ESAIM: PS 25 (2021) 325–345
https://doi.org/10.1051/ps/2021012

ESAIM: Probability and Statistics
www.esaim-ps.org

NEW ERROR BOUNDS FOR LAPLACE APPROXIMATION *VIA* STEIN'S METHOD

Robert E. Gaunt*

Abstract. We use Stein's method to obtain explicit bounds on the rate of convergence for the Laplace approximation of two different sums of independent random variables; one being a random sum of mean zero random variables and the other being a deterministic sum of mean zero random variables in which the normalisation sequence is random. We make technical advances to the framework of Pike and Ren [*ALEA Lat. Am. J. Probab. Math. Stat.* **11** (2014) 571–587] for Stein's method for Laplace approximation, which allows us to give bounds in the Kolmogorov and Wasserstein metrics. Under the additional assumption of vanishing third moments, we obtain faster convergence rates in smooth test function metrics. As part of the derivation of our bounds for the Laplace approximation for the deterministic sum, we obtain new bounds for the solution, and its first two derivatives, of the Rayleigh Stein equation.

Mathematics Subject Classification. 60F05, 62E17.

Received October 4, 2020. Accepted June 28, 2021.

1. Introduction

The central limit theorem states that for a sequence of independent and identically distribution (i.i.d.) random variables, $X_1, X_2, \ldots$, with zero mean and variance $\sigma^2 \in (0,\infty)$, the standardised sum $W_n = \frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n X_i$ convergences in distribution to the standard normal distribution, as $n \to \infty$. By modifying the sum W_n appropriately such that either the number of terms in the sum is random or the normalisation is random we can instead naturally arrive at an asymptotic Laplace distribution. Studying the rate of convergence to the Laplace distribution in these two settings, *via* Stein's method, is the subject of this paper.

More precisely, consider the Laplace distribution with parameters $a \in \mathbb{R}$ and $b \in (0,\infty)$ with probability density function

$$f_W(x) = \frac{1}{2b}\mathrm{e}^{-\frac{|x-a|}{b}}, \quad x \in \mathbb{R}. \tag{1.1}$$

If a random variable W has density (1.1), then we write $W \sim \mathrm{Laplace}(a,b)$. It is readily checked that $\mathbb{E}[W] = a$ and $\mathrm{Var}(W) = 2b^2$. For a comprehensive account of the properties and applications of the Laplace distribution, see [28].

Keywords and phrases: Stein's method, Laplace approximation, rate of convergence, random sums, Rayleigh distribution.

Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK.

* Corresponding author: robert.gaunt@manchester.ac.uk

The first limit theorem we consider concerns geometric sums, which arise in a variety of settings [26]. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with zero mean and variance $\sigma^2 \in (0,\infty)$ and let $N_p \sim \mathrm{Geo}(p)$ be independent of the X_i with probability mass function $P(N_p = k) = p(1-p)^{k-1}$, $k = 1, 2, \ldots$, $0 < p < 1$. Then, with an obvious abuse of notation,

$$S_p := \sqrt{p}\sum_{i=1}^{N_p} X_i \to_d \mathrm{Laplace}(0, \frac{\sigma}{\sqrt{2}}), \quad p \to 0.$$

This result is proved under the stronger assumption of symmetric X_i in [28], whilst weaker Lindeberg-type conditions for the existence of the distributional limit are given by [47].

The second limit theorem considered in this paper concerns the case in which the sum $\sum_{i=1}^n X_i$ is normalised by a random variable. Let B_n be a beta random variable with parameters 1 and $n \geq 1$ and probability density function

$$f_{B_n}(x) = n(1-x)^{n-1}, \quad 0 < x < 1.$$

We write $B_n \sim \mathrm{Beta}(1, n)$. As in the first limit theorem, let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with zero mean and variance $\sigma^2 \in (0,\infty)$. For $n \geq 2$, let $B_{n-1} \sim \mathrm{Beta}(1, n-1)$ be independent of the X_i. Then, Proposition 2.2.12 of [28] states that

$$T_n := B_{n-1}^{1/2}\sum_{i=1}^{n} X_i \to_d \mathrm{Laplace}(0, \frac{\sigma}{\sqrt{2}}), \quad n \to \infty.$$

For characterisations of the Laplace distribution involving the random variables S_p and T_n, see [25, 33, 34], respectively.

In this paper, we give explicit bounds on the distance, with respect to certain probability metrics, between the distributions of S_p and T_n and their limiting Laplace distributions *via* Stein's method, a powerful probabilistic technique that was introduced in 1972 by Charles Stein [45] for normal approximation. For a given target distribution q, the first step in Stein's method is to find a suitable operator $\mathcal{A}$ acting on a class of functions $\mathcal{F}$ such that $\mathbb{E}[\mathcal{A}f(Y)] = 0$ for all $f \in \mathcal{F}$ if and only if the random variable Y has distribution q. For the $N(\mu, \sigma^2)$ distribution, the classical *Stein operator* is $\mathcal{A}f(x) = \sigma^2 f'(x) - (x-\mu)f(x)$. This leads to the *Stein equation*

$$\mathcal{A}f_h(x) = h(x) - \mathbb{E}[h(Y)], \tag{1.2}$$

where the test function h is real-valued. The second step is to solve (1.2) for f_h (for which we require $f_h \in \mathcal{F}$) and obtain suitable bounds for the solution. Finally, to approximate the distribution of a random variable of interest W by the target distribution q, one may evaluate both sides of (1.2) at W, take expectations, absolute values, and suprema of both sides over a class of functions $\mathcal{H}$ to obtain

$$d_{\mathcal{H}}(W, Y) := \sup_{h\in\mathcal{H}} |\mathbb{E}[h(W)] - \mathbb{E}[h(Y)]| = \sup_{h\in\mathcal{H}} |\mathbb{E}[\mathcal{A}f_h(W)]|.$$

This is of interest because many important probability metrics are of the form $d_{\mathcal{H}}(W, Y)$, and in many settings bounding the expectation $\mathbb{E}[Af_h(W)]$ is relatively tractable. In particular, taking

$$\begin{aligned}
\mathcal{H}_{\mathrm{K}} &= \{\mathbf{1}(\cdot \leq z) \,|\, z \in \mathbb{R}\}, \\
\mathcal{H}_{\mathrm{W}} &= \{h : \mathbb{R} \to \mathbb{R} \,|\, h \text{ is Lipschitz}, \|h'\| \leq 1\}, \\
\mathcal{H}_{\mathrm{BW}} &= \{h : \mathbb{R} \to \mathbb{R} \,|\, h \text{ is Lipschitz}, \|h\| \leq 1 \text{ and } \|h'\| \leq 1\},
\end{aligned}$$

$$\mathcal{H}_2 = \{h : \mathbb{R} \to \mathbb{R} \,|\, h' \text{ is Lipschitz, } \|h''\| \le 1\},$$
$$\mathcal{H}_{1,2} = \{h : \mathbb{R} \to \mathbb{R} \,|\, h' \text{ is Lipschitz, } \|h'\| \le 1 \text{ and } \|h''\| \le 1\}$$

gives the Kolmogorov, Wasserstein and bounded Wasserstein distances, which we denote by d_{K}, d_{W} and d_{BW}, respectively, as well as two smooth test function metrics, which we denote by d_2 and $d_{1,2}$, respectively. (Here and throughout the paper $\|g\| := \|g\|_\infty = \sup_{x\in\mathbb{R}} |g(x)|$.) The d_2 and $d_{1,2}$ and similar smooth test function metrics are often found in applications of Stein's method in which 'fast' convergence rates are sought, see, for example, [3, 13, 21, 23].

Stein's method was adapted to the Laplace distribution by [38] (a number of their contributions are outlined in Sect. 2), and as an application they derived an explicit bound on the bounded Wasserstein distance between the distribution of S_p and its limiting Laplace distribution. Their approach, which involves the introduction of the so-called *centered equilibrium transformation* for Laplace approximation, mirrored that of [35], who used Stein's method for exponential approximation to give explicit bounds on the rate of convergence in a generalisation of a well-known result of Rényi [40] concerning the convergence of geometric sums of positive random variables to the exponential distribution. In this paper, we make technical improvements on the work of [38] (through Lem. 2.1 and Thm. 2.5) that allow for their framework of Laplace approximation by Stein's method to yield optimal order Kolmogorov and Wasserstein distance bounds, as well as faster convergence rates in the d_2 distance. As an application we are able to obtain the following theorem.

