsikas and Vaggelatou (2002)),

$$\boldsymbol{\zeta}_s(X, Y) = \frac{\mathbb{E}Y^s - \mathbb{E}X^s}{s!}, \ s = 1, 2, ... \tag{9}$$

3 Main Results

In what follows we assume that $X, X_1, X_2, ...$ are i.i.d. r.v.'s and we shall use the notations,

$$\mathbf{d}_s = \mathbf{d}\left(\tfrac{1}{\sqrt{s}}\sum_{i=1}^s X_i, \mathcal{N}\right), \quad \boldsymbol{\rho}_s = \boldsymbol{\rho}\left(\tfrac{1}{\sqrt{s}}\sum_{i=1}^s X_i, \mathcal{N}\right), \quad \boldsymbol{\zeta}_s = \boldsymbol{\zeta}_s(X, \mathcal{N}), \quad s = 1, 2, ... ,$$

($\mathbf{d}_0 = \boldsymbol{\rho}_0 = 0$), where $\mathcal{N}$ denotes the standard normal distribution. We shall also denote by μ_i the ith moment of $\mathcal{N}$. We remind that $\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 0$, $\mu_4 = 3$, and generally, $\mu_{2i} = \frac{(2i)!}{2^i i!}, \mu_{2i-1} = 0, i = 1, 2, ...$. By $\mathcal{M}_k$ we shall denote the class of all distributions with the same first $k-1$ moments as $\mathcal{N}$, i.e. $\mathcal{L}X \in \mathcal{M}_k$ when $\mathbb{E}X^i = \mu_i, i = 1, 2, ..., k-1$.

3.1 Bounds for the total variation distance

We begin with results concerning the total variation distance which, apart from their independent interest, are essential for the establishment of the relevant results regarding the Kolmogorov distance. The first theorem we present, furnishes a recursive inequality satisfied by $\mathbf{d}_n$ when the summands possess the same first $k-1$ moments as $\mathcal{N}$. Its proof employs Lindeberg's decomposition and smoothing inequalities of Lemmas 16, 17 (see Section 4).

Theorem 2 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \geq 1$, then*

$$\mathbf{d}_n \leq \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k \sum_{j=2}^{n-1} \frac{\mathbf{d}_{j-1}}{(n-j)^{k/2}} + \frac{n\left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k}{2(n-1)^{k/2}} + 2\mathbf{d}_1\mathbf{d}_{n-1}, \quad n \geq 2. \tag{10}$$

The above recursive scheme produces an upper bound for $\mathbf{d}_n$ in terms of $\mathbf{d}_1 = \mathbf{d}(X,\mathcal{N})$ and $\boldsymbol{\zeta}_k = \boldsymbol{\zeta}_k(X,\mathcal{N})$. In practice, if the explicit values of $\mathbf{d}_1$ or $\boldsymbol{\zeta}_k$ cannot be analytically evaluated in order to initiate the recursions, they can be upper bounded (for example, via the approach proposed by Cacoullos, Papadatos and Papathanasiou (1997) for $\mathbf{d}_1$) or they can be numerically evaluated.

For small k, the explicit value of $\left\|\varphi^{(k)}\right\|$ appearing in (10) can be derived using standard but lengthy and tedious calculus techniques. For large k it can always be accurately approximated via numerical integration. For example, after lengthy algebraic manipulations, we find

$$\begin{aligned}\left\|\varphi^{(3)}\right\| &= \tfrac{2+8e^{-3/2}}{\sqrt{2\pi}} \approx 1.510013, \quad \left\|\varphi^{(4)}\right\| = 4e^{\frac{-3-\sqrt{6}}{2}}\tfrac{e^{\sqrt{6}}\sqrt{3-\sqrt{6}}+\sqrt{3+\sqrt{6}}}{\sqrt{\pi/3}} \approx 2.800600,\\ \left\|\varphi^{(5)}\right\| &= \frac{3\sqrt{2}+16e^{-\frac{5+\sqrt{10}}{2}}\left(\sqrt{2}+\sqrt{5}-\left(\sqrt{2}-\sqrt{5}\right)e^{\sqrt{10}}\right)}{\sqrt{\pi}} \approx 5.910086. \qquad (11)\end{aligned}$$

We can also approximately find $\left\|\varphi^{(6)}\right\| \approx 13.8156$, $\left\|\varphi^{(7)}\right\| \approx 35.1479$, $\left\|\varphi^{(8)}\right\| \approx 96.11237$.

It is worth stressing that the recursive bound derived from (10) converges to zero only when $k \geq 3$ and the distribution $\mathcal{L}X$ of the summands is relatively close to $\mathcal{N}$, that is, when $\mathbf{d}_1 = \mathbf{d}(X,\mathcal{N})$ and $\boldsymbol{\zeta}_k = \boldsymbol{\zeta}_k(X,\mathcal{N})$ are relatively small. For example, when $k = 3$, numerical experimentation shows that the recursive bound of Theorem 2 vanishes when the point $(\mathbf{d}_1, \boldsymbol{\zeta}_3)$ of $\mathbb{R}^2$ lies roughly within the triangle with edges $(0,0), (0.5,0), (0,0.2)$. In the following definition we offer an explicit condition which guarantees the aforementioned distributional closeness. We shall be invoking condition $A(k)$ in order to deduce vanishing upper bounds for $\mathbf{d}_n$ (see Theorems 4, 5) or $\boldsymbol{\rho}_n$ (see Theorem 12).

Definition 3 *The distribution $\mathcal{L}X$ of a r.v. X, with $\mathcal{L}X \in \mathcal{M}_k$ for some $k \geq 3$, satisfies the condition $A(k)$ when*

$$\mathbf{d}_1 = \mathbf{d}(X,\mathcal{N}) < 2^{-k/2-1} \quad \textit{and} \quad \varepsilon_k \overset{\texttt{def}}{=} \frac{2^{k/2+1}\mathbf{d}_1}{1+a_k\mathbf{d}_1} + \delta_k \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k < 1,$$

where

$$\delta_k \overset{\texttt{def}}{=} \begin{cases} (\frac{5}{3})^{3/2} + \frac{1}{2}(\frac{5}{4})^{3/2}, & k=3 \\ \frac{3^{k/2}}{2}, & k \geq 4 \end{cases}, \quad \textit{and} \quad a_k \overset{\texttt{def}}{=} \begin{cases} \frac{3}{5}\left((4/3)^{k/2+1}-1\right)2^{k/2+1}, & k=3,4 \\ 2^{k/2+1}, & k \geq 5. \end{cases} \qquad (12)$$

Apparently, a closed form bound for $\mathbf{d}_n$ would be much more convenient than a recursive one. Such a bound is offered by the next theorem for i.i.d. X_i's having the same first $k-1$ moments as $\mathcal{N}$ with $k \geq 3$. This is not a restriction since the r.v. $(X - \mathbb{E}X)/\sqrt{VX}$ fulfills this requirement at least for $k = 3$. The closed form bound is proven inductively (see Section 4) by the aid of (10) and it is asymptotically equivalent to (albeit slightly worse than) the recursive bound of Theorem 2.

Theorem 4 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \geq 3$ and $\mathcal{L}X$ satisfies condition $A(k)$, then*

$$\mathbf{d}_n \leq \frac{n}{(n-1)^{k/2}} \frac{\left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k}{2(1-\varepsilon_k)} + \frac{1}{2}(2\mathbf{d}_1)^n \overset{\texttt{def}}{=} \mathbf{D}_{k,n}, \quad n \geq 2. \qquad (13)$$

In the next theorem, we construct a slightly more complicated bound in which the asymptotic constant is equal to $\left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k/2$, that is, ε_k in the denominator of (13) no longer appears. We achieve this by appropriately applying twice the result of Theorem 4 (see Section 4).

Theorem 5 *Let $\mathcal{L}X \in \mathcal{M}_k$ for some $k \geq 3$ and $\mathcal{L}X$ satisfies condition $A(k)$. Denote*

$$\varepsilon^*_{k,m} = \frac{2^{k/2+1}\mathbf{D}_{k,m}}{1 + a_k\mathbf{D}_{k,m}} + \frac{\delta_k \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k}{m^{k/2-1}},$$

where δ_k, a_k and $\mathbf{D}_{k,m}, m \geq 2$ are defined in (12) and Theorem 4 respectively ($\mathbf{D}_{k,1} \overset{\texttt{def}}{=} \mathbf{d}_1$). For $m \geq 1$ large enough such that $\varepsilon^*_{k,m} < 1, \mathbf{D}_{k,m} < 2^{-k/2-1}$ and for $h \geq 2$, we have that

$$\mathbf{d}_{hm} \leq \frac{(1-\frac{1}{h})^{-k/2} \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k}{2(hm)^{k/2-1}\left(1-\varepsilon^*_{k,m}\right)} + \frac{(2\mathbf{D}_{k,m})^h}{2} \overset{\texttt{def}}{=} \mathbf{D}^*_{k,h,m}. \tag{14}$$

The bound $\mathbf{D}^*_{k,h,m}$ is better than $\mathbf{D}_{k,hm}$ except perhaps of small values of m, h. Moreover, $\mathbf{D}_{k,m} \to 0$ and $\varepsilon^*_{k,m} \to 0$ as $m, h \to \infty$, and for $n = hm$,

$$\mathbf{d}_n \leq \mathbf{D}^*_{k,h,m} \sim \frac{\left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k}{2n^{k/2-1}}. \tag{15}$$

The asymptotic constant $\left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k/2$ can now be proven to be optimal, at least in the case $k = 3$ for a large class of distributions (see Corollary 8 below).

Remark 1. We can exploit once more the idea behind the construction of the bound in Theorem 5, and apply twice the recursive relation of Theorem 2 in order to achieve a sharper recursive bound for $\mathbf{d}_{hm}$. This can be easily accomplished as follows. We first derive an upper bound for $\mathbf{d}_m$ via the recursive relation (10) and then, after setting $Y_i = m^{-1/2}\sum_{j=1}^m X_{(i-1)m+j}$, we establish an upper bound for $\mathbf{d}_{hm} = \mathbf{d}(h^{-1/2}\sum_{i=1}^h Y_i, \mathcal{N}) \overset{\texttt{def}}{=} \mathbf{d}^*_h$, employing once more the recursion (10), but now for the sequence $Y_1, ..., Y_h$ (instead of $X_1, ..., X_h$), starting from $\mathbf{d}^*_1 = \mathbf{d}_m$, that is,

$$\mathbf{d}^*_i \leq \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}^*_k \sum_{j=2}^{i-1} \frac{\mathbf{d}^*_{j-1}}{(i-j)^{k/2}} + \frac{i\left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}^*_k}{2(i-1)^{k/2}} + 2\mathbf{d}^*_1\mathbf{d}^*_{i-1}, i = 2, 3, ..., h, \tag{16}$$

where $\boldsymbol{\zeta}^*_k = \boldsymbol{\zeta}_k(Y_1, \mathcal{N})$ can be replaced by its upper bound $m^{-k/2+1}\boldsymbol{\zeta}_k$ (see (5)).

Remark 2. It is worth stressing that the distance $\boldsymbol{\zeta}_k$ appearing in the above results can be easily evaluated when $\mathcal{L}X \leq_{k-cx} \mathcal{N}$ or $\mathcal{N} \leq_{k-cx} \mathcal{L}X$. In this case we deduce that (see (9)),

$$\boldsymbol{\zeta}_k(X, \mathcal{N}) = \frac{\left|\mathbb{E}X^k - \mu_k\right|}{k!}, \ k = 1, 2, ... \tag{17}$$

Otherwise, $\boldsymbol{\zeta}_k(X, \mathcal{N})$ can be upper bounded by $\left(\mathbb{E}|X|^k + \mathbb{E}|Z|^k\right)/k!$ where $\mathcal{L}Z = \mathcal{N}$, provided that $\mathbb{E}X^i = \mathbb{E}Z^i, i = 1, 2, ..., k-1$. It would thus be useful to describe a simple criterion for the verification of the s-convex order between two given distributions, which would enable the use of (17). Initially, define the number of sign-changes of a function ϕ on $\mathbb{R}$ by

$$S^-(\phi) \overset{\texttt{def}}{=} \sup\{S^-[\phi(x_1), \phi(x_2), ...\phi(x_n)] : x_1 < x_2 < ... < x_n \in \mathbb{R}, n \in \mathbb{N}\}, \tag{18}$$

where $S^-[y_1, y_2, ..., y_n]$ denotes the number of sign changes in the sequence $y_1, y_2, ..., y_n$ (zero terms are being discarded). According to the generalized Karlin-Novikoff cut-criterion (see Denuit, Lefèvre and Shaked (1998)) we have $X \leq_{s-cx} Y$ when (a) $\mathbb{E}(X^j - Y^j) = 0, j = 1, 2, ..., s-1$ and (b) their c.d.f.'s F_X, F_Y have $s-1$ crossing points (that is, $S^-(F_X - F_Y) = s-1$) while the last sign of $F_X - F_Y$ is a $+$. If X, Y possess Lebesgue densities f_X, f_Y, then condition (b) can be replaced by $S^-(f_X - f_Y) \le s$, while the last sign of $f_Y - f_X$ is a $+$. Hence, if the density of a continuous r.v. X crosses the standard normal density φ at most k times (and $\mathcal{L}X \in \mathcal{M}_k$) then $\mathcal{L}X \le_{k-cx} \mathcal{N}$ or $\mathcal{N} \le_{k-cx} \mathcal{L}X$. Many known distributions seem to satisfy these conditions when standardized with $k=3$ (e.g. gamma, Weibull, lognormal) or $k=4$ (e.g. uniform, Laplace, triangular).

Remark 3. Regarding the condition $A(k)$, we may encounter one of the following two cases:

(a) The distribution $\mathcal{L}X$ of the summands is relatively close to $\mathcal{N}$ so that condition $A(k)$ is satisfied (e.g., if $k=3$ and $\mathbf{d}_1 = 0.1$ then $\boldsymbol{\zeta}_3$ must be less than about 0.13) and hence we can employ directly Theorems 4, 5 (also in this case the recursive bound of Theorem 2 tends to zero).

(b) The condition $A(k)$ is not satisfied. If in this case $\mathbf{d}_1 < 1$ then we can overcome this condition by considering an appropriate subsequence of $n^{-1/2}\sum_{i=1}^n X_i, n = 1, 2, \ldots$. More specifically, we observe that there always exist an integer r large enough so that the distribution of the r.v. $X' = r^{-1/2}\sum_{i=1}^r X_i$ satisfies condition $A(k)$. This follows from the fact that $\mathbf{d}(X', \mathcal{N}) \to_{r\to\infty} 0$ (provided that $\mathbf{d}_1 < 1$; see Prokhorov (1952)) and $\boldsymbol{\zeta}_k(X', \mathcal{N}) \le r^{-k/2+1}\boldsymbol{\zeta}_k \to_{r\to\infty} 0$ (see (5)). Therefore, we can apply the above Theorems 4 and 5 for the i.i.d. r.v's $X_i' = r^{-1/2}\sum_{j=(i-1)r+1}^{ir} X_j, i = 1, 2, \ldots$ and deduce bounds for the distance $\mathbf{d}\left(n^{-1/2}\sum_{i=1}^n X_i', \mathcal{N}\right) = \mathbf{d}_{nr},\ n = 1, 2, \ldots$. This procedure is illustrated in the applications (Section 5). Note that, since $\mathcal{N}$ is absolutely continuous, the condition $\mathbf{d}_1 < 1$ implies that $\mathcal{L}X$ must have an *absolutely continuous part.*

In case we have constructed (for example, via the procedure described in Remark 3 or via Theorem 5) bounds for a subsequence of the form $\mathbf{d}_{hm}, h = 1, 2, \ldots$, we can use the following proposition to deduce a simple bound for $\mathbf{d}_{hm+j}$ and therefore gain bounds for the whole sequence above m, that is, $\mathbf{d}_i, i = m, m+1, \ldots$. For the proof of this proposition see Section 4.