Theorem 1.1. *Suppose $X_1, X_2, \dots$ is a sequence of independent random variables with $\mathbb{E}[X_i] = 0$ and $\mathbb{E}[X_i^2] = \sigma^2 \in (0,\infty)$. Let $N_p \sim \mathrm{Geo}(p)$, $0 < p < 1$, be independent of the X_i. Define $S_p = \sqrt{p}\sum_{i=1}^{N_p} X_i$ and let $Z \sim \mathrm{Laplace}(0, \frac{\sigma}{\sqrt{2}})$. Then*

$$d_{\mathrm{K}}(S_p, Z) \le \sqrt{2}\left(\frac{7}{2} + \sqrt{10}\right)\frac{\sqrt{p}}{\sigma}\sup_{i\ge1}\|F_{X_i}^{-1} - F_{X_i^L}^{-1}\|. \tag{1.3}$$

Suppose additionally that $\rho_3 = \sup_{i\ge1}\mathbb{E}[|X_i|^3] < \infty$. Then

$$d_{\mathrm{W}}(S_p, Z) \le 2\sigma\sqrt{p}\left(1 + \frac{\rho_3}{3\sigma^3}\right). \tag{1.4}$$

Let $k \ge 1$. Suppose that $\rho_{k+2} = \sup_{i\ge1}\mathbb{E}[|X_i|^{k+2}] < \infty$. Then

$$d_{\mathrm{K}}(S_p, Z) \le 11.56 \cdot 2^{\frac{k-1}{k+1}}(2p)^{\frac{k}{2(k+1)}}\left(\frac{\rho_k}{\sigma^k} + \frac{2\rho_{k+2}}{(k+1)(k+2)\sigma^{k+2}}\right)^{\frac{1}{(k+1)}}. \tag{1.5}$$

Finally, suppose that $X_1, X_2, \dots$ are identically distributed and that $\mathbb{E}[X_1^3] = 0$, $\mathbb{E}[X_1^4] < \infty$. Then

$$d_2(S_p, Z) \le \sigma^2 p\left[\frac{2-p}{1-p} + \frac{\mathbb{E}[X_1^4]}{6\sigma^4} + \frac{\sqrt{p}\log(1/p)}{\sqrt{2}(1-p)}\left(2 + \frac{\mathbb{E}[|X_1|^3]}{\sigma^3}\right)\right]. \tag{1.6}$$

Remark 1.2. The dependence on p in (1.5) is worse than in (1.3), but the bound may be preferable if $\sup_{i\ge1}\|F_{X_i}^{-1} - F_{X_i^L}^{-1}\|$ is difficult to compute or large. Note, though, that as k increases the exponent $\frac{k}{2(k+1)}$ of p in (1.5) approaches the exponent $\frac{1}{2}$ of (1.3).

We are also able to obtain a similar theorem for the deterministic sum T_n:

Theorem 1.3. *Let $n \geq 2$ and suppose that $X_1, \ldots, X_n$ are independent random variables with $\mathbb{E}[X_i] = 0$, $\mathbb{E}[X_i^2] = \sigma^2 \in (0, \infty)$ and $\mathbb{E}[|X_i|^3] < \infty$, for all $1 \leq i \leq n$. Then*

$$d_{\mathrm{K}}(T_n, Z) \leq \frac{0.5600}{\sigma^3 n^{3/2}} \sum_{i=1}^{n} \mathbb{E}[|X_i|^3] + \frac{1}{n}\left(1 + 2\left(1 - \frac{2}{n}\right)^{n-2}\right).$$

and

$$d_{\mathrm{W}}(T_n, Z) \leq \frac{2\sqrt{2}\sigma}{3n^{3/2}} \sum_{i=1}^{n} \left(2 + \frac{\mathbb{E}[|X_i|^3]}{\sigma^3}\right) + \frac{9.168\sigma}{n}.$$

In addition to the above assumptions, suppose that $\mathbb{E}[X_i^3] = 0$ and $\mathbb{E}[X_i^4] < \infty$, for all $1 \leq i \leq n$. Then

$$d_{1,2}(T_n, Z) \leq \frac{\sigma^2}{n^2} \sum_{i=1}^{n} \left(1 + \frac{\mathbb{E}[X_i^4]}{3\sigma^4}\right) + \frac{9.168\sigma}{n}.$$

Written in the notation of Theorem 1.1, the bounded Wasserstein distance bound of [38] reads $d_{\mathrm{BW}}(S_p, Z) \leq \sigma\sqrt{p}\big(1 + \frac{2\sqrt{2}}{\sigma}\big)\big(1 + \frac{\rho_3}{3\sigma^3}\big)$. We see that in addition to being given in a stronger metric, the Wasserstein distance bound (1.4) of Theorem 1.1 has a better dependence on σ (the bound of [38] has an extra factor of $\big(1 + \frac{2\sqrt{2}}{\sigma}\big)$ meaning that the bound has a worse dependence on σ if σ is 'small') and a smaller numerical constant if $\sigma < 2\sqrt{2}$ (the bound of [38] has the smaller numerical constant if $\sigma > 2\sqrt{2}$). The bound (1.4) also improves on the recent Wasserstein distance bound given in Theorem 5.10 of [18], in which Laplace approximations were obtained as part of a more general work on variance-gamma approximation. By working in a specialist Laplace framework, it is no surprise that we outperform the results of [18], and our Kolmogorov distance bound (1.3) is also an improvement on the analogous bound in Theorem 5.10 of that work. The $O(p)$ bound (1.6) is the first faster than $O(p^{1/2})$ bound for the random sum S_p in the literature. The faster convergence rate is a result of the vanishing third moment assumption, and as such complements a number of other 'matching moments' limit theorems that are found in the Stein's method literature, see, for example, [5, 13, 16, 19, 22, 29]. Theorem 1.3 gives the first bounds in the literature on the rate of convergence of the deterministic sum T_n to its asymptotic Laplace distribution. Again, under the assumption of vanishing third moments, we obtain a faster convergence rate. As part of our proof of the theorem, we obtain the first bounds in the literature for the solution, and its first two derivatives, of the Rayleigh Stein equation, which may be useful in future applications.

The rest of the paper is organised as follows. In Section 2, we obtain new bounds for the solution of the Laplace Stein equation (Lem. 2.1) and give general bounds for Laplace approximation involving the centered equilibrium distribution (Thm. 2.5). In Sections 3 and 4, we prove Theorems 1.1 and 1.3, respectively. In Section 5, we obtain new bounds for the solution of the Rayleigh Stein equation that are used in the proof of Theorem 1.3.

2. Stein's method for the Laplace distribution

In this section, we recall some of the theory developed by [38] for Stein's method for Laplace approximation and make some technical improvements that allow their framework for Laplace approximation to be applied in the Kolmogorov and Wasserstein metrics, as well as the d_2 metric when faster convergence rates are sought. We begin by recalling the following characterisation of the Laplace distribution ([38], Thm. 1.1).

Let W be a real-valued random variable. Then W follows the Laplace$(0, b)$ distribution if and only if

$$\mathbb{E}\big[b^2 f''(W) - f(W) + f(0)\big] = 0 \tag{2.1}$$

for all $f:\mathbb{R}\to\mathbb{R}$ such that f and f' are locally absolutely continuous and $\mathbb{E}|f'(Z)|<\infty$ and $\mathbb{E}|f''(Z)|<\infty$, for $Z\sim\mathrm{Laplace}(0,b)$. Based on this characterisation, [38] were led to the initial value problem

$$b^2f''(x)-f(x)=\tilde{h}(x),\quad f(0)=0, \tag{2.2}$$

where $\tilde{h}(x)=h(x)-\mathbb{E}[h(Z)]$, $Z\sim\mathrm{Laplace}(0,b)$.

At this point it is worth noting that an alternative Stein equation for the $\mathrm{Laplace}(0,b)$ distribution is given by

$$b^2xf''(x)+2b^2f'(x)-xf(x)=\tilde{h}(x), \tag{2.3}$$

which is a special case of the variance-gamma Stein equation of [15] (it is noted in Proposition 1.2 of [15] that the Laplace distribution is a special case of the variance-gamma distribution). A framework for variance-gamma approximation by Stein's method in the Kolmogorov and Wasserstein metrics was developed by [18], and a special case of this general framework gives a framework for Laplace approximation. However, the Stein equation (2.3) is more difficult to work with than (2.2) and it is therefore not surprising that all the comparable results for Laplace approximation obtained in this paper outperform those of [18]. We also remark that another Stein characterisation of the Laplace distribution is given by [1], as a special case of a general characterisation concerning infinitely divisible distributions, although the quantitative limit theorems derived in their work are quite different to ours.

Let us now focus on the initial value problem (2.2). The solution

$$f(x)=\frac{1}{2b}\bigg(\mathrm{e}^{x/b}\int_x^\infty\mathrm{e}^{-t/b}\tilde{h}(t)\,\mathrm{d}t+\mathrm{e}^{-x/b}\int_{-\infty}^x\mathrm{e}^{t/b}\tilde{h}(t)\,\mathrm{d}t\bigg) \tag{2.4}$$

was obtained by [38], as well as bounds for f and its first three derivatives. In the following lemma, we improve on Lemma 2.2 of [38] by obtaining bounds for f and its derivatives (of arbitrary order) that have smaller constants and hold for a larger class of functions. The latter improvement is crucial in enabling us to later obtain Kolmogorov and Wasserstein distance bounds for Laplace approximation.

Lemma 2.1. *Let $h:\mathbb{R}\to\mathbb{R}$ be a measurable function with $\mathbb{E}|h(Z)|<\infty$, where $Z\sim\mathrm{Laplace}(0,b)$. Let f be the solution (2.4) to the Stein equation (2.2). If h is bounded, then this is the unique bounded solution to (2.2). Moreover, the solution f and its first two derivatives satisfy the bounds*

$$\|f\|\le\|\tilde{h}\|,\quad \|f'\|\le\frac{1}{b}\|\tilde{h}\|,\quad \|f''\|\le\frac{2}{b^2}\|\tilde{h}\|. \tag{2.5}$$

Suppose that h is Lipschitz. Then

$$|f(x)|\le(2b+|x|)\|h'\|,\quad x\in\mathbb{R},$$

Now suppose that $h^{(k)}$ is Lipschitz, where $h^{(0)}\equiv h$. Then, for $k\ge0$,

$$\|f^{(k+1)}\|\le\|h^{(k+1)}\|,\quad \|f^{(k+2)}\|\le\frac{1}{b}\|h^{(k+1)}\|,\quad \|f^{(k+3)}\|\le\frac{2}{b^2}\|h^{(k+1)}\|. \tag{2.6}$$

Proof. It is easily verified that there is at most one bounded solution to (2.2). Suppose that u and v are solutions to (2.2). Then $w=u-v$ satisfies $w(0)=0$ and solves the differential equation $b^2w''(x)-w(x)=0$, the general solution to which is given by $w(x)=A\mathrm{e}^{x/b}+B\mathrm{e}^{-x/b}$. For $w(x)$ to be bounded for all $x\in\mathbb{R}$, we must take $A=B=0$, from which we conclude that $w=0$, so that $u=v$.

Now we establish the bounds in (2.5). Suppose h is bounded. We first note that, for all $x\in\mathbb{R}$,

$$\bigg|\mathrm{e}^{x/b}\int_x^\infty \mathrm{e}^{-t/b}\tilde{h}(t)\,\mathrm{d}t\bigg|\leq \|\tilde{h}\|\mathrm{e}^{x/b}\int_x^\infty \mathrm{e}^{-t/b}\,\mathrm{d}t=b\|\tilde{h}\|,$$

and

$$\bigg|\mathrm{e}^{-x/b}\int_{-\infty}^x \mathrm{e}^{t/b}\tilde{h}(t)\,\mathrm{d}t\bigg|\leq \|\tilde{h}\|\mathrm{e}^{-x/b}\int_{-\infty}^x \mathrm{e}^{t/b}\,\mathrm{d}t=b\|\tilde{h}\|.$$

Applying these inequalities into (2.4) gives the bound

$$\|f\|\leq\frac{1}{2b}\big(b\|\tilde{h}\|+b\|\tilde{h}\|\big)=\|\tilde{h}\|. \tag{2.7}$$

Differentiating both sides of (2.4) gives that

$$f'(x)=\frac{1}{2b}\bigg(\frac{1}{b}\mathrm{e}^{x/b}\int_x^\infty \mathrm{e}^{-t/b}\tilde{h}(t)\,\mathrm{d}t-\frac{1}{b}\mathrm{e}^{-x/b}\int_{-\infty}^x \mathrm{e}^{t/b}\tilde{h}(t)\,\mathrm{d}t\bigg), \tag{2.8}$$

and so

$$\|f'\|\leq\frac{1}{2b}\big(\|\tilde{h}\|+\|\tilde{h}\|\big)=\frac{1}{b}\|\tilde{h}\|.$$

From (2.2) and formula (2.4) we have that, for all $x\in\mathbb{R}$,

$$\begin{aligned}|f''(x)|&=\frac{1}{b^2}|\tilde{h}(x)+f(x)|\\&=\bigg|\frac{1}{b^2}\tilde{h}(x)+\frac{1}{2b^3}\bigg(\mathrm{e}^{x/b}\int_x^\infty \mathrm{e}^{-t/b}\tilde{h}(t)\,\mathrm{d}t+\mathrm{e}^{-x/b}\int_{-\infty}^x \mathrm{e}^{t/b}\tilde{h}(t)\,\mathrm{d}t\bigg)\bigg|\\&\leq\frac{1}{b^2}\|\tilde{h}\|+\frac{1}{2b^3}\big(b\|\tilde{h}\|+b\|\tilde{h}\|\big)=\frac{2}{b^2}\|\tilde{h}\|.\end{aligned}$$

Now we suppose that h is Lipschitz. We shall now prove the non-uniform bound for $|f(x)|$. By the mean value theorem, $|\tilde{h}(x)|\leq\|h'\|(|x|+\mathbb{E}|Z|)$, where $Z\sim\mathrm{Laplace}(0,b)$. Note that $\mathbb{E}|Z|=b$. Also, in anticipation of bounding $|f(x)|$ we note two integral inequalities: for $\lambda>0$ and $x\in\mathbb{R}$,

$$\mathrm{e}^{\lambda x}\int_x^\infty |t|\mathrm{e}^{-\lambda t}\,\mathrm{d}t<\frac{2}{\lambda^2}(1+\lambda|x|),\quad \mathrm{e}^{-\lambda x}\int_{-\infty}^x |t|\mathrm{e}^{\lambda t}\,\mathrm{d}t<\frac{2}{\lambda^2}(1+\lambda|x|).$$

We verify the first inequality; the second inequality is proved similarly. For $x\geq0$,

$$\mathrm{e}^{\lambda x}\int_x^\infty |t|\mathrm{e}^{-\lambda t}\,\mathrm{d}t=\frac{1}{\lambda^2}(1+\lambda x)$$

and, for $x<0$,

$$\mathrm{e}^{\lambda x}\int_x^\infty |t|\mathrm{e}^{-\lambda t}\,\mathrm{d}t=\mathrm{e}^{\lambda x}\bigg(-\int_x^0 t\mathrm{e}^{-\lambda t}\,\mathrm{d}t+\int_0^\infty t\mathrm{e}^{-\lambda t}\,\mathrm{d}t\bigg)$$

$$= \frac{1}{\lambda^2}\big(2\mathrm{e}^{\lambda x} - 1 - \lambda x\big) < \frac{1}{\lambda^2}(1 - \lambda x).$$

Putting all of the above together, we obtain, for $x \in \mathbb{R}$,

$$\begin{aligned} |f(x)| &\leq \frac{\|h'\|}{2b}\bigg(\mathrm{e}^{x/b}\int_x^\infty \mathrm{e}^{-t/b}(|t| + b)\,\mathrm{d}t + \mathrm{e}^{-x/b}\int_{-\infty}^x \mathrm{e}^{t/b}(|t| + b)\,\mathrm{d}t\bigg) \\ &\leq \frac{1}{2b}\bigg(2b^2\bigg(1 + \frac{|x|}{b}\bigg) + 2b^2\bigg) = \|h'\|(2b + |x|). \end{aligned}$$

Finally, we prove the uniform bounds. We note that applying integration by parts to (2.8) gives that

$$\begin{aligned} f'(x) &= \frac{1}{2b}\bigg\{\frac{1}{b}\mathrm{e}^{x/b}\bigg[b\mathrm{e}^{-x/b}\tilde{h}(x) + b\int_x^\infty \mathrm{e}^{-t/b}h'(t)\,\mathrm{d}t\bigg] \\ &\quad - \frac{1}{b}\mathrm{e}^{-x/b}\bigg[b\mathrm{e}^{x/b}\tilde{h}(x) - b\int_{-\infty}^x \mathrm{e}^{t/b}h'(t)\,\mathrm{d}t\bigg]\bigg\} \\ &= \frac{1}{2b}\bigg(\mathrm{e}^{x/b}\int_x^\infty \mathrm{e}^{-t/b}h'(t)\,\mathrm{d}t + \mathrm{e}^{-x/b}\int_{-\infty}^x \mathrm{e}^{t/b}h'(t)\,\mathrm{d}t\bigg). \end{aligned}$$