Proposition 6 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \ge 3$, then*

$$\mathbf{d}_{s+j} \le \mathbf{d}_s + \frac{j}{2s^{k/2}} \left\| \varphi^{(k)} \right\| \boldsymbol{\zeta}_k, \qquad s \ge 1, j \ge 1.$$

Next, utilizing Theorem 5, we present two asymptotic results that hold true when F_X has an absolutely continuous part, regardless of condition $A(k)$ (their proofs are given in Section 4). We remind that every c.d.f. F can be written as a convex combination of a singular c.d.f. F_s (discrete and/or singular continuous) and an absolutely continuous c.d.f. F_{ac}, that is, $F = pF_{ac} + (1-p)F_s, p \in [0,1]$. By saying that F has an absolutely continuous part, we mean that $p > 0$.

Corollary 7 *If $\mathcal{L}X \in \mathcal{M}_k$ for some $k \ge 3$ and $\mathcal{L}X$ has an absolutely continuous part, then*

$$\limsup_{n\to\infty} n^{k/2-1}\mathbf{d}\left(\tfrac{1}{\sqrt{n}}\sum_{i=1}^n X_i, \mathcal{N}\right) \le \tfrac{1}{2}\left\| \varphi^{(k)} \right\| \boldsymbol{\zeta}_k. \tag{19}$$

Note that Sirazhdinov and Mamatov (1962) proved that, if $\mathcal{L}X$ has an absolutely continuous part and $|EX^3| < \infty$ then

$$\mathbf{d}_n = \frac{1+4e^{-3/2}}{6\sqrt{2\pi}} \frac{|\mathbb{E}X^3|}{\sqrt{n}} + o(\tfrac{1}{\sqrt{n}}). \tag{20}$$

We show below that, when $k=3$, the bound of Theorem 5 is asymptotically equal to (20) for a large class of distributions. In contrast to (14), the above result cannot be used in practice for finite n since the last term is not explicitly bounded.

Corollary 8 *If $\mathcal{L}X$ has an absolutely continuous part, $|\mathbb{E}X^3| < \infty$ and $\mathcal{L}X \geq_{3-cx} \mathcal{N}$ or $\mathcal{L}X \leq_{3-cx} \mathcal{N}$, then the bound $\mathbf{D}^*_{3,h,m}$ of Theorem 5 is asymptotically optimal, that is*

$$\mathbf{d}_n \sim \mathbf{D}^*_{3,h,m} = \frac{1+4e^{-3/2}}{6\sqrt{2\pi}} \frac{|\mathbb{E}X^3|}{\sqrt{n}} + o(\tfrac{1}{\sqrt{n}}), \quad \textit{as } h, m \to \infty, n = hm.$$

3.2 Bounds for the Kolmogorov distance

We now turn our attention to results concerning the Kolmogorov distance $\boldsymbol{\rho}$, which is most often used for expressing the approximation error in the CLT. The results to follow are being established invoking also developments of the previous subsection. The next theorem is analogous to Theorem 2 (for its proof see Section 4) offering a recursive bound for $\boldsymbol{\rho}_n$ in terms of $\mathbf{d}_1$, $\boldsymbol{\zeta}_k$ and $\boldsymbol{\rho}_1$.

Theorem 9 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \geq 1$, then, for all $n \geq 2$,*

$$\boldsymbol{\rho}_n \leq \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k \sum_{j=2}^{n-1} \frac{\boldsymbol{\rho}_{j-1}}{(n-j)^{k/2}} + \frac{n \left\|\varphi^{(k-1)}\right\|_\infty \boldsymbol{\zeta}_k}{(n-1)^{k/2}} + 2\mathbf{d}_1 \boldsymbol{\rho}_{n-1}. \tag{21}$$

The explicit value of $\left\|\varphi^{(k)}\right\|_\infty$ can be found using lengthy calculus techniques (for small k) or it can be approximated via numerical integration (for large k). For example,

$$\left\|\varphi^{(2)}\right\|_\infty = \tfrac{1}{\sqrt{2\pi}}, \quad \left\|\varphi^{(3)}\right\|_\infty = \frac{e^{\frac{-3+\sqrt{6}}{2}}\sqrt{9-3\sqrt{6}}}{\sqrt{\pi}}, \quad \left\|\varphi^{(4)}\right\|_\infty = \tfrac{3}{\sqrt{2\pi}}, \quad \left\|\varphi^{(6)}\right\|_\infty = \tfrac{15}{\sqrt{2\pi}},$$

and, approximately, $\left\|\varphi^{(5)}\right\|_\infty \approx 2.30711$, $\left\|\varphi^{(7)}\right\|_\infty \approx 14.177977$.

Again, the recursive bound derived from (21) vanishes only when $k \geq 3$ and $\mathcal{L}X$ is relatively close to $\mathcal{N}$, that is, when $\mathbf{d}_1$ and $\boldsymbol{\zeta}_k$ are relatively small. In the following definition, which is analogous to Definition 3, we offer an explicit condition which guarantees this distributional closeness.

Definition 10 *The distribution $\mathcal{L}X$ of a r.v. X, with $\mathcal{L}X \in \mathcal{M}_k$ for some $k \geq 3$, satisfies the condition $B(k)$ when*

$$\mathbf{d}_1 < 2^{-k/2-1} \quad \textit{and} \quad \beta_k \overset{\texttt{def}}{=} b_k \mathbf{d}_1 + \delta_k \left\|\varphi^{(k)}\right\| \boldsymbol{\zeta}_k < 1$$

where δ_k is defined in (12) and $b_k = \frac{2^{k/2}}{3}\left(4 + \left\|\varphi^{(k)}\right\| \cdot \left\|\varphi^{(k-1)}\right\|_\infty^{-1}\right)$.

Next, from Theorem 9 we obtain a closed form bound for $\mathbf{d}_n$ (proven inductively in Section 4), which is slightly larger than (but asymptotically equal to) the recursive bound derived via (21).

Theorem 11 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \geq 3$, and $\mathcal{L}X$ satisfies condition $B(k)$ then*

$$\boldsymbol{\rho}_n \leq \frac{n\left\|\varphi^{(k-1)}\right\|_\infty \boldsymbol{\zeta}_k}{(n-1)^{k/2}(1-\beta_k)} + \frac{(2\mathbf{d}_1)^n}{2} \overset{\texttt{def}}{=} \mathbf{R}_{k,n}, \ n \geq 2. \tag{22}$$

By applying successively Theorems 4, 11 (see Section 4) we construct a slightly more complicated bound analogous to that of Theorem 5, with the asymptotic constant $\left\|\varphi^{(k-1)}\right\|_\infty \boldsymbol{\zeta}_k$.

Theorem 12 *Let $\mathcal{L}X \in \mathcal{M}_k$ for some $k \geq 3$ and $\mathcal{L}X$ satisfies condition $A(k)$. Define*

$$\beta^*_{k,m} = b_k \mathbf{D}_{k,m} + \frac{\delta_k \left\| \varphi^{(k)} \right\| \boldsymbol{\zeta}_k}{m^{k/2-1}},$$

where δ_k and $\mathbf{D}_{k,m}, m \geq 2$ are the same as in Theorem 4 ($\mathbf{D}_{k,1} \overset{\texttt{def}}{=} \mathbf{d}_1$), b_k is given in Definition 10. For $m \geq 1$ large enough such that $\beta^*_{k,m} < 1, \mathbf{D}_{k,m} < 2^{-k/2-1}$, we have

$$\boldsymbol{\rho}_{hm} \leq \frac{(1-\frac{1}{h})^{-k/2}}{(hm)^{k/2-1}} \frac{\left\| \varphi^{(k-1)} \right\|_\infty \boldsymbol{\zeta}_k}{1-\beta^*_{k,m}} + \frac{(2\mathbf{D}_{k,m})^h}{2} \overset{\texttt{def}}{=} \mathbf{R}^*_{k,h,m}, \quad h \geq 2. \tag{23}$$

It holds $\mathbf{R}^*_{k,h,m} \leq \mathbf{R}_{k,hm}$ except perhaps of small m, h. As $m, h \to \infty$ ($n = hm$), we have that

$$\boldsymbol{\rho}_n \leq \mathbf{R}^*_{k,h,m} \sim \frac{\left\| \varphi^{(k-1)} \right\|_\infty \boldsymbol{\zeta}_k}{n^{k/2-1}}. \tag{24}$$

The asymptotic constant $\left\| \varphi^{(k-1)} \right\|_\infty \boldsymbol{\zeta}_k$ is optimal for a large class of distributions (see Corollary 15) and all $k \geq 3$.

Remark 4. As we discussed in Remark 1, we can also apply here successively the recursive bounds of Theorems 2 and 9 and obtain a sharper recursive bound for $\boldsymbol{\rho}_{hm}$. More specifically, we initially construct an upper bound for $\mathbf{d}_m$ via the recursive relation (10) and then deduce an upper bound for $\boldsymbol{\rho}_{hm} = \boldsymbol{\rho}(h^{-1/2}\sum_{i=1}^h Y_i, \mathcal{N}) \overset{\texttt{def}}{=} \boldsymbol{\rho}^*_h$ employing the recursions of Theorem 9 for the i.i.d. r.v.'s $Y_1, ..., Y_h$ (see Remark 1) starting from $\mathbf{d}^*_1 = \mathbf{d}_m$, that is,

$$\boldsymbol{\rho}^*_i \leq \left\| \varphi^{(k)} \right\| \boldsymbol{\zeta}^*_k \sum_{j=2}^{i-1} \frac{\boldsymbol{\rho}^*_{j-1}}{(i-j)^{k/2}} + \tfrac{i}{(i-1)^{k/2}} \left\| \varphi^{(k-1)} \right\|_\infty \boldsymbol{\zeta}^*_k + 2\mathbf{d}^*_1 \boldsymbol{\rho}^*_{i-1}, \quad i = 2, 3, ..., h, \tag{25}$$

where $\boldsymbol{\rho}^*_1$ and $\boldsymbol{\zeta}^*_k = \boldsymbol{\zeta}_k(Y_1, \mathcal{N})$ can be replaced by their upper bounds $\mathbf{d}^*_1$ and $m^{-k/2+1}\boldsymbol{\zeta}_k$ (see (5)) respectively.

Remark 5. Similarly to the total variation distance case (see Remark 3), if the condition $A(k)$ or $B(k)$ is not satisfied in order to directly use Theorems 12 or 11 respectively, we can alternatively apply these theorems for the i.i.d. r.v's $X'_i = r^{-1/2}\sum_{j=(i-1)r+1}^{ir} X_j, i = 1, 2, ...$ for large enough r (provided that $\mathbf{d}_1 < 1$) and deduce bounds for the subsequence $\boldsymbol{\rho}_{nr}, n = 1, 2, ...$. This procedure is illustrated in the examples of applications (Section 5).

The following proposition is analogous to Proposition 6 (with a similar proof which is omitted). It can be used to bound $\boldsymbol{\rho}_i, i \geq m$ when we have at hand bounds for a subsequence $\boldsymbol{\rho}_{hm}, h \geq 1$.

Proposition 13 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \geq 3$ then*

$$\boldsymbol{\rho}_{s+j} \leq \boldsymbol{\rho}_s + \frac{j}{s^{k/2}} \left\| \varphi^{(k-1)} \right\|_\infty \boldsymbol{\zeta}_k, \quad s \geq 1, j \geq 1.$$

The next result is analogous to Corollary 7 with a similar proof which is omitted.

Corollary 14 *If $\mathcal{L}X \in \mathcal{M}_k$ for some $k \geq 3$ and $\mathcal{L}X$ has an absolutely continuous part, then*

$$\limsup_{n\to\infty} n^{k/2-1} \boldsymbol{\rho}\left(\tfrac{1}{\sqrt{n}} \sum_{i=1}^n X_i, \mathcal{N}\right) \leq \left\| \varphi^{(k-1)} \right\|_\infty \boldsymbol{\zeta}_k. \tag{26}$$

It is well known (e.g. see Petrov (1995), p.173) that if the i.i.d. r.v.'s $X, X_1, X_2, \dots$ follow a non-lattice distribution, with $\mathbb{E}X = 0, \mathbb{E}X^2 = 1$, $\limsup_{|t|\to\infty} |\mathbb{E}e^{\mathrm{i}tX}| < 1$ and $\mathbb{E}|X|^k < \infty$ for some integer $k \geq 3$, then the Edgeworth expansion of the c.d.f. F_n of $n^{-1/2}\sum_{i=1}^{n} X_i$ is

$$F_n(x) = \Phi(x) + \sum_{v=1}^{k-2} \frac{Q_\nu(x)}{n^{\nu/2}} + o(\frac{1}{n^{k/2-1}}), \tag{27}$$

uniformly in x (see also (2) for $k = 3$). The quantity $Q_\nu(x)$ can be expressed in terms of the Chebyshev-Hermite polynomials $\mathcal{H}_k$ (see (7)),

$$Q_\nu(x) = -\frac{e^{-x^2/2}}{\sqrt{2\pi}} \sum \mathcal{H}_{\nu+2s-1}(x) \prod_{m=1}^{v} \frac{1}{k_m!} \left(\frac{\gamma_{m+2}}{(m+2)!} \right)^{k_m}, \tag{28}$$

where γ_m denotes the cumulant of order m of the r.v. X, and the summation above is considered over all integer solutions $(k_1, \dots, k_\nu)$ of $k_1 + 2k_2 + \dots + \nu k_\nu = \nu$, and $s = k_1 + k_2 + \dots + k_\nu$. Note that, the aforementioned condition $\limsup_{|t|\to\infty} |\mathbb{E}e^{\mathrm{i}tX}| < 1$ is satisfied when X has an absolutely continuous part (for example, see Petrov (1995), p. 12). The estimates (27) describe the rate of convergence of F_n to Φ but, in contrast to the bounds of Theorem 12, they are practically useless if we are interested in estimates of the difference $F_n(x) - \Phi(x)$ for a given fixed value of n.

In the last corollary we prove (see Section 4) that, for a large class of distributions of the summands, the error bound of Theorem 12 asymptotically coincides with the approximation error implied by the Edgeworth expansion (27), and thus (23) is asymptotically optimal.