We recognise this representation of $f'(x)$ as being the same as the representation (2.4) of $f(x)$, with $\tilde{h}(t)$ replaced by $h'(t)$, and so we can immediately deduce the bounds in (2.6) for $\|f'\|$, $\|f''\|$ and $\|f^{(3)}\|$. Repeating the procedure inductively yields the bounds for $\|f^{(k+1)}\|$, $\|f^{(k+2)}\|$ and $\|f^{(k+3)}\|$, $k \geq 0$. □

The following distributional transformation, introduced by [38], is very natural in the context of Stein's method for Laplace approximation. Let W have mean zero and non-zero finite variance. Then we say that the random variable W^L has the *centered equilibrium distribution* with respect to W if

$$\mathbb{E}[f(W)] - f(0) = \frac{1}{2}\mathbb{E}[W^2]\mathbb{E}[f''(W^L)] \tag{2.9}$$

for all twice differentiable $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}|f(W)| < \infty$ and $\mathbb{E}|Wf'(W)| < \infty$. Stronger conditions were imposed on f by [38], but on examining the proof of their Theorem 3.2 it can be seen that the weaker conditions presented here are sufficient to ensure W^L exists and is unique. We also refer the reader to [7] for a generalisation of (2.9) to all random variables W with finite second moment, and we note that the centered equilibrium distribution is itself the Laplace analogue of the equilibrium distribution that is used in Stein's method for exponential approximation by [35]. Some useful properties of the centered equilibrium transformation are collected in Section 3 of [38] and Proposition 4.6 of [18]. In the sequel, the following moment relations will be important: assuming $\mathbb{E}[W^2] = 2b^2$, we have that, for $r \geq 0$,

$$\mathbb{E}[(W^L)^r] = \frac{\mathbb{E}[W^{r+2}]}{(r+1)(r+2)b^2}, \quad \mathbb{E}[|W^L|^r] = \frac{\mathbb{E}[|W|^{r+2}]}{(r+1)(r+2)b^2}. \tag{2.10}$$

The formulas in (2.10) are obtained by substituting $f_1(w) = w^{r+2}$ and $f_2(w) = |w|^{r+2}$, respectively, into (2.9) and using that $\mathbb{E}[W^2] = 2b^2$.

Theorem 2.5 below gives general bounds for Laplace approximation involving the centered equilibrium transformation. Bounds (2.15)–(2.19) of the theorem are the Laplace analogues of the bounds of Theorem 2.1 of [35], which give Kolmogorov and Wasserstein distance bounds in terms the absolute difference between a random variable W and its W-equilibrium transformation. We additionally provide a bound in the weaker d_2 metric, which is used to obtain the $O(p^{-1})$ bound (1.6) of Theorem 1.1. We mostly follow the approach of [35], but the approach used to obtain the d_2 metric bound is similar to that used by ([22], Thm. 3.1) to prove an analogous bound for the zero bias transformation. We begin by stating three lemmas. The proofs of Lemmas 2.2 and 2.4 are simple and hence omitted, and the proof of Lemma 2.3 follows immediately from the estimates of Lemma 2.1.

Lemma 2.2. *Let $Z \sim \mathrm{Laplace}(0, b)$. Then, for any random variable W,*

$$\mathbb{P}(\alpha \leq W \leq \beta) \leq \frac{\beta - \alpha}{2b} + 2d_{\mathrm{K}}(W, Z). \tag{2.11}$$

Lemma 2.3. *For any $a \in \mathbb{R}$ and any $\epsilon > 0$, define*

$$h_{a,\epsilon}(x) := \epsilon^{-1} \int_0^{\epsilon} \mathbf{1}(x + s \leq a)\,\mathrm{d}s. \tag{2.12}$$

Let $f_{a,\epsilon}$ be the solution (2.4) with test function $h_{a,\epsilon}$. Let $h_{a,0}(x) = \mathbf{1}(x \leq a)$ and define $f_{a,0}$ accordingly. Then

$$\|f_{a,\epsilon}\| \leq 1, \tag{2.13}$$

$$\|f'_{a,\epsilon}\| \leq \frac{1}{b}, \tag{2.14}$$

$$\|f''_{a,\epsilon}\| \leq \frac{2}{b^2}.$$

Lemma 2.4. *Let W be a real-valued random variable and let $Z \sim \mathrm{Laplace}(0, b)$. Then, for any $\epsilon > 0$,*

$$d_{\mathrm{K}}(W, Z) \leq \frac{\epsilon}{2b} + \sup_{a \in \mathbb{R}} |\mathbb{E}[h_{a,\epsilon}(W)] - \mathbb{E}[h_{a,\epsilon}(Z)]|,$$

with $h_{a,\epsilon}$ defined as in Lemma 2.3.

Theorem 2.5. *Let W be random variable with zero mean and variance $2b^2 \in (0, \infty)$, and let W^L have the W-centered equilibrium distribution. Then, for any $\beta > 0$,*

$$d_{\mathrm{K}}(W, Z) \leq \frac{(7/2 + \sqrt{10})\beta}{b} + 3\bigg(1 + \sqrt{\frac{2}{5}}\bigg)\mathbb{P}(|W - W^L| > \beta), \tag{2.15}$$

$$d_{\mathrm{K}}(W^L, Z) \leq \frac{\beta}{b} + 2\mathbb{P}(|W - W^L| > \beta). \tag{2.16}$$

Suppose further that $\mathbb{E}[|W|^3] < \infty$. Then

$$d_{\mathrm{W}}(W, Z) \leq 2\mathbb{E}|W - W^L|, \tag{2.17}$$

$$d_{\mathrm{W}}(W^L, Z) \leq \mathbb{E}|W - W^L|, \tag{2.18}$$

$$d_{\mathrm{K}}(W^L, Z) \leq \frac{1}{b}\mathbb{E}|W - W^L|. \tag{2.19}$$

Suppose now that $\mathbb{E}[W^4] < \infty$. Then

$$d_2(W, Z) \leq b\mathbb{E}[|\mathbb{E}[W - W^L \,|\, W]|] + \mathbb{E}[(W - W^L)^2]. \tag{2.20}$$

Remark 2.6. Analogues of inequalities (2.15)–(2.19) for variance-gamma approximation were given in Theorem 4.10 of [18], which as special cases give bounds for Laplace approximation in terms of the centered equilibrium distribution. In all cases, our bounds improve on the bounds of [18].

Proof. For ease of notation, we let $\kappa = d_{\mathrm{K}}(W,Z)$. We also let $\Delta := W - W^L$ and $I_1 := \mathbf{1}(|\Delta| \le \beta)$. Let f be the solution of the Laplace$(0,b)$ Stein equation with test function $h_{a,\epsilon}$, as given in (2.12). Note that the expectation $\mathbb{E}[f''(W^L)]$ is well defined, since $\|f''\| < \infty$ (see Lem. 2.3). By the Laplace Stein equation (2.2), we have

$$\begin{aligned}\mathbb{E}[h(W)] - \mathbb{E}[h(Z)] &= \mathbb{E}[b^2 f''(W) - f(W)] \\ &= b^2\mathbb{E}[I_1(f''(W) - f''(W^L))] + b^2\mathbb{E}[(1-I_1)(f''(W) - f''(W^L))] \\ &=: J_1 + J_2.\end{aligned}$$

Using the bound (2.13) we have

$$\begin{aligned}|J_2| &= |\mathbb{E}[(1-I_1)(f(W) - f(W^L) + \tilde{h}_{a,\epsilon}(W) - \tilde{h}_{a,\epsilon}(W^L))]| \\ &= |\mathbb{E}[(1-I_1)(f(W) - f(W^L) + h_{a,\epsilon}(W) - h_{a,\epsilon}(W^L))]| \\ &\le (2\|f\| + 1)\mathbb{P}(|\Delta| > \beta) \\ &\le 3\mathbb{P}(|\Delta| > \beta).\end{aligned}$$

We also have

$$\begin{aligned}J_1 &= \mathbb{E}\bigg[I_1 \int_0^{-\Delta} b^2 f^{(3)}(W+t)\,\mathrm{d}t\bigg] \\ &= \mathbb{E}\bigg[I_1 \int_0^{-\Delta} \big\{f'(W+t) - \epsilon^{-1}\mathbf{1}(a-\epsilon \le W+t \le a)\big\}\,\mathrm{d}t\bigg] \\ &\le \|f'\|\mathbb{E}|I_1\Delta| + \epsilon^{-1}\int_{-\beta}^{0} \mathbb{P}(a-\epsilon \le W+t \le a)\,\mathrm{d}t \\ &\le \frac{\beta}{b} + \frac{\beta}{2b} + 2\beta\epsilon^{-1}\kappa = \frac{3\beta}{2b} + 2\beta\epsilon^{-1}\kappa,\end{aligned}$$

where we used inequality (2.14) and Lemma 2.2 to obtain the last inequality. By a similar argument,

$$J_1 \ge -\frac{3\beta}{2b} - 2\beta\epsilon^{-1}\kappa,$$

and so we conclude that

$$|J_1| \le \frac{3\beta}{2b} + 2\beta\epsilon^{-1}\kappa.$$

We now apply Lemma 2.4 and take the convenient choice $\epsilon = \eta\beta$, $\eta > 2$, to obtain

$$\kappa \le 3\mathbb{P}(|\Delta| > \beta) + \frac{3\beta + \epsilon}{2b} + 2\beta\epsilon^{-1}\kappa = 3\mathbb{P}(|\Delta| > \beta) + \frac{(3+\eta)\beta}{2b} + \frac{2\kappa}{\eta},$$

which on rearranging yields

$$\kappa \le \frac{3\eta}{\eta - 2}\mathbb{P}(|\Delta| > \beta) + \frac{(3\eta + \eta^2)\beta}{2b(\eta - 2)}. \tag{2.21}$$

Choosing $\eta = 2 + \sqrt{10}$ minimises the second term in (2.21) and yields the bound (2.15). We elected to minimise the second term because in some applications the first term vanishes; as an example, see the proof of inequality (3.2).

Now we prove inequality (2.16). We have

$$\begin{aligned}\mathbb{E}[b^2f''(W^L)-f(W^L)]&=\mathbb{E}[f(W)-f(W^L)]\\&=\mathbb{E}[I_1(f(W)-f(W^L))]+\mathbb{E}[(1-I_1)(f(W)-f(W^L))].\end{aligned}$$

By the mean value theorem, applying the triangle inequality and then using the bounds (2.13) and (2.14) we obtain

$$\begin{aligned}\mathbb{E}[b^2f''(W^L)-f(W^L)]&\le\|f'\|\mathbb{E}|I_1\Delta|+2\|f\|\mathbb{P}(|\Delta|>\beta)\\&\le\frac{\beta}{b}+2\mathbb{P}(|\Delta|>\beta),\end{aligned}$$

yielding inequality (2.16).

Now suppose that $\mathbb{E}[|W|^3]<\infty$. By the absolute moment relation (2.10), this assumption guarantees that $\mathbb{E}|W^L|<\infty$. Let $h\in\mathcal{H}_{\mathrm{W}}$. We have

$$\begin{aligned}|\mathbb{E}[h(W)]-\mathbb{E}[h(Z)]|&=|\mathbb{E}[b^2f''(W)-f(W)]|=b^2|\mathbb{E}[f''(W)-f''(W^L)]|\\&\le b^2\|f^{(3)}\|\mathbb{E}|W-W^L|\le 2\mathbb{E}|W-W^L|,\end{aligned}$$

where we used the bound $\|f^{(3)}\|\le\frac{2}{b^2}\|h'\|$ of Lemma 2.1 in the final step. This proves inequality (2.17). Also,

$$\begin{aligned}\big|\mathbb{E}\big[b^2f''(W^L)-f(W^L)\big]\big|&=\big|\mathbb{E}f(W)-\mathbb{E}f(W^L)\big|\\&\le\|f'\|\mathbb{E}|W-W^L|.\end{aligned}\tag{2.22}$$

Using inequality $\|f'\|\le\|h'\|$ of Lemma 2.1 to (2.22) gives (2.18). Suppose now that $h\in\mathcal{H}_{\mathrm{K}}$. Then using the bound $\|f'\|\le\frac{1}{b}\|h\|$ gives us (2.19).

Finally, let $h\in\mathcal{H}_2$. Suppose that $\mathbb{E}[W^4]<\infty$, which, by the moment relation (2.10), ensures that $\mathbb{E}[(W^L)^2]<\infty$. By Taylor expansion we have

$$\begin{aligned}|\mathbb{E}[b^2f''(W)-f(W)]|&=b^2|\mathbb{E}[f''(W)-f''(W^L)]|\\&\le b^2|\mathbb{E}[f^{(3)}(W)(W-W^L)]|+\frac{b^2}{2}\|f^{(4)}\|\mathbb{E}[(W-W^L)^2]\\&=b^2|\mathbb{E}[f^{(3)}(W)\mathbb{E}[W-W^L\,|\,W]]|+\frac{b^2}{2}\|f^{(4)}\|\mathbb{E}[(W-W^L)^2]\\&\le b^2\|f^{(3)}\||\mathbb{E}[|\mathbb{E}[W-W^L\,|\,W]|]+\frac{b^2}{2}\|f^{(4)}\|\mathbb{E}[(W-W^L)^2].\end{aligned}$$

Applying the bounds $\|f^{(3)}\|\le\frac{1}{b}\|h''\|$ and $\|f^{(4)}\|\le\frac{2}{b^2}\|h''\|$ from Lemma 2.1 then yields the bound (2.20), as required. □

Corollary 2.7. *Let $k\ge1$ and suppose that $\mathbb{E}[|W|^{k+2}]<\infty$. Then*

$$d_{\mathrm{K}}(W,Z)\le 11.56\bigg(\frac{\mathbb{E}[|W-W^L|^k]}{b^k}\bigg)^{1/(k+1)}.\tag{2.23}$$

Proof. Applying Markov's inequality to (2.15) gives

$$d_{\mathrm{K}}(W,Z) \leq \frac{(7/2+\sqrt{10})\beta}{b} + 3\left(1+\sqrt{\frac{2}{5}}\right)\frac{\mathbb{E}[|W-W^L|^k]}{\beta^k},$$

whence on setting $\beta = (b\mathbb{E}[|W-W^L|^k])^{1/(k+1)}$ we obtain (2.23). □

3. Proof of Theorem 1.1

We begin by proving the following general theorem, which improves on Theorem 4.4 of [38] and Theorem 5.9 of [18]. The improvement comes from smaller constants than in both of those theorems and by giving the bounds in metrics stronger than the bounded Wasserstein metric bounds of [38]. Very recently, [37] have obtained an optimal order Wasserstein distance bound for a multivariate generalisation of the following theorem. In their result $X_1, X_2, \dots$ are i.i.d. random vectors, the limiting distribution is a centered multivariate symmetric Laplace distribution (see [28]) and an explicit constant is not given in their bound.

Theorem 3.1. *Suppose that $X_1, X_2, \dots$ is a sequence of independent random variables, with $\mathbb{E}[X_i] = 0$ and $\mathbb{E}[X_i^2] = \sigma_i^2 \in (0,\infty)$. Let N be a positive, integer-valued random variable with finite mean μ, which is independent of the X_i. Define $\sigma^2 = \frac{1}{\mu}\mathbb{E}\big[\big(\sum_{i=1}^N X_i\big)^2\big] = \frac{1}{\mu}\mathbb{E}\big[\sum_{i=1}^N \sigma_i^2\big]$. Also, let M be a random variable satisfying*

$$\mathbb{P}(M=m) = \frac{\sigma_m^2}{\mu\sigma^2}\mathbb{P}(N \geq m), \quad m = 1,2,\dots.$$

Define $W_\mu = \frac{1}{\sqrt{\mu}}\sum_{i=1}^N X_i$ and let $Z \sim \mathrm{Laplace}(0, \frac{\sigma}{\sqrt{2}})$. Then

$$d_{\mathrm{W}}(W_\mu, Z) \leq 2\mu^{-1/2}\big\{\mathbb{E}|X_M - X_M^L| + \sup_{i\geq 1}\sigma_i \mathbb{E}\big[|N-M|^{\frac{1}{2}}\big]\big\}. \tag{3.1}$$

Now suppose that $|X_i| \leq C$ for all i and $|N-M| \leq K$. Then we have

$$d_{\mathrm{K}}(W_\mu, Z) \leq \frac{\sqrt{2}(7/2+\sqrt{10})}{\sigma\sqrt{\mu}}\Big\{\sup_{i\geq 1}\|F_{X_i}^{-1} - F_{X_i^L}^{-1}\| + CK\Big\}, \tag{3.2}$$

and if $K = 0$ the bound also holds for unbounded X_i.