Corollary 15 *If $\mathcal{L}X \in \mathcal{M}_k$ for some integer $k \geq 3$, $\mathbb{E}|X|^k < \infty$, $\mathcal{L}X$ has an absolutely continuous part, and $\mathcal{L}X \geq_{k-cx} \mathcal{N}$ or $\mathcal{L}X \leq_{k-cx} \mathcal{N}$, then the bound $\mathbf{R}^*_{k,h,m}$ of Theorem 12 is asymptotically optimal, that is,*

$$\boldsymbol{\rho}_n \sim \mathbf{R}^*_{k,h,m} = \frac{\left\|\varphi^{(k-1)}\right\|_\infty}{n^{k/2-1}} \frac{\left|\mathbb{E}X^k - \mu_k\right|}{k!} + o(\tfrac{1}{n^{k/2-1}}), \quad \textit{as } h, m \to \infty, n = hm.$$

4 Proofs

The following Lemma 16 offers some well known *smoothing inequalities* for the metrics $\mathbf{d}$ and $\boldsymbol{\rho}$ that have been employed by many authors in the past for the derivation of Berry-Esseen-type or Poisson approximation results. The proof of part (a) can be found in Rachev (1991), p. 274 (see also Boutsikas and Vaggelatou (2009) for a more general result), while the proof of part (b) can be found in Senatov (1980) (see also Rachev (1991), p. 323).

Lemma 16 *(a) If the random vectors $X, Y \in \mathbb{R}^r$ are independent of $Z, W \in \mathbb{R}^r$ then*

$$\mathbf{d}(Z + X, Z + Y) \leq 2\mathbf{d}(X, Y)\mathbf{d}(Z, W) + \mathbf{d}(W + X, W + Y).$$

(b) If the random variables $X, Y \in \mathbb{R}$ are independent of $Z, W \in \mathbb{R}$ then

$$\begin{aligned} \boldsymbol{\rho}(Z + X, Z + Y) &\leq 2\mathbf{d}(X, Y)\boldsymbol{\rho}(Z, W) + \boldsymbol{\rho}(W + X, W + Y), \\ \boldsymbol{\rho}(Z + X, Z + Y) &\leq 2\boldsymbol{\rho}(X, Y)\mathbf{d}(Z, W) + \boldsymbol{\rho}(W + X, W + Y). \end{aligned}$$

Note that, the above inequalities are usually stated for independent X, Y, Z, W, but they are also valid under the weaker assumption that X, Y are independent of Z, W.

Several variations of the following lemma can be found in a number of papers in the literature. For example, see Rachev (1991), p. 325, Zolotarev (1983), p. 294 or Rachev and Ruschendorf (1990) for similar results. For discrete analogues (leading to Poisson approximation results) see Boutsikas and Vaggelatou (2009).

Lemma 17 *Let Z be a r.v. independent of the r.v.'s X, Y, with density f_Z, k-times differentiable on $\mathbb{R}$, and denote by $f_Z^{(k)}$ its kth order derivative ($f^0 = f$). If $\boldsymbol{\zeta}_k(X,Y) < \infty,\ k \geq 1$, then,*

$$\mathbf{d}(X+Z, Y+Z) \leq \frac{1}{2}\left\|f_Z^{(k)}\right\| \boldsymbol{\zeta}_k(X,Y), \quad \text{and} \quad \boldsymbol{\rho}(X+Z, Y+Z) \leq \left\|f_Z^{(k-1)}\right\|_\infty \boldsymbol{\zeta}_k(X,Y).$$

where $\|f\| = \int_{-\infty}^{\infty}|f(x)|\,dx$, and $\|f\|_\infty = \sup_x |f(x)|$.

Proof. Set $H(x) = \bar{F}_X(x) - \bar{F}_Y(x)$ where $\bar{F}_X(x) = P(X > x)$. Define $H^{[0]}(x) = H(x)$ and recursively set

$$H^{[i]}(t) = \int_t^\infty H^{[i-1]}(x)dx, \quad i = 1, 2, ..., k-1.$$

Using the fact that $(i+1)\int_x^\infty \mathbb{E}(X-t)_+^i dt = \mathbb{E}(X-x)_+^{i+1}$, it can be easily verified recursively that (see also Denuit, Lefèvre and Shaked (1998)),

$$i!H^{[i]}(t) = \mathbb{E}(X-t)_+^i - \mathbb{E}(Y-t)_+^i, \quad i = 0, 1, ..., k-1.$$

Denote by $f_X^{(i)}$ the ith order derivative of the p.d.f. f_X of a r.v. X. We observe that the r.v.'s $X+Z, Y+Z$ possess a Lebesgue density and

$$\begin{aligned} f_{X+Z}(w) - f_{Y+Z}(w) &= -\int_{-\infty}^{\infty} f_Z(w-z)dH(z) = -\int_{-\infty}^{\infty} H^{[0]}(z) f_Z^{(1)}(w-z)dz \\ &= \int_{-\infty}^{\infty} H^{[1]}(z)\ f_Z^{(2)}(w-z)dz = ... = (-1)^k \int_{-\infty}^{\infty} H^{[k-1]}(z)\ f_Z^{(k)}(w-z)dz. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbf{d}(X+Z, Y+Z) &= \tfrac{1}{2}\int_{-\infty}^{\infty}\left|\int_{-\infty}^{\infty} H^{[k-1]}(z)\ f_Z^{(k)}(w-z)dz\right| dw \leq \tfrac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|H^{[k-1]}(z)\ f_Z^{(k)}(w-z)\right| dw\ dz \\ &= \tfrac{1}{2}\int_{-\infty}^{\infty}\left|\frac{\mathbb{E}(X-z)_+^{k-1} - \mathbb{E}(Y-z)_+^{k-1}}{(k-1)!}\right| \left\|f_Z^{(k)}\right\| dz = \frac{1}{2}\left\|f_Z^{(k)}\right\| \boldsymbol{\zeta}_k(X,Y). \end{aligned}$$

Likewise, $\boldsymbol{\rho}(X+Z, Y+Z)$ is bounded above by

$$\sup_w \left|\bar{F}_{X+Z}(w) - \bar{F}_{Y+Z}(w)\right| = \sup_w \left|\int_{-\infty}^{\infty} \bar{H}^{[k-1]}(z)\ f_Z^{(k-1)}(w-z)dz\right| \leq \left\|f_Z^{(k-1)}\right\|_\infty \boldsymbol{\zeta}_k(X,Y).$$

■

We are now in possession of the machinery needed in order to prove our first main result.

Proof. (**of Theorem 2**) Let $Z_1, Z_2, ...$ be i.i.d. standard normal r.v.'s, independent of X_i's. From the triangle inequality for $\mathbf{d}$ we deduce that (Lindeberg decomposition),

$$\mathbf{d}_n = \mathbf{d}\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} Z_i\right) \leq \sum_{j=1}^{n} \mathbf{d}\left(\sum_{i=1}^{j} X_i + \sum_{i=j+1}^{n} Z_i, \sum_{i=1}^{j-1} X_i + \sum_{i=j}^{n} Z_i\right). \tag{29}$$

Employing Lemma 16 for the r.v.'s $X = X_j + \Sigma_{i=j+1}^{n} Z_i, Y = \Sigma_{i=j}^{n} Z_i, Z = \Sigma_{i=1}^{j-1} X_i, W = \Sigma_{i=1}^{j-1} Z_i$ we get for $n \geq 2, j = 1, 2, ..., n$,

$$\begin{aligned} &\mathbf{d}\left(\sum_{i=1}^{j} X_i + \sum_{i=j+1}^{n} Z_i, \sum_{i=1}^{j-1} X_i + \sum_{i=j}^{n} Z_i\right) \\ \leq\ \ & 2\mathbf{d}\left(X_j + \sum_{i=j+1}^{n} Z_i, Z_j + \sum_{i=j+1}^{n} Z_i\right)\mathbf{d}_{j-1} + \mathbf{d}\left(X_j + \sum_{i=1}^{n} Z_i - Z_j, \sum_{i=1}^{n} Z_i\right), \end{aligned} \tag{30}$$

where $\mathbf{d}_0 = 0$. Invoking Lemma 17, for $j = 1, 2, ..., n-1$, the right part of the above inequality is upper bounded by (see also (8))

$$\boldsymbol{\zeta}_k(X_j, Z_j)\left\|\varphi_{0,n-j}^{(k)}\right\|\mathbf{d}_{j-1} + \tfrac{1}{2}\boldsymbol{\zeta}_k(X_j, Z_j)\left\|\varphi_{0,n-1}^{(k)}\right\| = \boldsymbol{\zeta}_k \tfrac{\|\varphi^{(k)}\|}{(n-j)^{k/2}}\mathbf{d}_{j-1} + \tfrac{1}{2}\boldsymbol{\zeta}_k \tfrac{\|\varphi^{(k)}\|}{(n-1)^{k/2}}, \tag{31}$$

(φ_{μ,σ^2} denotes the density of $\mathcal{N}_{\mu,\sigma^2}$) while for $j = n$ it is upper bounded by

$$2\mathbf{d}(X_n, Z_n)\mathbf{d}_{n-1} + \tfrac{1}{2}\boldsymbol{\zeta}_k(X_n, Z_n)\left\|\varphi_{0,n-1}^{(k)}\right\| = 2\mathbf{d}_1\mathbf{d}_{n-1} + \tfrac{1}{2}\boldsymbol{\zeta}_k \tfrac{\|\varphi^{(k)}\|}{(n-1)^{k/2}}. \tag{32}$$

Finally, combining (29) (30) (31) and (32) we obtain (10). ■

The following two technical results are essential for the proof of Theorem 4.

Lemma 18 *For every couple of integers $k \geq 3, m \geq 4$, we have*

$$b_{k,m} \overset{\texttt{def}}{=} \frac{(m-1)^{k/2}}{m} \sum_{j=3}^{m-1} \frac{(j-1)}{(m-j)^{k/2}(j-2)^{k/2}} \leq \delta_k, \tag{33}$$

where $\delta_k = \frac{3^{k/2}}{2}$ for $k \geq 4$, and $\delta_3 = (\frac{5}{3})^{3/2} + \frac{1}{2}(\frac{5}{4})^{3/2} \approx 2.8504$.

Proof. Initially, assume that $k = 3, m \geq 4$. It can be verified that the function

$$f(x) = \tfrac{(m-1)^{3/2}(x-1)}{m(m-x)^{3/2}(x-2)^{3/2}}, \quad x \in (2, m),$$

is convex and therefore,

$$b_{3,m} = \sum_{i=3}^{m-1} f(i) \leq \sum_{i=3}^{m-1} \int_{i-1/2}^{i+1/2} f(x)dx = \int_{3-1/2}^{(m-1)+1/2} f(x)dx = \tfrac{(m-1)^{3/2}4(m-3)}{(m-2)^2\sqrt{2m-5}} \overset{\texttt{def}}{=} h(m).$$

The above function $h : [4, \infty) \to \mathbb{R}$ is decreasing for $m > m_0 = \frac{23+\sqrt{73}}{6} \approx 5.257$ ($h'(x) = 0$ in $[4, \infty)$ only at $x = m_0$), converging to $2\sqrt{2} \approx 2.82843$ as $m \to \infty$. Therefore, $b_{3,m} \leq h(100) \approx 2.8498 < \delta_3$

for all $m \geq 100$. For $4 \leq m < 100$ it can now be numerically checked that $\max\{b_{3,4}, b_{3,5}, ..., b_{3,99}\} = b_{3,6} = \delta_3$ and hence (33) is valid for $k = 3, m \geq 4$. Similarly, for $k = 4, m \geq 4$ we deduce that

$$b_{4,m} \leq \int_{3-1/2}^{m-1+1/2} \tfrac{(m-1)^2(x-1)}{m(m-x)^2(x-2)^2} dx = \tfrac{(m-1)^2(24+(m-5)4m+2(2m-5)\ln(2m-5))}{(m-2)^3(2m-5)} \overset{\texttt{def}}{=} g(m).$$

The function $g : [4, \infty) \to \mathbb{R}$ is decreasing, converging to 2, and therefore, $b_{4,m} \leq g(6) \approx 4.19881 < \delta_4 = 4.5$ for all $m \geq 6$. For $4 \leq m < 6$ it can be numerically checked that $\max\{b_{4,4}, b_{4,5}\} = b_{4,4} = \delta_4$ and hence (33) is again valid for $k = 4, m \geq 4$. Suppose now that $\sup_{m\geq 7} b_{k,m} \leq \frac{3^{k/2}}{2} = \delta_k$ for some $k \geq 4$. We have shown above that this inequality is valid for $k = 4$. For $m \geq 7$ it can be easily verified that $\frac{m-1}{(m-j)(j-2)} \leq 1, j = 4, ..., m-2$, and thus, for $m \geq 7$,

$$\begin{aligned} b_{k+1,m} &= \sum_{j=3}^{m-1} \frac{(m-1)^{\frac{k+1}{2}}(j-1)}{m(m-j)^{\frac{k+1}{2}}(j-2)^{\frac{k+1}{2}}} \leq b_{k,m} + \left(\tfrac{2(m-1)^{k/2}}{m(m-3)^{k/2}} + \tfrac{(m-1)^{k/2}(m-2)}{m(m-3)^{k/2}}\right)\left(\sqrt{\tfrac{m-1}{m-3}} - 1\right) \\ &\leq \tfrac{3^{k/2}}{2} + \tfrac{(m-1)^{k/2}}{(m-3)^{k/2}}\left(\sqrt{\tfrac{m-1}{m-3}} - 1\right) \leq \tfrac{3^{k/2}}{2} + \tfrac{3^{k/2}}{2^{k/2}}\left(\sqrt{\tfrac{3}{2}} - 1\right) \leq \tfrac{3^{(k+1)/2}}{2}. \end{aligned}$$

Therefore $\sup_{m\geq 7} b_{k+1,m} \leq \delta_{k+1}$ and inductively, $\sup_{m\geq 7} b_{k,m} \leq \frac{3^{k/2}}{2} = \delta_k, k = 4, 5, ...$. It finally suffices to prove that $b_{k,m} \leq \delta_k$ for $m = 4, 5, 6$, and $k > 4$.For $m = 4$, it is $b_{k,4} = \delta_k, k \geq 4$. For $m = 5$ it is $b_{k,5} = 2^{k/2} \leq \delta_k, k \geq 4$. For $m = 6$ we have

$$b_{k,6} = \tfrac{5^{k/2}}{3^{k/2}} + \tfrac{5^{k/2}}{2 \cdot 4^{k/2}} \leq \tfrac{3^{k/2}}{2} = \delta_k, \quad k \geq 4,$$

and the proof is completed. ■

Lemma 19 *For every couple of integers* $k \geq 3, m \geq 5$ *we have*

$$A_{k,m} \overset{\texttt{def}}{=} \frac{1}{m} \sum_{j=3}^{m-1} \frac{(m-1)^{k/2}}{(2^{k/2})^{j-2}(m-j)^{k/2}} \leq \frac{2}{5}.$$

Proof. We use induction. For $m = 5$ we have $A_{k,5} = 2/5$ and suppose that $A_{k,m} \leq 2/5$ for some $m \geq 5$. We show that also $A_{k,m+1} \leq 2/5$. Indeed, $A_{k,m+1} = g_1(k,m)A_{k,m} + g_2(k,m)$ where the functions

$$g_1(k,m) = \tfrac{2^{-k/2}m^{k/2+1}}{(m+1)(m-1)^{k/2}}, \quad g_2(k,m) = \tfrac{2^{-k/2}m^{k/2}}{(m+1)(m-2)^{k/2}},$$

are decreasing with respect to m and k ($m \geq 5, k \geq 3$) and therefore, for $m \geq 5, k \geq 3$,

$$A_{k,m+1} \leq \tfrac{5^{3/2+1}}{6 \cdot 8^{3/2}} \tfrac{2}{5} + \tfrac{5^{3/2}}{6 \cdot 6^{3/2}} \approx 0.29149 < \tfrac{2}{5}.$$

■

Proof. (**of Theorem 4**) Denote for simplicity $c_k = \left\|\varphi^{(k)}\right\|$. From Theorem 2,

$$\mathbf{d}_n \leq c_k \boldsymbol{\zeta}_k \sum_{j=2}^{n-1} \tfrac{\mathbf{d}_{j-1}}{(n-j)^{k/2}} + 2\mathbf{d}_1\mathbf{d}_{n-1} + \tfrac{nc_k\boldsymbol{\zeta}_k}{2(n-1)^{k/2}}. \tag{34}$$

For $n = 2$, inequality (34) leads to

$$\mathbf{d}_2 \leq c_k\boldsymbol{\zeta}_k + 2\mathbf{d}_1^2 \leq \tfrac{c_k\boldsymbol{\zeta}_k}{1-\varepsilon_k} + \tfrac{1}{2}(2\mathbf{d}_1)^2 = \mathbf{D}_{k,2}, \tag{35}$$

and thus (13) is valid. Likewise, for $n = 3$, (34) along with (35) yields

$$\mathbf{d}_3 \leq c_k \boldsymbol{\zeta}_k \mathbf{d}_1 + 2\mathbf{d}_1\mathbf{d}_2 + \tfrac{3c_k\boldsymbol{\zeta}_k}{2^{k/2+1}} \leq c_k \boldsymbol{\zeta}_k \mathbf{d}_1 + 2\mathbf{d}_1(\boldsymbol{\zeta}_k c_k + 2\mathbf{d}_1^2) + \tfrac{3c_k\boldsymbol{\zeta}_k}{2^{k/2+1}} \leq \mathbf{D}_{k,3},$$

because $a_k \leq 2^{k/2+1}$, and therefore (13) is valid also for $n = 3$.Assume now that (13) holds for $n = 2, 3, ..., m-1$ $(m \geq 4)$. We shall show that it also holds for $n = m$. Invoking (34) we get

$$\mathbf{d}_m \leq \tfrac{c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-2)^{k/2}} + c_k\boldsymbol{\zeta}_k \sum_{j=3}^{m-1} \tfrac{\mathbf{d}_{j-1}}{(m-j)^{k/2}} + 2\mathbf{d}_1\mathbf{d}_{m-1} + \tfrac{mc_k\boldsymbol{\zeta}_k}{2(m-1)^{k/2}},$$

and hence, using the induction assumption, we get

$$\mathbf{d}_m \leq \tfrac{c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-2)^{k/2}} + \sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{k/2}} \left(\tfrac{(j-1)c_k\boldsymbol{\zeta}_k}{2(j-2)^{k/2}(1-\varepsilon_k)} + \tfrac{(2\mathbf{d}_1)^{j-1}}{2} \right) + \tfrac{\mathbf{d}_1(m-1)c_k\boldsymbol{\zeta}_k}{(m-2)^{k/2}(1-\varepsilon_k)} + \tfrac{(2\mathbf{d}_1)^m}{2} + \tfrac{mc_k\boldsymbol{\zeta}_k}{2(m-1)^{k/2}}.$$

To ensure that the above bound is less than or equal to $\mathbf{D}_{k,m}$, it suffices to guarantee that

$$\sum_{j=3}^{m-1} \frac{c_k\boldsymbol{\zeta}_k\left(\frac{j-1}{2(j-2)^{k/2}}\frac{c_k\boldsymbol{\zeta}_k}{1-\varepsilon_k} + \frac{(2\mathbf{d}_1)^{j-1}}{2}\right)}{(m-j)^{k/2}} + \frac{(m-\varepsilon_k)c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(1-\varepsilon_k)(m-2)^{k/2}} + \frac{mc_k\boldsymbol{\zeta}_k}{2(m-1)^{k/2}} \leq \frac{\frac{m}{(m-1)^{k/2}}c_k\boldsymbol{\zeta}_k}{2(1-\varepsilon_k)}, \tag{36}$$

or, multiplying both sides of (36) by $1 - \varepsilon_k \in (0, 1)$, it suffices to show that

$$\sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{k/2}} \left(\tfrac{(j-1)c_k\boldsymbol{\zeta}_k}{2(j-2)^{k/2}} + \tfrac{(2\mathbf{d}_1)^{j-1}}{2}(1-\varepsilon_k) \right) + \tfrac{(m-\varepsilon_k)c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-2)^{k/2}} \leq \tfrac{mc_k\boldsymbol{\zeta}_k\varepsilon_k}{2(m-1)^{k/2}}.$$

The above relation now follows readily when we add by parts the following two inequalities

$$\sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{k/2}} \tfrac{(j-1)c_k\boldsymbol{\zeta}_k}{2(j-2)^{k/2}} \leq \tfrac{mc_k\boldsymbol{\zeta}_k}{2(m-1)^{k/2}}\delta_k c_k\boldsymbol{\zeta}_k, \tag{37}$$

$$\sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{k/2}} \tfrac{(2\mathbf{d}_1)^{j-1}}{2}(1-\varepsilon_k) + \tfrac{(m-\varepsilon_k)c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-2)^{k/2}} \leq \tfrac{mc_k\boldsymbol{\zeta}_k}{2(m-1)^{k/2}} \tfrac{2^{k/2+1}\mathbf{d}_1}{1+a_k\mathbf{d}_1}. \tag{38}$$

Inequality (37) is straightforward from Lemma 18. Hence it remains to prove (38). First we prove it for $m = 4, k \geq 3$. For $m = 4$, the left part of inequality (38) reduces to $(2\mathbf{d}_1(1-\varepsilon_k) + \frac{4-\varepsilon_k}{2^{k/2}})c_k\boldsymbol{\zeta}_k\mathbf{d}_1$,which, since $\mathbf{d}_1 \leq 2^{-k/2-1}$, and $a_k \leq \frac{2^{k/2+1}}{5}(\frac{4^{k/2+1}}{3^{k/2}} - 3)$, is bounded above by

$$\begin{aligned}
\tfrac{(5-2\varepsilon_k)c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{2^{k/2}} \quad &\leq \quad \left(5 - 2\tfrac{2^{k/2+1}\mathbf{d}_1}{1+a_k\mathbf{d}_1}\right)\tfrac{c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{2^{k/2}} = \tfrac{5+\left(5a_k-2^{k/2+2}\right)\mathbf{d}_1}{(1+a_k\mathbf{d}_1)2^{k/2}}c_k\boldsymbol{\zeta}_k\mathbf{d}_1 \\
&\leq \quad \frac{5+\left(2^{k/2+1}\left(\frac{4^{k/2+1}}{3^{k/2}}-3\right)-2^{k/2+2}\right)2^{-k/2-1}}{(1+a_k\mathbf{d}_1)2^{k/2}}c_k\boldsymbol{\zeta}_k\mathbf{d}_1 = \tfrac{4c_k\boldsymbol{\zeta}_k}{2\cdot 3^{k/2}}\tfrac{2^{k/2+1}\mathbf{d}_1}{1+a_k\mathbf{d}_1},
\end{aligned}$$

and (38) is true for $m = 4, k \geq 3$. For $m \geq 5, k \geq 3$, invoking Lemma 19 and the assumption $\mathbf{d}_1 \leq 2^{-k/2-1}$, we ascertain that the left part of inequality (38) is bounded above by

$$(\sum_{j=3}^{m-1} \tfrac{(2^{-k/2})^{j-2}(1-\varepsilon_k)}{(m-j)^{k/2}} + \tfrac{m}{(m-2)^{k/2}})c_k\boldsymbol{\zeta}_k\mathbf{d}_1 \leq (\tfrac{2m(1-\frac{2^{k/2+1}\mathbf{d}_1}{1+a_k\mathbf{d}_1})}{5(m-1)^{k/2}} + \tfrac{m}{(m-2)^{k/2}})c_k\boldsymbol{\zeta}_k\mathbf{d}_1 \leq \tfrac{\eta_k\ mc_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-1)^{k/2}(1+a_k\mathbf{d}_1)}, \tag{39}$$

where $\eta_k = \frac{2}{5}\left(1 + a_k\mathbf{d}_1 - 2^{k/2+1}\mathbf{d}_1\right) + \frac{4^{k/2}}{3^{k/2}}(1 + a_k\mathbf{d}_1)$. Taking into account that $a_k \le 2^{k/2+1}$ we arrive at $\eta_k \le \frac{2}{5} + 2\frac{4^{k/2}}{3^{k/2}} \le 2^{k/2}$ for $k \ge 4$. Besides, $\eta_3 \le 2^{3/2}$ (since $\mathbf{d}_1 < 2^{-3/2-1}$) and hence $\eta_k \le 2^{k/2}$ for all $k \ge 3$ which ensures that (39) leads to (38) and the proof is completed. ■

Next, we employ twice Theorem 4 in order to get the improved bound of Theorem 5.

Proof. (**of Theorem 5**) From theorem's assumptions we have $\varepsilon_k < 1$, $\mathbf{d}_1 < 2^{-k/2-1}$ and therefore we can employ Theorem 4 for the r.v.'s $X_1, X_2, ...$ to get the inequality

$$\mathbf{d}_m \le \mathbf{D}_{k,m} = \frac{m\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k}{2(m-1)^{k/2}(1-\varepsilon_k)} + \frac{1}{2}(2\mathbf{d}_1)^m. \tag{40}$$

Let now $Y_i = m^{-1/2}\sum_{j=1}^m X_{(i-1)m+j}$, $i = 1, 2, ...$ and let $Z, Z_1, Z_2, ..., Z_m$ be i.i.d. standard normal r.v.'s. Observe that $Y_1, Y_2, ...$ are i.i.d. r.v.'s with $\mathcal{L}Y_i \in \mathcal{M}_k$. Moreover, we ascertain that $\varepsilon_{k,m} \le \varepsilon^*_{k,m} < 1$ and $\mathbf{d}_{1,m} \le \mathbf{D}_{k,m} < 2^{-k/2-1}$ where (see also (5)),

$$\begin{aligned}
&\mathbf{d}_{h,m} \overset{\mathbf{def}}{=} \mathbf{d}\Big(\tfrac{1}{\sqrt{h}}\sum_{i=1}^{h} Y_i, \mathcal{N}\Big) = \mathbf{d}\Big(\tfrac{1}{\sqrt{hm}}\sum_{i=1}^{hm} X_i, \mathcal{N}\Big) = \mathbf{d}_{hm},\\
&\boldsymbol{\zeta}_{k,m} \overset{\mathbf{def}}{=} \boldsymbol{\zeta}_k(Y_1, \mathcal{N}) = \boldsymbol{\zeta}_k\Big(\tfrac{1}{\sqrt{m}}\sum_{i=1}^{m} X_i, \mathcal{N}\Big) \le \tfrac{\boldsymbol{\zeta}_k}{m^{k/2-1}}, \quad \varepsilon_{k,m} \overset{\mathbf{def}}{=} \tfrac{2^{k/2+1}\mathbf{d}_{1,m}}{1+a_k\mathbf{d}_{1,m}} + \delta_k\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_{k,m}.
\end{aligned}$$

Therefore, we are eligible to apply Theorem 4 for the r.v.'s $Y_1, Y_2, ...$ and get

$$\mathbf{d}_{h,m} \le \frac{h\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_{k,m}}{2(h-1)^{k/2}\left(1-\varepsilon_{k,m}\right)} + \frac{1}{2}(2\mathbf{d}_{1,m})^h, \quad h \ge 2. \tag{41}$$

Since $\varepsilon_{k,m} \le \varepsilon^*_{k,m}$, $\mathbf{d}_{1,m} \le \mathbf{D}_{k,m}$ and $\boldsymbol{\zeta}_{k,m} \le \frac{\boldsymbol{\zeta}_k}{m^{k/2-1}}$ we finally conclude (14). ■

Proof. (**of Proposition 6**) Let $Z_1, Z_2, ...$ be a sequence of i.i.d. standard normal r.v.'s, independent also of X_i's. From the triangle inequality for $\mathbf{d}$ we deduce that,

$$\mathbf{d}_{s+j} = \mathbf{d}\left(\sum_{i=1}^{s+j} X_i, \sum_{i=1}^{s+j} Z_i\right) \le \mathbf{d}\left(\sum_{i=1}^{s+j} X_i, \sum_{i=1}^{s} Z_i + \sum_{i=s+1}^{s+j} X_i\right) + \mathbf{d}\left(\sum_{i=1}^{s} Z_i + \sum_{i=s+1}^{s+j} X_i, \sum_{i=1}^{s+j} Z_i\right). \tag{42}$$

Invoking the regularity property, the first term in the right part of (42) is less than or equal to $\mathbf{d}_s$. The proof is completed observing that, from Lemma 17 and the subadditivity property of the metric $\boldsymbol{\zeta}$, the second term in the right part of (42) is bounded above by,

$$\frac{1}{2}\left\|\varphi^{(k)}_{0,s}\right\|\boldsymbol{\zeta}_k\left(\sum_{i=s+1}^{s+j} X_i, \sum_{i=s+1}^{s+j} Z_i\right) \le \frac{\left\|\varphi^{(k)}\right\|}{2s^{k/2}}\sum_{i=s+1}^{s+j}\boldsymbol{\zeta}_k(X_i, Z_i) \le \frac{j\left\|\varphi^{(k)}\right\|}{2s^{k/2}}\boldsymbol{\zeta}_k.$$

■

Proof. (**of Corollary 7**) Taking into account that F_X has an absolutely continuous part, it follows that $\mathbf{d}_n \to 0$ as $n \to \infty$ (see Prokhorov (1952) or, for example, Sirazhdinov and Mamatov (1962)). Hence, we can choose an integer r_0 large enough so that, for all $r \ge r_0$, it holds $\mathbf{d}_r < 2^{-k/2-1}$ and $\varepsilon_{k,r} < 1$, where (see also (5)),

$$\varepsilon_{k,r} \overset{\mathbf{def}}{=} \frac{2^{k/2+1}\mathbf{d}_r}{1+a_k\mathbf{d}_r} + \delta_k\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k\left(\sum_{i=1}^{r}\frac{X_i}{\sqrt{r}}, \mathcal{N}\right) \le \frac{2^{k/2+1}\mathbf{d}_r}{1+a_k\mathbf{d}_r} + \frac{\delta_k\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k}{r^{k/2-1}}. \tag{43}$$