Proof. It was shown in the proof of Theorem 4.4 of [38] that $W_\mu^L = \mu^{-1/2}\big(\sum_{i=1}^{M-1} X_i + X_M^L\big)$. We take X_m^L to be independent of M, N, and X_k for all k. Therefore

$$W_\mu^L - W_\mu = \mu^{-1/2}\Bigg\{(X_M^L - X_M) + \mathrm{sgn}(M-N)\sum_{i=(M\wedge N)+1}^{N\vee M} X_i\Bigg\}.$$

Substituting into (2.17) and bounding $\mathbb{E}\big|\sum_{i=(M\wedge N)+1}^{N\vee M} X_i\big| \leq \sup_{i\geq 1}\sigma_i\mathbb{E}\big[|N-M|^{1/2}\big]$ (see the proof of Theorem 4.4 of [38]) gives us (3.1). Recall from (2.15) that

$$d_{\mathrm{K}}(W_\mu, Z) \leq \frac{(7/2+\sqrt{10})\beta}{b} + \frac{15+3\sqrt{10}}{5}\mathbb{P}(|W_\mu - W_\mu^L| > \beta). \tag{3.3}$$

On setting $\beta = \mu^{-1/2}\big\{\sup_{i\geq 1}\|F_{X_i}^{-1} - F_{X_i^L}^{-1}\| + CK\big\}$, and using Strassen's theorem we deduce (3.2) from (3.3) (recalling that $b = \frac{\sigma}{\sqrt{2}}$). The assertion after inequality (3.2) follows similarly. □

Proof of Theorem 1.1. To ease notation, in this proof we drop the subscripts from S_p and N_p. As noted by [38], the assumptions imposed on N and the X_i imply that $\mathcal{L}(M) = \mathcal{L}(N)$, meaning that we can take $M = N$. Inequality (1.3) now follows from inequality (3.2). To prove inequality (1.4), we note the following simple inequality (see [38])

$$\mathbb{E}|X_N - X_N^L| \leq \sup_{i\geq 1}\mathbb{E}|X_i| + \sup_{i\geq 1}\mathbb{E}|X_i^L| = \sup_{i\geq 1}\mathbb{E}|X_i| + \sup_{i\geq 1}\frac{\mathbb{E}[|X_i|^3]}{3\sigma^2} \leq \sigma + \frac{\rho_3}{3\sigma^2},$$

where in the final step the Cauchy-Schwarz inequality was applied. We are now able to obtain (1.4) from (3.1).

To prove inequality (1.5), we apply inequality (2.23) of Corollary 2.7. We use the assumption that $\sup_{i\geq 1}\mathbb{E}[X_i^{k+2}] < \infty$, the moment relation (2.10) and the simple inequality $|a+b|^r \leq 2^{r-1}(|a|^r + |b|^r)$, $r \geq 1$, to obtain the bound

$$\begin{aligned}\mathbb{E}[|S - S^L|^k] &= p^{k/2}\mathbb{E}[|X_N - X_N^L|^k] \\ &\leq 2^{k-1}p^{k/2}(\mathbb{E}[|X_N|^k] + \mathbb{E}[|X_N^L|^k]) \\ &\leq 2^{k-1}p^{k/2}\bigg(\rho_k + \frac{2\rho_{k+2}}{(k+1)(k+2)\sigma^2}\bigg). \end{aligned} \tag{3.4}$$

Substituting into (2.23) then yields inequality (1.5).

We end by establishing inequality (1.6). We now assume that $X_1, X_2, \ldots$ are identically distributed with $\mathbb{E}[X_1^3] = 0$ and $\mathbb{E}[X_1^4] < \infty$. We prove inequality (1.6) by applying inequality (2.20) of Theorem 2.5. We proceed similarly to we did in obtaining (3.4), but this time use the independence of X_N and X_N^L to obtain

$$\mathbb{E}[(S - S^L)^2] = p\mathbb{E}[(X_N - X_N^L)^2] = p(\mathbb{E}[X_N^2] + \mathbb{E}[(X_N^L)^2]) = p\bigg(\sigma^2 + \frac{\mathbb{E}[X_1^4]}{6\sigma^2}\bigg). \tag{3.5}$$

We now bound $\mathbb{E}[|\mathbb{E}[S - S^L \,|\, S]|]$. We have

$$\mathbb{E}[S - S^L \,|\, S] = \sqrt{p}\mathbb{E}[X_N - X_N^L \,|\, S] = \sqrt{p}\big(\mathbb{E}[X_N \,|\, S] - \mathbb{E}[X_N^L]\big),$$

as X_N^L and S are independent. Also, due to the assumption that $\mathbb{E}[X_i^3] = 0$ for all $i \geq 1$, we have, by (2.10), that $\mathbb{E}[X_N^L] = \frac{1}{3\sigma^2}\mathbb{E}[X_N^3] = 0$. By the tower property of conditional expectation we then have

$$\begin{aligned}\mathbb{E}[S - S^L \,|\, S] &= \sqrt{p}\mathbb{E}[\mathbb{E}[X_N \,|S, N] \,|\, S] \\ &= \mathbb{E}\bigg[\frac{S}{N}\,\bigg|\, S\bigg],\end{aligned}$$

where we used that because the X_i are i.i.d., and therefore exchangeable, $\mathbb{E}[X_N \,|\, S, N] = S/(\sqrt{p}N)$. Therefore

$$\begin{aligned}\mathbb{E}[|\mathbb{E}[S - S^L \,|\, S]|] &= \mathbb{E}\bigg[\bigg|\mathbb{E}\bigg[\frac{S}{N}\,\bigg|\, S\bigg]\bigg|\bigg] \\ &\leq \mathbb{E}\bigg[\mathbb{E}\bigg[\frac{|S|}{N}\,\bigg|\, S\bigg]\bigg]\end{aligned}$$

$$= \mathbb{E}\left[\frac{|S|}{N}\right] = \sqrt{p}\sum_{n=1}^{\infty}\frac{1}{n}\mathbb{E}\left|\sum_{i=1}^{n}X_i\right|\mathbb{P}(N=n). \tag{3.6}$$

Taking $h(x)=|x|$ in inequality (4.4) (note that $h\in\mathcal{H}_{\mathrm{W}}$) gives the inequality

$$\left|\frac{1}{\sqrt{n}}\mathbb{E}\left|\sum_{i=1}^{n}X_i\right| - \sqrt{\frac{2}{\pi}}\sigma\right| \le \frac{\sigma}{\sqrt{n}}\left(2+\frac{\mathbb{E}[|X_1|^3]}{\sigma^3}\right)$$

(see [4] for a similar bound), and on applying this inequality to (3.6) we obtain the bound

$$\mathbb{E}[|\mathbb{E}[S-S^L\,|\,S]|] \le \sqrt{\frac{2p}{\pi}}\sigma\mathbb{E}[N^{-1/2}] + \sqrt{p}\sigma\left(2+\frac{\mathbb{E}[|X_1|^3]}{\sigma^3}\right)\mathbb{E}[N^{-1}]. \tag{3.7}$$

The expectation $\mathbb{E}[N^{-1}]$ is easily evaluated:

$$\mathbb{E}[N^{-1}] = \sum_{n=1}^{\infty}\frac{p(1-p)^{n-1}}{n} = \frac{p\log(1/p)}{1-p}.$$

We can bound $\mathbb{E}[N^{-1/2}]$ through an application of the integral test:

$$\begin{aligned}\frac{1-p}{p}\mathbb{E}[N^{-1/2}] &= \sum_{n=1}^{\infty}\frac{(1-p)^n}{\sqrt{n}} < \int_0^{\infty}\frac{(1-p)^x}{\sqrt{x}}\,\mathrm{d}x = \int_0^{\infty}\frac{\exp(x\log(1-p))}{\sqrt{x}}\,\mathrm{d}x \\ &= \sqrt{\frac{2}{-\log(1-p)}}\int_0^{\infty}\mathrm{e}^{-t^2/2}\,\mathrm{d}t = \sqrt{\frac{\pi}{-\log(1-p)}} < \sqrt{\frac{\pi}{p}},\end{aligned}$$

where we used the standard inequality $\log(1+x)<x$, for $x>-1$, in the last step. Plugging the estimates for $\mathbb{E}[N^{-1/2}]$ and $\mathbb{E}[N^{-1}]$ into (3.7) then yields the bound

$$\mathbb{E}[|\mathbb{E}[S-S^L\,|\,S]|] < \frac{\sqrt{2}\sigma p}{1-p} + \frac{\sigma p^{3/2}\log(1/p)}{1-p}\left(2+\frac{\mathbb{E}[|X_1|^3]}{\sigma^3}\right). \tag{3.8}$$

Finally, inserting (3.5) and inequality (3.8) into (2.20) yields the desired bound. □

4. Proof of Theorem 1.3

Let $Z\sim\mathrm{Laplace}(0,\frac{\sigma}{\sqrt{2}})$ and recall that $T_n = B_{n-1}^{1/2}\sum_{i=1}^{n}X_i$, where the $X_1,\ldots,X_n$ are independent random variables with zero mean and variance $\sigma^2\in(0,\infty)$. Then we have the representations

$$\begin{aligned}T_n &=_d U_nV_n,\\ Z &=_d UV,\end{aligned}$$

where $U_n=\sqrt{nB_{n-1}}$, $V_n=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i$, U follows the Rayleigh distribution with density function $f_U(x)=2x\mathrm{e}^{-x^2}$, $x>0$, and $V\sim N(0,\sigma^2)$ are mutually independent random variables. This representation of the Laplace distribution is given in ([28], Prop. 2.2.1). In the limit $n\to\infty$, U_n converges in distribution to U, and, by the central limit theorem, V_n converges in distribution to V. Indeed, $\mathbb{P}(U_n\le u)=1-(1-u^2/n)^{n-1}$, $u\in(0,\sqrt{n})$, which converges to $1-\mathrm{e}^{-u^2}$ as $n\to\infty$. We prove Theorem 1.3 by obtaining explicit bounds on the distance between the distributions of U_n and U and the distributions of V_n and V with respect to suitable probability metrics and then combine these bounds to bound the distance between $\mathcal{L}(T_n)$ and the Laplace$(0, \frac{\sigma}{\sqrt{2}})$ distribution. We combine these bounds through the following lemma.