Next, set $Y = r^{-1/2}\sum_{i=1}^{r} X_i$ for some $r \geq r_0$ and let $Y_1, Y_2, ...$ be a sequence of independent copies of Y. It is easy to verify that $\mathcal{L}Y_i \in \mathcal{M}_k$ and since $r \geq r_0$ we have guaranteed that $\mathbf{d}(Y, \mathcal{N}) = \mathbf{d}_r < 2^{-k/2-1}$ and $\varepsilon_{k,r} < 1$. Therefore we can apply Theorem 4 for $Y_1, Y_2, ...$ and get,

$$\mathbf{d}_{mr} = \mathbf{d}\left(\sum_{i=1}^{rm} \tfrac{X_i}{\sqrt{rm}}, \mathcal{N}\right) = \mathbf{d}\left(\sum_{i=1}^{m} \tfrac{Y_i}{\sqrt{m}}, \mathcal{N}\right) \leq \tfrac{m\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k(Y,\mathcal{N})}{(m-1)^{k/2}2(1-\varepsilon_{k,r})} + \tfrac{(2\mathbf{d}_r)^m}{2}. \tag{44}$$

The above is valid for $m \geq 2, r \geq r_0$. Observe now that every integer n can be written as $n = mr+j$ with $m = r = [\sqrt{n}]$ (the integer part of $\sqrt{n}$) and some integer $j \in [0, 2\sqrt{n}]$. Therefore, for large enough $n : [\sqrt{n}] \geq \max\{2, r_0\}$, relation (44) combined with Proposition 6 yields

$$n^{\frac{k-2}{2}}\mathbf{d}_n = n^{\frac{k-2}{2}}\mathbf{d}_{mr+j} \leq n^{\frac{k-2}{2}}\mathbf{d}_{mr} + \tfrac{n^{\frac{k-1}{2}}\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k}{2(mr)^{\frac{k}{2}}} \leq \tfrac{mn^{\frac{k-2}{2}}\left\|\varphi^{(k)}\right\|\frac{\boldsymbol{\zeta}_k}{r^{k/2-1}}}{(m-1)^{\frac{k}{2}}2(1-\varepsilon_{k,r})} + \tfrac{n^{\frac{k-2}{2}}}{2^{\frac{mk}{2}+1}} + \tfrac{n^{\frac{k-1}{2}}\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k}{(mr)^{\frac{k}{2}}},$$

where $m = r = [\sqrt{n}]$. Finally, letting $n \to \infty$ it is easy to see that the right part of the above inequality converges to $\frac{1}{2}\left\|\varphi^{(k)}\right\|\boldsymbol{\zeta}_k$, which confirms the asymptotic relation (19). ■

Proof. (**of Corollary 8**) For $k = 3$ we have $\left\|\varphi^{(3)}\right\| = \frac{2+8e^{-3/2}}{\sqrt{2\pi}}$ and since $\mathcal{L}X \geq_{3-cx} \mathcal{N}$ or $\mathcal{L}X \leq_{3-cx} \mathcal{N}$, it follows that $\boldsymbol{\zeta}_3(X, \mathcal{N}) = \left|\mathbb{E}X^3\right|/3!$ (see (9)). If $\mathbb{E}X^3 = 0$ then necessarily $\mathcal{L}X = \mathcal{N}$ and thus also $\mathbf{d}_{hm} = 0$. Otherwise, we have

$$\lim_{h,m\to\infty}\sqrt{hm}\mathbf{D}^{*}_{3,h,m} = \tfrac{\left\|\varphi^{(3)}\right\|\boldsymbol{\zeta}_3}{2} = \tfrac{1+4e^{-3/2}}{6\sqrt{2\pi}}\left|\mathbb{E}X^3\right| > 0.$$

Since F_X has an absolutely continuous part and $|\mathbb{E}X^3| < \infty$, relation (20) ensures that $\sqrt{n}\mathbf{d}_n$ converges to the same quantity as the above and thus $\mathbf{d}_{hm} \sim \mathbf{D}^{*}_{3,h,m}$. ■

The following proof is analogous to that of Theorem 2.

Proof. (**of Theorem 9**) Let $Z_1, Z_2, ...$ be a sequence of i.i.d. standard normal r.v.'s, independent of X_i's. From the triangle inequality we deduce that,

$$\boldsymbol{\rho}_n = \boldsymbol{\rho}\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} Z_i\right) \leq \sum_{j=1}^{n}\boldsymbol{\rho}\left(\sum_{i=1}^{j} X_i + \sum_{i=j+1}^{n} Z_i, \sum_{i=1}^{j-1} X_i + \sum_{i=j}^{n} Z_i\right). \tag{45}$$

Employing Lemma 16 as in the proof of Theorem 2, we get for $n \geq 2, j = 1, 2, ..., n$,

$$\begin{aligned} &\boldsymbol{\rho}\left(\sum_{i=1}^{j} X_i + \sum_{i=j+1}^{n} Z_i, \sum_{i=1}^{j-1} X_i + \sum_{i=j}^{n} Z_i\right) \\ \leq\ & 2\mathbf{d}\left(X_j + \sum_{i=j+1}^{n} Z_i, Z_j + \sum_{i=j+1}^{n} Z_i\right)\boldsymbol{\rho}_{j-1} + \boldsymbol{\rho}\left(X_j + \sum_{i=1}^{n} Z_i - Z_j, \sum_{i=1}^{n} Z_i\right), \end{aligned} \tag{46}$$

where $\boldsymbol{\rho}_0 = 0$. Invoking Lemma 17, for $j = 1, 2, ..., n-1$, the right part of (46) is upper bounded by (see also (8))

$$\boldsymbol{\zeta}_k(X_j, Z_j)\left\|\varphi^{(k)}_{0,n-j}\right\|\boldsymbol{\rho}_{j-1} + \boldsymbol{\zeta}_k(X_j, Z_j)\left\|\varphi^{(k-1)}_{0,n-1}\right\|_\infty = \tfrac{\boldsymbol{\zeta}_k\left\|\varphi^{(k)}\right\|}{(n-j)^{k/2}}\boldsymbol{\rho}_{j-1} + \boldsymbol{\zeta}_k\tfrac{\left\|\varphi^{(k-1)}\right\|_\infty}{(n-1)^{k/2}}, \tag{47}$$

while for $j = n$, is bounded above by

$$2\mathbf{d}(X_n, Z_n)\boldsymbol{\rho}_{n-1} + \boldsymbol{\zeta}_k(X_n, Z_n)\left\|\varphi^{(k-1)}_{0,n-1}\right\|_\infty = 2\mathbf{d}_1\boldsymbol{\rho}_{n-1} + \boldsymbol{\zeta}_k\tfrac{\left\|\varphi^{(k-1)}\right\|_\infty}{(n-1)^{k/2}}. \tag{48}$$

Inequality (21) is now immediate from (45), (46), (47) and (48). ■

The proof of Theorem 11 is similar to the proof of Theorem 4 above, but this time it involves both the Kolmogorov and the total variation distance and thus the technical details are different.

Proof. (**of Theorem 11**) Denote $c_k = \left\|\varphi^{(k)}\right\|$ and $\epsilon_k = \left\|\varphi^{(k)}\right\|_\infty$. From Theorem 9,

$$\boldsymbol{\rho}_n \le c_k \boldsymbol{\zeta}_k \sum_{j=2}^{n-1} \tfrac{\boldsymbol{\rho}_{j-1}}{(n-j)^{k/2}} + 2\mathbf{d}_1 \boldsymbol{\rho}_{n-1} + \tfrac{n\epsilon_{k-1}\boldsymbol{\zeta}_k}{(n-1)^{k/2}},\ n = 2, 3, \ldots . \tag{49}$$

For $n = 2$, recalling also that $\boldsymbol{\rho}_1 \le \mathbf{d}_1$, inequality (49) reduces to,

$$\boldsymbol{\rho}_2 \le 2\epsilon_{k-1}\boldsymbol{\zeta}_k + 2\mathbf{d}_1\boldsymbol{\rho}_1 \le \tfrac{2\boldsymbol{\zeta}_k \epsilon_{k-1}}{1-\beta_k} + \tfrac{1}{2}(2\mathbf{d}_1)^2 = \mathbf{R}_{k,2},$$

and therefore (22) is valid for $n = 2$. Similarly, for $n = 3$, (49) yields

$$\begin{aligned} \boldsymbol{\rho}_3 &\le c_k\boldsymbol{\zeta}_k\boldsymbol{\rho}_1 + \tfrac{3\epsilon_{k-1}\boldsymbol{\zeta}_k}{2^{k/2}} + 2\mathbf{d}_1\boldsymbol{\rho}_2 \le c_k\boldsymbol{\zeta}_k\boldsymbol{\rho}_1 + \tfrac{3\epsilon_{k-1}\boldsymbol{\zeta}_k}{2^{k/2}} + 2\mathbf{d}_1\left(2\epsilon_{k-1}\boldsymbol{\zeta}_k + 2\mathbf{d}_1\boldsymbol{\rho}_1\right) \\ &\le \tfrac{c_k\boldsymbol{\zeta}_k\mathbf{d}_1 + 4\epsilon_{k-1}\boldsymbol{\zeta}_k\mathbf{d}_1}{(1-b_k\mathbf{d}_1)} + \tfrac{3\epsilon_{k-1}\boldsymbol{\zeta}_k}{2^{k/2}} + \tfrac{(2\mathbf{d}_1)^3}{2} = \tfrac{3\boldsymbol{\zeta}_k\epsilon_{k-1}}{2^{k/2}(1-b_k\mathbf{d}_1)} + \tfrac{(2\mathbf{d}_1)^3}{2} \le \mathbf{R}_{k,3}, \end{aligned}$$

and thus (22) is valid for $n = 3$. Assume now that (22) is valid for $n = 2, 3, \ldots, m-1$ ($m \ge 4$). We shall show that it is also valid for $n = m$. From (49),

$$\boldsymbol{\rho}_m \le \tfrac{c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-2)^{k/2}} + c_k\boldsymbol{\zeta}_k \sum_{j=3}^{m-1} \tfrac{\boldsymbol{\rho}_{j-1}}{(m-j)^{k/2}} + 2\mathbf{d}_1\boldsymbol{\rho}_{m-1} + \tfrac{m\epsilon_{k-1}\boldsymbol{\zeta}_k}{(m-1)^{k/2}},$$

(we also use that $\mathbf{d}_1 \le \boldsymbol{\rho}_1$) and hence, using the induction assumption,

$$\boldsymbol{\rho}_m \le \tfrac{c_k\boldsymbol{\zeta}_k\mathbf{d}_1}{(m-2)^{\frac{k}{2}}} + \sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{\frac{k}{2}}} \left(\tfrac{(j-1)\epsilon_{k-1}\boldsymbol{\zeta}_k}{(j-2)^{k/2}(1-\beta_k)} + \tfrac{(2\mathbf{d}_1)^{j-1}}{2}\right) + 2\mathbf{d}_1\left(\tfrac{(m-1)\epsilon_{k-1}\boldsymbol{\zeta}_k}{(m-2)^{k/2}(1-\beta_k)} + \tfrac{(2\mathbf{d}_1)^{m-1}}{2}\right) + \tfrac{m\epsilon_{k-1}\boldsymbol{\zeta}_k}{(m-1)^{\frac{k}{2}}}.$$

In order to show that the above quantity is less than or equal to $\mathbf{R}_{k,m}$ it suffices to show that

$$c_k\boldsymbol{\zeta}_k \sum_{j=3}^{m-1} \frac{\frac{(j-1)\epsilon_{k-1}\boldsymbol{\zeta}_k}{(j-2)^{k/2}(1-\beta_k)} + \frac{(2\mathbf{d}_1)^{j-1}}{2}}{(m-j)^{k/2}} + \boldsymbol{\zeta}_k\mathbf{d}_1\epsilon_{k-1}\frac{\frac{c_k}{\epsilon_{k-1}} + \frac{2(m-1)}{(1-\beta_k)}}{(m-2)^{k/2}} + \frac{m\epsilon_{k-1}\boldsymbol{\zeta}_k}{(m-1)^{k/2}} \le \frac{\frac{m}{(m-1)^{k/2}}\epsilon_{k-1}\boldsymbol{\zeta}_k}{1-\beta_k},$$

or, multiplying both sides by $1 - \beta_k \in (0, 1)$, it suffices to show that

$$c_k\boldsymbol{\zeta}_k \sum_{j=3}^{m-1} \frac{\frac{(j-1)\epsilon_{k-1}\boldsymbol{\zeta}_k}{(j-2)^{k/2}} + \frac{(1-\beta_k)(2\mathbf{d}_1)^{j-1}}{2}}{(m-j)^{k/2}} + \boldsymbol{\zeta}_k\mathbf{d}_1\epsilon_{k-1}\frac{\frac{c_k(1-\beta_k)}{\epsilon_{k-1}} + 2(m-1)}{(m-2)^{k/2}} \le \frac{m\epsilon_{k-1}\boldsymbol{\zeta}_k\beta_k}{(m-1)^{k/2}}.$$

The above relation is now readily derived by adding by parts the following two inequalities

$$\sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{k/2}} \tfrac{(j-1)\epsilon_{k-1}\boldsymbol{\zeta}_k}{(j-2)^{k/2}} \le \tfrac{m\epsilon_{k-1}\boldsymbol{\zeta}_k}{(m-1)^{k/2}} \delta_k c_k \boldsymbol{\zeta}_k, \tag{50}$$

$$\sum_{j=3}^{m-1} \tfrac{c_k\boldsymbol{\zeta}_k}{(m-j)^{k/2}} \tfrac{(1-\beta_k)(2\mathbf{d}_1)^{j-1}}{2} + \frac{\boldsymbol{\zeta}_k\mathbf{d}_1\epsilon_{k-1}\left(\frac{c_k(1-\beta_k)}{\epsilon_{k-1}} + 2(m-1)\right)}{(m-2)^{k/2}} \le \tfrac{m\epsilon_{k-1}\boldsymbol{\zeta}_k}{(m-1)^{k/2}} b_k \mathbf{d}_1. \tag{51}$$

Inequality (50) follows directly from Lemma 18 ($m \geq 4$). It now remains to prove (51). For $m = 4, k \geq 3$ the left part of (51) is equal to (we use the fact that $\beta_k > 0, \mathbf{d}_1 < 2^{-k/2-1}$),

$$c_k \boldsymbol{\zeta}_k \tfrac{1-\beta_k}{2} (2\mathbf{d}_1)^2 + \boldsymbol{\zeta}_k \mathbf{d}_1 \epsilon_{k-1} \frac{\frac{c_k(1-\beta_k)}{\epsilon_{k-1}}+6}{(4-2)^{k/2}} \leq \epsilon_{k-1} \boldsymbol{\zeta}_k \mathbf{d}_1 \frac{6+\frac{2c_k}{\epsilon_{k-1}}}{2^{k/2}} \leq \frac{4\epsilon_{k-1}\boldsymbol{\zeta}_k}{3^{k/2}} b_k \mathbf{d}_1,$$

and hence (51) is valid for $m = 4$. When $m \geq 5$, it suffices to show that

$$\tfrac{c_k}{\epsilon_{k-1}} \sum_{j=3}^{m-1} \tfrac{(2\mathbf{d}_1)^{j-2}}{(m-j)^{k/2}} + \frac{\frac{c_k}{\epsilon_{k-1}}+2(m-1)}{(m-2)^{k/2}} \leq \tfrac{m b_k}{(m-1)^{k/2}}.$$