Lemma 4.1. *Let Y_1, Y_2, Z_1, Z_2 be real-valued random variables. Then*

$$\begin{aligned} d_{\mathrm{K}}(Y_1Z_1, Y_2Z_2) &\le d_{\mathrm{K}}(Y_1, Y_2) + d_{\mathrm{K}}(Z_1, Z_2), \\ d_{\mathrm{W}}(Y_1Z_1, Y_2Z_2) &\le \mathbb{E}|Z_1| d_{\mathrm{W}}(Y_1, Y_2) + \mathbb{E}|Y_2| d_{\mathrm{W}}(Z_1, Z_2), \\ d_{1,2}(Y_1Z_1, Y_2Z_2) &\le \mathbb{E}|Z_1| d_{\mathrm{W}}(Y_1, Y_2) + \mathbb{E}[Y_2^2] d_2(Z_1, Z_2), \end{aligned} \tag{4.1}$$

where each inequality holds provided the expectations in the the right-hand side of the inequality exist.

Proof. We prove the bound for $d_{1,2}$; the bounds for d_{K} and d_{W} are obtained through similar and slightly simpler arguments. Let $h \in \mathcal{H}_{1,2}$. Then, by the triangle inequality and conditioning,

$$\begin{aligned} &|\mathbb{E}[h(Y_1Z_1)] - \mathbb{E}[h(Y_2Z_2)]| \\ &\le |\mathbb{E}[h(Y_1Z_1)] - \mathbb{E}[h(Y_2Z_1)]| + |\mathbb{E}[h(Y_2Z_1)] - \mathbb{E}[h(Y_2Z_2)]| \\ &= |\mathbb{E}[\mathbb{E}[h(Y_1Z_1) - h(Y_2Z_1)] \,|\, Z_1]]| + |\mathbb{E}[\mathbb{E}[h(Y_2Z_1) - h(Y_2Z_2)] \,|\, Y_2]]| \\ &\le \mathbb{E}[|\mathbb{E}[h(Y_1Z_1) - h(Y_2Z_1)] \,|\, Z_1]|] + \mathbb{E}[|\mathbb{E}[h(Y_2Z_1) - h(Y_2Z_2)] \,|\, Y_2]|]. \end{aligned} \tag{4.2}$$

Now, for $a \in \mathbb{R} \setminus \{0\}$ and real-valued random variables X and Y we have that

$$\begin{aligned} |\mathbb{E}[h(aX)] - \mathbb{E}[h(aY)]| &\le d_{\mathrm{W}}(aX, aY) = a d_{\mathrm{W}}(X, Y), \\ |\mathbb{E}[h(aX)] - \mathbb{E}[h(aY)]| &\le d_2(aX, aY) = a^2 d_2(X, Y), \end{aligned}$$

since $\mathcal{H}_{1,2} \subset \mathcal{H}_{\mathrm{W}}$ and $\mathcal{H}_{1,2} \subset \mathcal{H}_2$. Applying these inequalities to (4.2) we obtain that, for $h \in \mathcal{H}_{1,2}$,

$$\begin{aligned} |\mathbb{E}[h(Y_1Z_1)] - \mathbb{E}[h(Y_2Z_2)]| &\le \mathbb{E}[|Z_1 d_{\mathrm{W}}(Y_1, Y_2)|] + \mathbb{E}[|Y_2^2 d_2(Z_1, Z_2)|] \\ &= \mathbb{E}|Z_1| d_{\mathrm{W}}(Y_1, Y_2) + \mathbb{E}[Y_2^2] d_2(Z_1, Z_2) \end{aligned} \tag{4.3}$$

The bound (4.3) holds for all $h \in \mathcal{H}_{1,2}$, and as $d_{1,2}(Y_1Z_1, Y_2Z_2) = \sup_{h\in\mathcal{H}_{1,2}} |\mathbb{E}[h(Y_1Z_1)] - \mathbb{E}[h(Y_2Z_2)]|$ it follows that inequality (4.1) holds. □

There is a vast literature on bounds for $d_{\mathcal{H}}(V_n, V)$. We will make use of three bounds from the literature for the cases $\mathcal{H}_{\mathrm{K}}$, $\mathcal{H}_{\mathrm{W}}$ and $\mathcal{H}_2$.

Theorem 4.2 (Shevtsova [43])**.** *Let $X_1, \ldots, X_n$ be independent random variables with $\mathbb{E}[X_i] = 0$, $\mathrm{Var}(X_i) = \sigma^2 \in (0, \infty)$ and $\mathbb{E}[|X_i|^3] < \infty$, for all $1 \le i \le n$. Denote $V_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ and let $V \sim N(0, \sigma^2)$. Then*

$$d_{\mathrm{K}}(V_n, V) \le \frac{C_0}{\sigma^3 n^{3/2}} \sum_{i=1}^n \mathbb{E}[|X_i|^3],$$

where $C_0 = 0.5600$.

Theorem 4.3 (Reinert [39])**.** *Under the same assumptions as Theorem 4.2, we have that, for $h \in \mathcal{H}_{\mathrm{W}}$,*

$$|\mathbb{E}[h(V_n)] - \mathbb{E}[h(V)]| \le \frac{\sigma}{n^{3/2}} \sum_{i=1}^n \left(2 + \frac{\mathbb{E}[|X_i|^3]}{\sigma^3}\right). \tag{4.4}$$

Consequently,

$$d_{\mathrm{W}}(V_n, V) \leq \frac{\sigma}{n^{3/2}} \sum_{i=1}^{n} \left(2 + \frac{\mathbb{E}[|X_i|^3]}{\sigma^3}\right). \tag{4.5}$$

Theorem 4.4 (Gaunt [16])**.** *Let $X_1, \ldots, X_n$ be independent random variables with $\mathbb{E}[X_i] = 0$, $\mathrm{Var}(X_i) = \sigma^2 \in (0, \infty)$, $\mathbb{E}[X_i^3] = 0$ and $\mathbb{E}[X_i^4] < \infty$, for all $1 \leq i \leq n$. Then*

$$d_2(V_n, V) \leq \frac{\sigma^2}{n^2} \sum_{i=1}^{n} \left(1 + \frac{\mathbb{E}[X_i^4]}{3\sigma^4}\right). \tag{4.6}$$

Remark 4.5. The Berry-Esseen Theorem 4.2, with a larger constant C_0, was proved independently by Berry [2] and Esseen [12] in the early 1940s, and since then several works have improved on the constant with the best estimate of $C_0 = 0.5600$ due to [43]. For i.i.d. random variables $X_1, \ldots, X_n$, the constant improves to $C_0 = 0.4748$ [44]. The assumption of bounded third absolute moments can also be reduced at the expense of a slightly more complicated bound with bigger constants [14]. Theorem 4.3 is formulated slightly differently in Theorem 2.1 of [39], but by re-scaling we obtain the bound (4.5). This is also the case for Theorem 4.4, and we additionally obtain an improved constant in (4.6) by using the bound $\|f^{(4)}\| \leq 2\|h''\|$ (due to [5]) for the solution of the standard normal Stein equation $f''(x) - xf'(x) = h(x) - \mathbb{E}[N]$, $N \sim N(0,1)$, rather than the bound $\|f^{(4)}\| \leq 3\|h''\|$ that was used in proof of Theorem 3.1 of [16].

As the Rayleigh distribution is a special case of the generalized gamma distribution, the following lemma follows as a special case of Proposition 2.3 of [17].

Lemma 4.6. *Let U denote a Rayleigh random variable with probability density function $p_U(x) = 2x\mathrm{e}^{-x^2}$, $x > 0$. Suppose that $f : (0, \infty) \to \mathbb{R}$ is differentiable and such that $\mathbb{E}|Uf'(U)| < \infty$, $\mathbb{E}|f(U)| < \infty$ and $\mathbb{E}|U^2 f(U)| < \infty$. Then*

$$\mathbb{E}[\mathcal{A}_U f(U)] = 0,$$

where $\mathcal{A}_U f(x) = xf'(x) + (2 - 2x^2)f(x)$.