The above follows from Lemma 19 and the fact that $\mathbf{d}_1 < 2^{-k/2-1}$. In detail,

$$\sum_{j=3}^{m-1} \tfrac{c_k (2\mathbf{d}_1)^{j-2}}{\epsilon_{k-1}(m-j)^{k/2}} + \frac{\frac{c_k}{\epsilon_{k-1}}+2(m-1)}{(m-2)^{k/2}} \leq \frac{2c_k m}{5\epsilon_{k-1}(m-1)^{\frac{k}{2}}} + \frac{\frac{c_k}{\epsilon_{k-1}}+2(m-1)}{(m-2)^{k/2}} \leq m \frac{\frac{2c_k}{5\epsilon_{k-1}}+\frac{4^{k/2}}{3^{k/2}}(\frac{c_k}{5\epsilon_{k-1}}+2)}{(m-1)^{k/2}} \leq \frac{m b_k}{(m-1)^{\frac{k}{2}}},$$

and the proof is completed. ■

Proof. (**of Theorem 12**) Since $\varepsilon_k < 1,\ \mathbf{d}_1 < 2^{-k/2-1}$, we apply Theorem 4 for the r.v.'s X_1, X_2,... and deduce (40). Consider the i.i.d. r.v.'s $Y_i = m^{-1/2} \sum_{j=1}^{m} X_{(i-1)m+j},\ i = 1, 2, ...$ and verify that $\mathcal{L}Y_i \in \mathcal{M}_k$. We observe that $\beta_{k,m} \leq \beta^*_{k,m} < 1,\ \mathbf{d}_{1,m} \leq \mathbf{D}_{k,m} < 2^{-k/2-1}$ where $\mathbf{d}_{h,m}, \boldsymbol{\zeta}_{k,m}$ are defined in the proof of Theorem 5 and

$$\boldsymbol{\rho}_{h,m} \overset{\text{def}}{=} \boldsymbol{\rho}\left(\tfrac{1}{\sqrt{h}} \sum_{i=1}^{h} Y_i, \mathcal{N}\right) = \boldsymbol{\rho}_{hm}, \qquad \beta_{k,m} \overset{\text{def}}{=} b_k \mathbf{d}_{1,m} + \delta_k \left\| \varphi^{(k)} \right\| \boldsymbol{\zeta}_{k,m}.$$

Hence, assumptions of Theorem 11 are satisfied for the r.v.'s $Y_1, Y_2, ...$ and thus

$$\boldsymbol{\rho}_{h,m} \leq \tfrac{h \left\|\varphi^{(k-1)}\right\|_\infty \boldsymbol{\zeta}_{k,m}}{(h-1)^{k/2}\left(1-\beta_{k,m}\right)} + \tfrac{(2\mathbf{d}_{1,m})^h}{2}, h \geq 2. \tag{52}$$

Since $\beta_{k,m} \leq \beta^*_{k,m}, \mathbf{d}_{1,m} \leq \mathbf{D}_{k,m}$ and $\boldsymbol{\zeta}_{k,m} \leq \frac{\boldsymbol{\zeta}_k}{m^{k/2-1}}$ we finally conclude (23). ■

Proof. (**of Corollary 15**) The fact that $\mathcal{L}X \geq_{k-cx} \mathcal{N}$ or $\mathcal{L}X \leq_{k-cx} \mathcal{N}$ guarantees that $\mathbb{E}X^i = \mu_i, i = 1, 2, ..., k-1$. If, in addition, $\mathbb{E}X^k = \mu_k$ then $\mathcal{L}X = \mathcal{N}$ and $\boldsymbol{\rho}_n = \mathbf{R}^*_{k,h,m} = 0$. We proceed for the non trivial case $\left|\mathbb{E}X^k - \mu_k\right| > 0$. Since $\mathcal{L}X \in \mathcal{M}_k$, the distribution of X has the same cumulants as $\mathcal{N}$ up to order $k-1$, that is $\gamma_1 = 0, \gamma_2 = 1$ and $\gamma_i = 0, i = 3, 4, ..., k-1$. It can also be verified that $\gamma_k = \mathbb{E}X^k - \mu_k$. From formula (28) we now ascertain that $Q_\nu(x) = 0$ for $v < k-2$, because $\gamma_i = 0, i = 3, 4, ..., k-1$. For $\nu = k-2$, the only non zero term of the sum in (28) is the one with $k_1 = 0, ..., k_{\nu-1} = 0, k_\nu = 1$. Therefore

$$Q_{k-2}(x) = -\tfrac{e^{-x^2/2}}{\sqrt{2\pi}} \mathcal{H}_{k-1}(x) \tfrac{\gamma_k}{k!} = -\tfrac{e^{-x^2/2}}{\sqrt{2\pi}} \mathcal{H}_{k-1}(x) \tfrac{\mathbb{E}X^k - \mu_k}{k!}. \tag{53}$$

Thus, from the Edgeworth expansion of F_n we deduce that,

$$F_n(x) - \Phi(x) = \sum_{v=1}^{k-2} \tfrac{Q_\nu(x)}{n^{\nu/2}} + o(\tfrac{1}{n^{k/2-1}}) = \tfrac{Q_{k-2}(x)}{n^{k/2-1}} + o(\tfrac{1}{n^{k/2-1}}), \tag{54}$$

uniformly in x. By virtue of the uniform convergence in (54), there exists a non-negative sequence $a_n, n \in \mathbb{N}$ independent of x, with $a_n \to 0$, such that

$$\left| \left| n^{k/2-1}(F_n(x) - \Phi(x)) \right| - \left| Q_{k-2}(x) \right| \right| \leq \left| n^{k/2-1}(F_n(x) - \Phi(x)) - Q_{k-2}(x) \right| \leq a_n,$$

and thus $|Q_{k-2}(x)| - a_n \le n^{k/2-1}\,|F_n(x) - \Phi(x)| \le |Q_{k-2}(x)| + a_n$, which reveals that

$$\tfrac{\sup_x |Q_{k-2}(x)|}{n^{k/2-1}} - \tfrac{a_n}{n^{k/2-1}} \le \boldsymbol{\rho}_n \le \tfrac{\sup_x |Q_{k-2}(x)|}{n^{k/2-1}} + \tfrac{a_n}{n^{k/2-1}}.$$

We finally ascertain that (see (53),(6))

$$\boldsymbol{\rho}_n = \tfrac{\sup_x |Q_{k-2}(x)|}{n^{k/2-1}} + o(\tfrac{1}{n^{k/2-1}}) = \tfrac{\left\|\varphi^{(k-1)}\right\|_\infty}{n^{k/2-1}} \tfrac{\left|\mathbb{E}X^k - \mu_k\right|}{k!} + o(\tfrac{1}{n^{k/2-1}}).$$

The above combined with (24) and (9) completes the proof . ■

5 Examples of applications

The purpose of this section is to illustrate the applicability and the performance of our main results. For this task, we compute the proposed normal approximation error bounds for standardized sums of i.i.d. r.v.'s (via both $\mathbf{d}$ and $\boldsymbol{\rho}$) considering three different distributions for the summands: (a) the *gamma distribution* (where $k = 3$), (b) the *uniform distribution* (where $k = 4$) and (c) a proper *mixture of two uniform distributions* ($k = 8$). The first two distributions were chosen because, for these cases, there exist in the literature sharp normal approximation error bounds comparable to ours, derived via different methods; see Adell and Lekuona (2008) for the gamma distribution, and Uspensky (1937) and Sherman (1971) for the uniform distribution. Moreover, in these cases we are able to calculate (numerically, at some acceptable accuracy) the exact Kolmogorov and total variation distances between the distribution of the standardized sum and $\mathcal{N}$. This made it possible to evaluate (numerically) the performance of our bounds compared to (i) other available ones and (ii) the exact value of the respective distances. The case (c) was chosen as a simple example where a fast convergence rate in the CLT can be observed ($O(n^{-3})$). All computations were easily performed using *Mathematica (Wolfram Research)* software

5.1 Normal approximation to Gamma Distribution ($k = 3$)

We shall examine the weak convergence of the sum of i.i.d. Y_i's when each Y_i follows the gamma distribution with parameters $a \in \{1, 2, ...\}, \lambda > 0$ (denoted by $\mathcal{G}_{a,\lambda}$). Thus $\mathbb{E}Y^k = \lambda^{-k}(a + k - 1)...(a + 1)a$. It suffices to study the case $\mathcal{L}Y_i = \mathcal{G}_{a,1}$ since then $\mathcal{L}Y_i/\lambda = \mathcal{G}_{a,\lambda}$. Considering the r.v.'s, $X_i = (Y_i - a)/\sqrt{a}$, we have $\mathbb{E}X_1 = 0, \mathbb{E}X_1^2 = 1, \mathbb{E}X_1^3 = 2a^{-1/2}$ and thus $\mathcal{L}X_1 \in \mathcal{M}_k$ for $k = 3$. We have that,

$$\boldsymbol{\zeta}_3 = \boldsymbol{\zeta}_3(X_1, \mathcal{N}) = \frac{\mathbb{E}X_1^3 - \mathbb{E}Z^3}{3!} = \frac{\mathbb{E}X_1^3}{3!} = \frac{1}{3a^{1/2}}, \tag{55}$$

where Z is a r.v. with $\mathcal{L}Z = \mathcal{N}$. The above follows from (9) and the fact that $X_1 \ge_{3-cx} Z$ which can be justified by the Karlin-Novikoff cut-criterion (see Remark 2). The relevant details are: if W is a r.v. such that $\mathcal{L}W = \mathcal{G}_{1,1}$ then $\mathbb{E}(W^i - (Z+1)^i) = 0, i = 1, 2$ and the difference of the p.d.f.s $f_{Z+1} - f_W$ changes sign 3 times (exhibits opposite signs on the intervals $(0, -1), (-1, q_1), (q_1, q_2), (q_2, \infty)$ where $q_1 = 2 - (3 - \ln 2\pi)^{1/2}, q_2 = 2 + (3 - \ln 2\pi)^{1/2}$), whereas the last sign of $f_W(x) - f_{Z+1}(x)$ is a $+$. Hence (see Remark 2) $Z + 1 \le_{3-cx} W$ and thus

$$X_1 =_d (\textstyle\sum_{i=1}^a W_i - a)/\sqrt{a} \ge_{3-cx} \textstyle\sum_{i=1}^a Z_i/\sqrt{a} =_d Z, \quad a = 1, 2, ...$$

where $Z_1, Z_2, ...$ and W_1, W_2 are independent copies of Z, W (note that s-convex orders are scale and shift invariant, and are closed under convolution).

If we choose $a = 1$ we compute via (4) that $\mathbf{d}_1 = \mathbf{d}\left(X_1, \mathcal{N}\right) \approx 0.319215$ which is not less than $2^{-3/2-1}$ and therefore we cannot directly apply Theorem 5. We can overcome this simply by choosing larger a. This is in agreement with the approach described in Remark 3 since the $a-$fold convolution of $\mathcal{G}_{1,1}$ is $\mathcal{G}_{a,1}$. For the rest of the example let us conveniently choose $a = 20$.

(i) **Closed form bound for the total variation distance**. We find numerically that $\mathbf{d}_1 \approx 0.0569934 < 2^{-3/2-1}$, while, by virtue of (55), we get $\boldsymbol{\zeta}_3 = 20^{-1/2}/3 \approx 0.0745356$. Thus, $\varepsilon_3 = \frac{2^{3/2+1}\mathbf{d}_1}{1+a_3\mathbf{d}_1} + \delta_3 \left\|\varphi^{(3)}\right\| \boldsymbol{\zeta}_3 \approx 0.588668 < 1$, where $\left\|\varphi^{(3)}\right\| = (2 + 8e^{-3/2})/\sqrt{2\pi}$ and the values of a_3, δ_3 are given in (12). Hence we can now apply Theorem 5. For $h \geq 2$ and m large enough such that $\varepsilon^*_{3,m} < 1, \mathbf{D}_{3,m} < 2^{-3/2-1}$, we derive

$$\mathbf{d}_{hm} \leq \mathbf{D}^*_{3,h,m} = \frac{(1-\frac{1}{h})^{-3/2}\left\|\varphi^{(3)}\right\| \boldsymbol{\zeta}_3}{2(hm)^{1/2}\left(1-\varepsilon^*_{3,m}\right)} + \frac{(2\mathbf{D}_{3,m})^h}{2}, \tag{56}$$

where $\varepsilon^*_{3,m} = \frac{2^{3/2+1}\mathbf{D}_{3,m}}{1+a_3\mathbf{D}_{3,m}} + \frac{\delta_3\left\|\varphi^{(3)}\right\|\boldsymbol{\zeta}_3}{m^{3/2-1}}$ and $\mathbf{D}_{3,m} = \frac{m}{(m-1)^{3/2}}\frac{\left\|\varphi^{(3)}\right\|\boldsymbol{\zeta}_3}{2(1-\varepsilon_3)} + \frac{(2\mathbf{d}_1)^m}{2}$. Moreover, as $h, m \to \infty$,

$$\mathbf{d}_{hm} \sim \mathbf{D}^*_{3,h,m} \sim \frac{\left\|\varphi^{(3)}\right\|}{2\sqrt{hm}}\frac{\mathbb{E}X_1^3}{3!} = \frac{2+8e^{-3/2}}{6\sqrt{2\pi}}\frac{1}{\sqrt{hma}} \approx \frac{0.25167}{\sqrt{hma}}.$$

and the bound (56) is asymptotically optimal. Equivalently, if $\mathcal{N}_{\mu,\sigma^2}$ denotes the normal distribution with mean μ and variance σ^2, we can write

$$\mathbf{d}_n = \mathbf{d}\left(\mathcal{G}_{n,\lambda},\ \mathcal{N}_{n/\lambda,n/\lambda^2}\right) \sim \mathbf{D}^*_{3,h,m} \sim \frac{2+8e^{-3/2}}{6\sqrt{2\pi}}\frac{1}{\sqrt{n}},$$

where the h, m in $\mathbf{D}^*_{3,h,m}$ are chosen so that $n = ahm$.

(ii) **Recursive bound for the total variation distance.** Alternatively, we can construct an upper bound for $\mathbf{d}_{hm}$ via the recursive relations described in Remark 1. More specifically we find an upper bound for $\mathbf{d}_m$ using the recursions (10), starting from $\mathbf{d}_1$ given above and $k = 3$, and then derive an upper bound of $\mathbf{d}^*_h (= \mathbf{d}_{hm})$ by initiating the recursive relations (16), now starting from $\mathbf{d}^*_1 = \mathbf{d}_m$. This recursive bound is always slightly better than (although asymptotically equal to) the closed form bound $\mathbf{D}^*_{3,h,m}$.

(iii) **Closed form bound for the Kolmogorov distance**: By Theorem 12, for $h \geq 2$ and m large enough such that $\beta^*_{3,m} < 1, \mathbf{D}_{3,m} < 2^{-3/2-1}$, we derive

$$\boldsymbol{\rho}_{hm} \leq \mathbf{R}^*_{3,h,m} = \frac{(1-\frac{1}{h})^{-3/2}}{(hm)^{3/2-1}}\frac{\left\|\varphi^{(2)}\right\|_\infty \boldsymbol{\zeta}_3}{1-\beta^*_{3,m}} + \frac{(2\mathbf{D}_{3,m})^h}{2}, \tag{57}$$

where $\beta^*_{3,m} = b_3\mathbf{D}_{3,m} + \delta_3\left\|\varphi^{(3)}\right\|\boldsymbol{\zeta}_3 m^{1-3/2},\ b_3 = \frac{2^{3/2}}{3}(4 + \left\|\varphi^{(3)}\right\| / \left\|\varphi^{(2)}\right\|_\infty)$, $\mathbf{D}_{3,m}$ is the same as in (i) above and $\left\|\varphi^{(2)}\right\|_\infty = (2\pi)^{-1/2}$. Moreover, as $h, m \to \infty$,

$$\boldsymbol{\rho}_{hm} \sim \mathbf{R}^*_{3,h,m} \sim \frac{\left\|\varphi^{(2)}\right\|_\infty}{\sqrt{hm}}\frac{\mathbb{E}X_1^3}{3!} = \frac{1}{3\sqrt{2\pi hma}} \approx \frac{0.13298}{\sqrt{ahm}},$$

and the bound (57) is asymptotically optimal. Equivalently, we can write

$$\boldsymbol{\rho}\left(\mathcal{G}_{n,\lambda},\ \mathcal{N}_{n/\lambda,n/\lambda^2}\right) = \boldsymbol{\rho}_{hm} \sim \mathbf{R}^*_{3,h,m} \sim \frac{1}{3\sqrt{2\pi n}},$$

where the h, m in $\mathbf{R}^*_{3,h,m}$ are chosen so that $n = ahm$.

(iv) **Recursive bound for the Kolmogorov distance.** According to Remark 4, we can bound $\mathbf{d}_m$ using the recursions (10) starting from $\mathbf{d}_1$, and then bound $\boldsymbol{\rho}^*_h (= \boldsymbol{\rho}_{hm})$ by initiating the recursions (25), now starting from $\boldsymbol{\rho}^*_1 \leq \mathbf{d}^*_1 = \mathbf{d}_m$.

(v) **Numerical comparisons**. In Table 1 we compute the bounds presented above for several values of the parameters. In this application, we can numerically evaluate the exact values of $\mathbf{d}_n$ and $\boldsymbol{\rho}_n$, since the sum under study is again gamma distributed. More specifically, in the left part of Table 1, we give the value of the *closed form bound* $\mathbf{D}^*_{3,h,m}$ (see 56), the *recursive bound* (see Paragraph (ii)), and the *exact value* of $\mathbf{d}_{hm}$ (numerically evaluated using (4)). In the right part of Table 1 we compute the analogous quantities concerning the Kolmogorov distance. That is, $\mathbf{R}^*_{3,h,m}$ (see 57), the *recursive bound* described in Paragraph (iv), and the *exact value* of $\boldsymbol{\rho}_{hm}$ evaluated numerically. Approximate values of the classical Berry-Esseen bound (labelled BE_n) and the (near optimal) bound of Adell and Lekuona (2008) (labelled AL_n) are also included in the right part of the table which concerns the Kolmogorov distance. The parameter $n = amh$ (where we have chosen $a = 20$) is also included.

Table 1 (*normal approximation error estimates for the gamma distribution*)

parameters		*Total variation distance*			*Kolmogorov distance*				
n amh	m, h	$\mathbf{D}^*_{3,h,m}$	recursive bound(ii)	$\mathbf{d}_{hm}$ (exact)	$\mathbf{R}^*_{3,h,m}$	recursive bound(iv)	$\boldsymbol{\rho}_{hm}$ (exact)	AL_n	BE_n
500	5,5	.032438	.022306	.011261	.036391	.012292	.005947	.015767	.076192
1000	10,5	.016964	.013900	.007961	.011165	.007523	.004205	.009180	.053876
10000	50,10	.003471	.003195	.002517	.001927	.001693	.001330	.001977	.017037
50000	100,25	.001340	.001261	.001126	.000730	.000667	.000595	.000794	.007619
100000	100,50	.000919	.000863	.000796	.000501	.000456	.000421	.000547	.005387

As it was expected by their asymptotic optimality, the bounds $\mathbf{D}^*_{3,h,m}$ and $\mathbf{R}^*_{3,h,m}$, as well as the recursive bounds (ii) and (iv) are very close to the true values of $\mathbf{d}_{hm}$ and $\boldsymbol{\rho}_{hm}$ respectively when n is large. The recursive bounds are slightly better but require more computing time. The bounds AL_n presented by Adell and Lekuona (2008) perform also very well, but asymptotically $AL_n \sim \frac{(6-e)/e}{3\sqrt{2\pi n}}$ while $\mathbf{R}^*_{3,h,m} \sim \frac{1}{3\sqrt{2\pi n}}$ where $(6-e)/e = 1.2073$. Finally, the classical Berry-Esseen bound, although it possesses the right order $O(n^{-1/2})$, it performs rather poorly since $BE_n = 0.7056\frac{2(6-e)/e}{\sqrt{n}}$, which, asymptotically, is nearly 13 times larger than the optimal $\mathbf{R}^*_{3,h,m}$.

5.2 Normal approximation to a sum of uniform r.v.'s ($k = 4$)

In this application we examine the weak convergence of the sum of i.i.d. r.v.'s $Y_1, Y_2, ...$ when every Y_i follows the uniform distribution in (0,1) (denoted by $\mathcal{U}_{0,1}$). Considering the standardized i.i.d. r.v.'s $W_i = (Y_i - 1/2)/\sqrt{1/12}, i = 1, 2, ...$ we have $\mathcal{L}W_1 \in \mathcal{M}_k$ for $k = 4$.The p.d.f. of W_i's is $f_W(x) = (2a)^{-1}, x \in (-a, a)$ where $a = \sqrt{3}$. We first need to verify that $\mathbf{d}_1 < 2^{-4/2-1} = 0.125$ and $\varepsilon_4 < 1$ so that Theorem 5 can be applied. The exact value of $\mathbf{d}\,(W_1, \mathcal{N})$ can be easily expressed in terms of the function Φ and its approximate value is 0.197678. Thus we cannot directly apply Theorem 5. According to Remark 3, we can alternatively apply Theorem 5 for the distribution of $(W_1 + ... + W_r)/\sqrt{r}$, for appropriate r. It suffices to take $r = 2$ and we set $X = (W_1 + W_2)/\sqrt{2}$. Denote by $X_1, X_2, ...$ a sequence of independent copies of X. Clearly, $\mathcal{L}X$ is the triangular distribution with support on the interval $(-\sqrt{6}, \sqrt{6})$ and mode at 0, equal to $6^{-1/2}$. We now find that $\mathbf{d}_1 = \mathbf{d}(X, \mathcal{N}) \approx 0.051247$. Moreover, we have

$$\boldsymbol{\zeta}_4 = \boldsymbol{\zeta}_4(X, \mathcal{N}) = \frac{\mathbb{E}Z^4 - \mathbb{E}X_1^4}{4!} = \frac{3 - \frac{12}{5}}{4!} = \frac{1}{40}.$$

where Z is a r.v. with $\mathcal{L}Z = \mathcal{N}$. The above equality follows from (9) because $Z \geq_{4-cx} X_1$. This can be easily verified by the Karlin-Novikoff cut-criterion (see Remark 2). Indeed, we have $\mathbb{E}W^i = \mu_i, i = 1, 2, 3$ and $f_Z(x) - f_W(x)$ changes sign 4 times while its last sign is a $+$. Hence $Z \geq_{4-cx} W$ and therefore $X_1 =_d (W_1 + W_2)/\sqrt{2} \leq_{4-cx} (Z_1 + Z_2)/\sqrt{2} =_d Z$, where Z_1, Z_2 are independent copies of Z (s-convex orders are scale invariant, and closed under convolution).

(i) **Closed form bound for the total variation distance**. We apply Theorem 5 for the r.v.'s $X_1, X_2, \ldots$ and $k = 4$ observing that $\mathbf{d}_1 \approx 0.051247 < 2^{-4/2-1}$, $\boldsymbol{\zeta}_4 = 1/40$ and $\varepsilon_4 \approx 0.621685 < 1$. For $h \geq 2$ and m large enough such that $\varepsilon^*_{4,m} < 1, \mathbf{D}_{4,m} < 2^{-4/2-1}$, it follows that

$$\mathbf{d}_{hm} \leq \mathbf{D}^*_{4,h,m} = \frac{(1 - \frac{1}{h})^{-2} \left\|\varphi^{(4)}\right\| \boldsymbol{\zeta}_4}{2(hm)\left(1 - \varepsilon^*_{4,m}\right)} + \frac{(2\mathbf{D}_{4,m})^h}{2}, \tag{58}$$

where $\varepsilon^*_{4,m} = \frac{8\mathbf{D}_{4,m}}{1+(296/45)\mathbf{D}_{4,m}} + \frac{9\left\|\varphi^{(4)}\right\|\boldsymbol{\zeta}_4}{2m}$, $\mathbf{D}_{4,m} = \frac{m}{(m-1)^2} \frac{\left\|\varphi^{(4)}\right\|\boldsymbol{\zeta}_4}{2(1-\varepsilon_4)} + \frac{(2\mathbf{d}_1)^m}{2}$ and $\left\|\varphi^{(4)}\right\|$ is given in (11). Moreover, as $h, m \to \infty$ (see (15)),

$$\mathbf{D}^*_{4,h,m} \sim \frac{\left\|\varphi^{(4)}\right\| \boldsymbol{\zeta}_4}{2hm} = \frac{\left\|\varphi^{(4)}\right\| \frac{3-\mathbb{E}X^4}{4!}}{2hm} = \frac{\left\|\varphi^{(4)}\right\| \frac{1}{40}}{2hm} \approx \frac{0.035008}{hm},$$

and since $n^{-1/2}\sum_{i=1}^{n} X_i =_d (2n)^{-1/2}\sum_{i=1}^{2n} \frac{Y_i-1/2}{\sqrt{1/12}}$ we can equivalently write

$$\mathbf{d}\left(\sum_{i=1}^{n} \frac{Y_i-1/2}{\sqrt{n/12}}, \mathcal{N}\right) = \mathbf{d}_{hm} \leq \mathbf{D}^*_{4,h,m} \sim \frac{\left\|\varphi^{(4)}\right\|}{40n} \approx \frac{0.070015}{n},$$

where h, m are chosen so that $n = 2hm$.

(ii) **Recursive bound for the total variation distance.** According to Remark 4, we first bound $\mathbf{d}_m$ using recursions (10) starting from $\mathbf{d}_1$ given above and then bound $\mathbf{d}^*_h (= \mathbf{d}_{hm})$ by initiating recursions (16), now starting from $\mathbf{d}^*_1 = \mathbf{d}_m$.

(iii) **Closed form bound for the Kolmogorov distance**: We apply Theorem 12 for $X_1, X_2, \ldots$ defined above. For $h \geq 2$ and m large enough such that $\beta^*_{4,m} < 1, \mathbf{D}_{4,m} < 2^{-4/2-1}$, we derive

$$\boldsymbol{\rho}_{hm} \leq \mathbf{R}^*_{4,h,m} = \frac{(1 - \frac{1}{h})^{-2}}{hm} \frac{\left\|\varphi^{(3)}\right\|_\infty \boldsymbol{\zeta}_4}{1 - \beta^*_{4,m}} + \frac{(2\mathbf{D}_{4,m})^h}{2}, \tag{59}$$

where $\beta^*_{4,m} = b_4\mathbf{D}_{4,m} + \frac{9}{2}\left\|\varphi^{(4)}\right\| \boldsymbol{\zeta}_4 m^{1-4/2}$, $b_4 = \frac{4}{3}\left(4 + \left\|\varphi^{(4)}\right\| / \left\|\varphi^{(3)}\right\|_\infty\right)$. Moreover,

$$\boldsymbol{\rho}_{hm} \sim \mathbf{R}^*_{4,h,m} \sim \frac{\left\|\varphi^{(3)}\right\|_\infty}{hm} \frac{\mu_4 - \mathbb{E}X^4}{4!} = \frac{e^{\frac{\sqrt{6}-3}{2}}\sqrt{\frac{3(3-\sqrt{6})}{\pi}}}{40hm}, \quad h, m \to \infty,$$

and the bound (59) is asymptotically optimal. Equivalently, we write

$$\boldsymbol{\rho}\left(\sum_{i=1}^{n} \frac{Y_i-1/2}{\sqrt{n/12}}, \mathcal{N}\right) = \boldsymbol{\rho}_{hm} \sim \mathbf{R}^*_{4,h,m} \sim \frac{\left\|\varphi^{(3)}\right\|_\infty \frac{3-\mathbb{E}W^4}{4!}}{n} = \frac{e^{\frac{\sqrt{6}-3}{2}}\sqrt{\frac{3(3-\sqrt{6})}{\pi}}}{20n} \approx \frac{0.027529}{n}, \tag{60}$$

where h, m in $\mathbf{R}^*_{4,h,m}$ are chosen so that $n = 2hm$.

(iv) **Recursive bound for $\boldsymbol{\rho}$.** We first bound $\mathbf{d}_m$ using recursions (10) starting from $\mathbf{d}_1$ and then bound $\boldsymbol{\rho}^*_h (= \boldsymbol{\rho}_{hm})$ using the recursions (25), starting from $\boldsymbol{\rho}^*_1 \le \mathbf{d}^*_1 = \mathbf{d}_m$.

(v) **Numerical comparisons**. We can numerically evaluate $\mathbf{d}_n$ and $\boldsymbol{\rho}_n$ in order to check the performance of our bounds by using the following well known exact formula for the p.d.f. g and c.d.f. G of a sum of n i.i.d. uniform (0,1) r.v.'s,

$$g_n(x) = \sum_{j=0}^{n} \tfrac{(-1)^j}{2(n-1)!}\tbinom{n}{j}(x-j)^{n-1} sgn(x-j), \quad G_n(x) = \sum_{j=0}^{n} \tfrac{(-1)^j}{2n!}\tbinom{n}{j}(x-j)^{n} sgn(x-j) + \tfrac{1}{2}. \quad (61)$$

It follows easily that the p.d.f. and c.d.f. of the standardized sum $(n/12)^{-1/2}\sum_{i=1}^n (Y_i - 1/2)$ is $f(x) = \sqrt{\frac{n}{12}} g_n(x\sqrt{\frac{n}{12}} + \frac{n}{2})$ and $F(x) = G_n(x\sqrt{\frac{n}{12}} + \frac{n}{2})$ respectively.

In the left part of Table 2 we compute the value of the *closed form bound* $\mathbf{D}^*_{4,h,m}$ (see 58), the *recursive bound* (see Paragraph (ii) above), and the *exact value* of $\mathbf{d}_{hm}$ (evaluated numerically via formula (4)). In the right part of Table 2 we include the analogous quantities concerning the Kolmogorov distance. That is, $\mathbf{R}^*_{4,h,m}$ (59), the *recursive bound* described in Paragraph (iv), and the *exact value* of $\boldsymbol{\rho}_{mh}$ evaluated numerically. Numeric values of the classical Berry-Esseen bound (labelled BE) and the bound of Uspensky (1937) (labelled US) as was improved by Sherman (1971) (see also Seoh (2002)), are also computed. The parameter $n = 2mh$ is also included in Table 2.

Table 2 (*normal approximation error estimates for the sum of uniform iid r.v.'s*)

parameters		*Total variation distance*			*Kolmogorov distance*				
n	m, h	$\mathbf{D}^*_{4,h,m}$	recursive bound(ii)	$\mathbf{d}_{hm}$ (exact)	$\mathbf{R}^*_{4,h,m}$	recursive bound(iv)	$\boldsymbol{\rho}_{hm}$ (exact)	US_n	BE_n
50	5,5	.002947	.002480	.001409	.001467	.000998	.000553	.000849	0.129627
100	10,5	.001238	.001152	.000702	.000518	.000456	.000276	.000424	0.091660
300	15,10	.000312	.000295	.000234	.000127	.000116	.000092	.000141	0.052920
500	25,10	.000181	.000175	.000140	.000072	.000069	.000055	.000085	0.040992
1000	25,20	.000081	.000078	.000070	.000032	.000031	.000028	.000042	0.028986
2000	50,20	.000040	.000039	.000035	.000016	.000015	.000014	.000021	0.020496

The closed form bounds $\mathbf{D}^*_{4,h,m}, \mathbf{R}^*_{4,h,m}$, and the recursive ones (ii),(iv), are very close to the exact $\mathbf{d}_{hm}, \boldsymbol{\rho}_{hm}$ respectively, especially for large n. The bound of Uspensky (1937) and Sherman (1971) performs very well, however it does not incorporate an asymptotically optimal constant since, $US_n \sim 2/(15\pi n) \approx 0.042441/n$ while $\mathbf{R}^*_{4,h,m} \approx 0.027529/n$ optimally (see (60)). Finally, the classical Berry-Esseen bound, $BE_n = 0.7056 \cdot 3\sqrt{3}/(4\sqrt{n})$, does not even possess the correct order $O(n^{-1})$ (as implied by the Edgeworth expansion) and, as was expected, it performs very poorly.

5.3 A simple example with $k = 8$

In this last example we present a simple case where a fast convergence rate in the CLT can be obtained. We consider a sum of i.i.d. r.v.'s $W_1, W_2, \ldots$, each distributed as $W = IU + (1 - I)V$, where U, V, I are independent r.v.'s, U, V are uniformly distributed in the intervals $(-c_1, c_1)$ and $(-c_2, c_2)$ respectively and I follows a Bernoulli distribution with parameter p. The symmetry of $\mathcal{L}W$ implies that $\mathbb{E}W^s = 0$ for s odd. Moreover,

$$\mathbb{E}W^s = p\mathbb{E}U^s + (1-p)\mathbb{E}V^s = \frac{pc_1^s + (1-p)c_2^s}{s+1} \text{ for } s \text{ even.}$$

Therefore, if we choose $c_1 = \sqrt{5-\sqrt{10}}, c_2 = \sqrt{5+\sqrt{10}}$ and $p = (5+\sqrt{10})/10$ we get that $\mathbb{E}W^2 = 1, \mathbb{E}W^4 = 3, \mathbb{E}W^6 = 15, \mathbb{E}W^8 = 275/3$ and therefore $\mathcal{L}W \in \mathcal{M}_k$ for $k = 8$.

Since $\mathbf{d}\,(W, \mathcal{N}) \approx 0.149156$ we cannot directly apply Theorem 5. Again, we alternatively apply Theorem 5 for $X = (W_1 + W_2)/\sqrt{2}$. From the fact that $Z \geq_{8-cx} X$, (9) yields

$$\boldsymbol{\zeta}_8 = \boldsymbol{\zeta}_8(X, \mathcal{N}) = \frac{\mathbb{E}Z^8 - \mathbb{E}X^8}{8!} = \frac{\frac{8!}{2^4 4!} - \frac{310}{3}}{8!} = \frac{5}{3 \cdot 8!},$$

where Z is a r.v. with $\mathcal{L}Z = \mathcal{N}$. The relation $Z \geq_{8-cx} X$ can be derived by observing that $\mathbb{E}W^i = \mu_i, i = 1, 2, ..., 7$, whereas $f_Z(x) - f_W(x)$ changes sign 8 times with the last sigh being a +. Hence (see Remark 2) $Z \geq_{8-cx} W$ and thus $X =_d (W_1 + W_2)/\sqrt{2} \leq_{8-cx} (Z_1 + Z_2)/\sqrt{2} =_d Z$ where Z_1, Z_2 are independent copies of Z.

(i) **Closed form bound for the total variation distance**. We apply Theorem 5 for the r.v.'s $X_1, X_2, ...$ (independent copies of X). We have $k = 8$, $\mathbf{d}_1 \approx 0.022161 < 2^{-8/2-1}$, $\boldsymbol{\zeta}_8 = \frac{5}{3 \cdot 8!}$ and $\varepsilon_8 \approx 0.575817 < 1$. For $h \geq 2$ and m large enough such that $\varepsilon^*_{8,m} < 1, \mathbf{D}_{8,m} < 2^{-5}$ we get

$$\mathbf{d}_{hm} \leq \mathbf{D}^*_{8,h,m} = \frac{(1-\frac{1}{h})^{-4} \left\|\varphi^{(8)}\right\| \boldsymbol{\zeta}_8}{2(hm)^3 \left(1-\varepsilon^*_{8,m}\right)} + \frac{(2\mathbf{D}_{8,m})^h}{2}, \tag{62}$$

where $\varepsilon^*_{8,m}$ and $\mathbf{D}_{8,m}$ are defined in Theorems 4, 5 and $\left\|\varphi^{(8)}\right\| \approx 96.11237$. Moreover

$$\mathbf{d}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i, \mathcal{N}\right) = \mathbf{d}_{hm} \leq \mathbf{D}^*_{8,h,m} \sim \frac{4\left\|\varphi^{(8)}\right\| \boldsymbol{\zeta}_8}{n^3} \approx \frac{0.015892}{n^3}, \tag{63}$$

where $h, m \to \infty$ and $n = 2hm$.

(ii) **Closed form bound for the Kolmogorov distance**: We apply Theorem 12 for $X_1, X_2, ...$. For $h \geq 2$ and m large enough such that $\beta^*_{8,m} < 1, \mathbf{D}_{8,m} < 2^{-5}$ we derive

$$\boldsymbol{\rho}_{hm} \leq \mathbf{R}^*_{8,h,m} = \frac{(1-\frac{1}{h})^{-4}}{(hm)^3} \frac{\left\|\varphi^{(7)}\right\|_\infty \boldsymbol{\zeta}_8}{1-\beta^*_{8,m}} + \frac{(2\mathbf{D}_{8,m})^h}{2}, \tag{64}$$

where $\beta^*_{8,m}$ is given in Theorem 12 and $\left\|\varphi^{(7)}\right\|_\infty \approx 14.177977$. Moreover,

$$\boldsymbol{\rho}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_i, \mathcal{N}\right) = \boldsymbol{\rho}_{hm} \sim \mathbf{R}^*_{8,h,m} \sim \frac{\left\|\varphi^{(7)}\right\|_\infty}{n^3} \frac{\mu_8 - \mathbb{E}W^8}{8!} \approx \frac{0.00468849}{n^3}, \tag{65}$$

where $h, m \to \infty$ and $n = 2hm$.

(iii) **Numerical comparisons**. It is not hard to verify that

$$P\left(\sum_{i=1}^{n} W_i \leq x\right) = \sum_{m=1}^{n-1} \binom{n}{m} p^m (1-p)^{n-m} \int_{-mc_1}^{mc_1} R_{n-m}(x-c) dH_m(c) + (1-p)^n R_n(x) + p^n H_n(x), \tag{66}$$

where $H_m(x) = G_m(\frac{x}{2c_1} + \frac{m}{2})$, $R_m(x) = G_m(\frac{x}{2c_2} + \frac{m}{2})$, ($G_n$ is given in (61)). From (66) and its derivative with respect to x we can numerically find $\mathbf{d}_n, \boldsymbol{\rho}_n$ for small n and up to some accuracy.

In the left part of Table 3 we compute the *closed form bound* $\mathbf{D}^*_{8,h,m}$ (see 62), the *recursive bound* (see Remark (4)), and the *exact value* of $\mathbf{d}_n$. In the right part of Table 3 we compute the analogous quantities concerning the Kolmogorov distance. The parameter $n = 2mh$ is also included. Since it was very hard to compute $\mathbf{d}_n, \boldsymbol{\rho}_n$ for n greater than about 30-40, we also compute in the table the asymptotic values of the bounds (see (63), (65)). The classical Berry-Esseen bound is of order $O(n^{-1/2})$ and there is no point in comparing it with $\mathbf{R}^*_{8,h,m}$, which is of order $O(n^{-3})$. (the values with a preceding "*" were computed by choosing $m = 2, h = n/4$).

Table 3 (*normal approximation error estimates*)

	Total variation distance				*Kolmogorov distance*			
$n\ (m,h)$	$\mathbf{D}^*_{8,h,m}$	recursive bound(ii)	$\mathbf{d}_n$ exact	(63) asympt	$\mathbf{R}^*_{8,h,m}$	recursive bound(iv)	$\boldsymbol{\rho}_n$ exact	(65) asympt
$12_{(1,6)}$	$4.47\ 10^{-5}$	$2.27\ 10^{-5}$	$1.07\ 10^{-5}$	$9.20\ 10^{-6}$	$^*4.01\ 10^{-5}$	$7.11\ 10^{-6}$	$3.14\ 10^{-6}$	$2.71\ 10^{-6}$
$24_{(1,12)}$	$^*3.26\ 10^{-6}$	$1.76\ 10^{-6}$	$1.24\ 10^{-6}$	$1.15\ 10^{-6}$	$^*1.83\ 10^{-6}$	$5.27\ 10^{-7}$	$3.65\ 10^{-7}$	$3.39\ 10^{-7}$
$32_{(1,16)}$	$^*1.13\ 10^{-6}$	$6.72\ 10^{-7}$		$4.85\ 10^{-7}$	$^*6.3\ 10^{-7}$	$2.00\ 10^{-7}$	$1.51\ 10^{-7}$	$1.43\ 10^{-7}$
$48_{(1,24)}$	$^*2.78\ 10^{-7}$	$1.81\ 10^{-7}$		$1.44\ 10^{-7}$	$^*1.56\ 10^{-7}$	$5.37\ 10^{-8}$	$4.42\ 10^{-8}$	$4.24\ 10^{-8}$
$200_{(5,20)}$	$2.45\ 10^{-9}$	$2.44\ 10^{-9}$		$1.99\ 10^{-9}$	$7.24\ 10^{-10}$	$7.19\ 10^{-10}$		$5.86\ 10^{-10}$
$1000_{(5,100)}$	$1.66\ 10^{-11}$	$1.65\ 10^{-11}$		$1.59\ 10^{-11}$	$4.91\ 10^{-12}$	$4.88\ 10^{-12}$		$4.69\ 10^{-12}$

References

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

[2] A.D. Barbour and L.H.Y.Chen (2005) *An introduction to Stein's method*, Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore, Singapore university press and world scientific publishing, Singapore.

[3] Barbour A.D. and Hall P. (1984) Stein's method and the Berry-Esseen theorem. *Australian J. Statist.* **26**, 8-15.

[4] Berry, A.C. (1941) The accuracy of Gaussian approximation to the sum of independent variates. *Trans. American Math. Soc.* **4**, 122-136.

[5] Boutsikas, M.V. and Vaggelatou, E. (2002). On the distance between convex-ordered random variables. *Adv. Appl. Probab.* **34**, 349-374.

[6] Boutsikas, M.V. and Vaggelatou, E. (2009) A new method for obtaining sharp compound Poisson approximation error estimates for sums of locally dependent random variables. *Bernoulli* (to appear)

[7] Cacoullos T., Papadatos N. and Papathanasiou V. (1997) Variance Inequalities for covariance kernels and applications to Central Limit Theorems. *Theory Probab. Appl.* **42**, 149-155.

[8] Chen L.H.Y. and Shao Q.M. (2001) A non-uniform Berry–Esseen bound via Stein's method. *Probab. Theory Related Fields*, **120**, 236–254.

[9] Denuit M., Lefèvre C. and Shaked M. (1998) The s-convex orders among real random variables, with applications. *Mathematical Inequalities & Applications* **1**, 585-613.

[10] Esseen, C.G. (1942). On the Liapounov limit of error in the theory of probability. *Ark. Mat. Astr. Fy.* **28**A, 1-19.

[11] Petrov, V.V. (1995). *Limit Theorems of Probability Theory.* Clarendon Press, Oxford.

[12] Prokhorov Yu V. (1952) On a local limit theorem for densities. *Dokl. Akad. Nauk SSSR*, **83**, 797-800.

[13] Rachev S.T. (1991) *Probability Metrics and the Stability of Stochastic Models.* Wiley, Chichester, New York.

[14] Rachev S.T. and Ruschendorf L. (1990) Approximation of sums by compound Poisson distributions with respect to stop-loss distances. *Advances in Applied Probability* **22**, 350-374.

[15] Senatov V.V. (1998). *Normal approximation: new results, methods and problems.* VSP (Utrecht, The Netherlands)

[16] Senatov V.V. (1980) Uniform estimates of the rate of convergence in the multi-dimensional CLT. *Theory Prob. Appl.* **25**, 745-759.

[17] Seoh M. (2002) On Improving Uspensky-Sherman's Normal Approximation by an Edgeworth-Expansion Approximation. *Bull. Belg. Math. Soc.* **9**, 65-72.

[18] Seoh, M. and Hallin, M. (1997) When does Edgeworth beat Berry and Esseen? Numerical evaluations of Edgeworth expansions. *J. Statist. Plann. Inference* **63**, 19–38.

[19] Shaked, M. and Shanthikumar, J.G. (1994) *Stochastic orders and their applications.* Academic Press, New York.

[20] Sherman, R. (1971). Error of the normal approximation to the sum of N random variables. *Biometrika*, **58**, 396-398.

[21] Shevtsova I.G. (2007) Sharpening of the upper bound of the absolute constant in the Berry–Esseen inequality. *Theory Probab. Appl.,* **51**, 549–553.

[22] Sirazhdinov S.Kh. and Mamatov M. (1962) On convergence in the mean for densities. *Theory Probab. Appl.*, **7**, 424-429.

[23] Stein, C. (1970) A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, *Proc. Sixth Berkeley Symp. Math. Stat. Prob.*, **2**, 583–602, Univ. California Press. Berkeley, Calif.

[24] Uspensky, J.V. (1937) *Introduction to Mathematical Probability.* McGraw-Hill, New York.

[25] Zolotarev V.M. (1983) Probability metrics. *Theory Prob. Appl.*, **28**, 278-302.

[26] Zolotarev V.M. (1997) *Modern theory of summation of random variables.* VSP

Michael V. Boutsikas
Postal address: Department of Statistics and Insurance Science, University of Piraeus,
Karaoli and Dimitriou str. 80, Piraeus 18534, Greece
email address: mbouts@unipi.gr