Lemma 4.7. *Let $U_n = \sqrt{nB_{n-1}}$, where $B_{n-1} \sim \mathrm{Beta}(1, n-1)$. Suppose that $f : (0, \sqrt{n}) \to \mathbb{R}$ is differentiable and such that $\mathbb{E}|U_n f'(U_n)| < \infty$, $\mathbb{E}|U_n^3 f'(U_n)| < \infty$, $\mathbb{E}|f(U_n)| < \infty$ and $\mathbb{E}|U_n^2 f(U_n)| < \infty$. Then*

$$\mathbb{E}[\mathcal{A}_{U_n} f(U_n)] = 0. \tag{4.7}$$

where $\mathcal{A}_{U_n} f(x) = x(1 - x^2/n)f'(x) + (2 - 2x^2)f(x)$.

Proof. Define the operator T_r by $T_r y(x) = xy'(x) + ry(x)$, $r \in \mathbb{R}$. In this notation, the classical Stein operator for the Beta$(1, n-1)$ distribution is given by $A_{B_{n-1}} y(x) = T_1 y(x) - xT_n y(x)$ [6, 23]. Let $C_n = B_{n-1}^{1/2}$ and let $g : (0,1) \to \mathbb{R}$ by such that $\mathbb{E}|C_n g'(C_n)| < \infty$, $\mathbb{E}|C_n^3 g'(C_n)| < \infty$, $\mathbb{E}|g(C_n)| < \infty$ and $\mathbb{E}|C_n^2 g(C_n)| < \infty$. Then, by equation (15) of [20],

$$\mathbb{E}[T_2 g(C_n) - C_n^2 T_{2n} g(C_n)] = 0. \tag{4.8}$$

(The conditions on g that are stated above are not specified in [20], but on examining their analysis one can see that these conditions ensure that (4.8) holds.) That is

$$\mathbb{E}\big[C_n(1 - C_n^2)g'(C_n) + (2n - 2nC_n^2)g(C_n)\big] = 0. \tag{4.9}$$

We have that $U_n =_d \sqrt{n}C_n$, and on rescaling we deduce (4.7) from (4.9). □

In the following lemma, the bound (4.10) is proved purely for reasons of exposition, as an improved bound will be stated in Remark 4.9. Proving both the Kolmogorov and Wasserstein distance bounds requires very little more work than only proving the Wasserstein distance bound.

Lemma 4.8. *Let the random variables U_n and U be defined as above. Then, for $n \geq 2$,*

$$d_{\mathrm{K}}(U_n, U) \leq \frac{2}{n}, \tag{4.10}$$

$$d_{\mathrm{W}}(U_n, U) \leq \frac{11.49}{n}. \tag{4.11}$$

Proof. Let the Stein operators A_U and A_{U_n} be defined as in Lemmas 4.6 and 4.7, respectively. Suppose that $h : (0, \infty) \to \mathbb{R}$ is either bounded or Lipschitz. Let f be the solution of the Rayleigh$(1/\sqrt{2})$ Stein equation $A_U f(x) = h(x) - \mathbb{E}[h(U)]$, which by Lemma 5.4, we know satisfies the bounds

$$\|xf'(x)\| \leq \frac{2}{2^{-1}} \times \frac{1}{2}\|h - \mathbb{E}[h(U)]\| \leq 2, \quad h \in \mathcal{H}_{\mathrm{K}}, \tag{4.12}$$

$$\|f'\| \leq \frac{6.11}{2^{-3/2}} \times \frac{1}{2}\|h'\| \leq 8.6408, \quad h \in \mathcal{H}_{\mathrm{W}}. \tag{4.13}$$

Then

$$\begin{aligned} |\mathbb{E}[h(U_n)] - \mathbb{E}[h(U)]| &= |\mathbb{E}[\mathcal{A}_U f(U_n)]| = |\mathbb{E}[\mathcal{A}_U f(U_n) - \mathcal{A}_{U_n} f(U_n)]| \\ &= \frac{1}{n}|\mathbb{E}[U_n^3 f'(U_n)]| \\ &\leq \frac{1}{n} \min\Big\{\|xf'(x)\|\mathbb{E}[U_n^2], \|f'\|\mathbb{E}[U_n^3]\Big\}. \end{aligned} \tag{4.14}$$

That $\mathbb{E}[\mathcal{A}_{U_n} f(U_n)] = 0$ follows from the assumptions on h and the estimates of Lemma 5.4 for the solution of the Rayleigh Stein equation. Now, $\mathbb{E}[U_n^2] = 1$ and

$$\begin{aligned} \mathbb{E}[U_n^3] &= n^{3/2}\mathbb{E}[B_{n-1}^{3/2}] = n^{3/2}\int_0^1 (n-1)x^{3/2}(1-x)^{n-2}\,\mathrm{d}t = n^{3/2}(n-1)B\big(\tfrac{5}{2}, n-1\big) \\ &= n^{3/2}(n-1)\frac{\Gamma(5/2)\Gamma(n-1)}{\Gamma(n+3/2)} = \frac{3\sqrt{\pi}n^{3/2}\Gamma(n)}{4\Gamma(n+3/2)}, \end{aligned} \tag{4.15}$$

where $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,\mathrm{d}x$ is the beta function, and we used the standard formulas $u\Gamma(u) = \Gamma(u+1)$ and $\Gamma(5/2) = 3\sqrt{\pi}/4$. Now $n^{3/2}\Gamma(n)/\Gamma(n+3/2)$ is an increasing function of n on $(0,\infty)$ [24]. Therefore, for $n \geq 2$,

$$\mathbb{E}[U_n^3] \leq \frac{3\sqrt{\pi}}{4} \lim_{n\to\infty} \frac{n^{3/2}\Gamma(n)}{\Gamma(n+3/2)} = \frac{3\sqrt{\pi}}{4},$$

where the limit follows from [32, formula (5.6.4)]. Applying the bounds (4.12) and (4.13) together with the bounds for $\mathbb{E}[U_n^2]$ and $\mathbb{E}[U_n^3]$ to (4.14) then yields the bounds (4.10) and (4.11). □

Remark 4.9. The following bounds will appear in the supplementary material of the arXiv version of the preprint [11]. For $n \geq 2$,

$$d_{\mathrm{K}}(U_n, U) \leq \frac{1}{n}\bigg(1 + 2\bigg(1 - \frac{2}{n}\bigg)^{n-2}\bigg), \tag{4.16}$$

and

$$\begin{aligned} d_{\mathrm{W}}(U_n, U) \leq & -\frac{\sqrt{\pi}\Gamma(n)}{4\sqrt{n}\Gamma(n+1/2)} \\ & + 2\sqrt{2}\frac{n-1}{n^n} \cdot \frac{(n-2)^n n(40 + 11(n-4)n) + (n-2)^3 n^n {}_2F_1(-\frac{1}{2}, 3-n; \frac{1}{2}; \frac{2}{n})}{(n-2)^2(2n-5)(2n-3)(2n-1)}, \end{aligned} \tag{4.17}$$

where ${}_2F_1(a, b; c; x)$ is the Gaussian hypergeometric function. (We define $0^0 := 1$, but this is irrelevant because the bound (4.16) is greater than 1 in this case.) These bounds were obtained using a recent technique of [11] for bounding distances between distributions that builds upon the formalism of [10] for new representations of solutions to Stein equations. For another recent approach to bounding distances between distributions, see [9].

Our Kolmogorov distance bound (4.10) outperforms (4.16) when $n = 2$ (although in this case the upper bound of 1 is trivial), but for all $n \geq 3$ the reverse is true. Numerical calculations carried using *Mathematica* suggest that the Wasserstein bound (4.17) improves on our bound (4.11) for all $n \geq 2$, although verifying this assertion analytically seems to be difficult. Our bound is of course much simpler and the dependence on n is very clear. For this reason, we will use the bound (4.11) in our proof of Theorem 1.3.

Proof of Theorem 1.3. Recall that $T_n =_d U_n V_n$ and $Z =_d UV$. Then, by Lemma 4.1,

$$d_{\mathrm{K}}(T_n, Z) \leq d_{\mathrm{K}}(U_n, U) + d_{\mathrm{K}}(V_n, V), \tag{4.18}$$

$$d_{\mathrm{W}}(T_n, Z) \leq \mathbb{E}|V| d_{\mathrm{W}}(U_n, U) + \mathbb{E}[U_n] d_{\mathrm{W}}(V_n, V), \tag{4.19}$$

$$d_{1,2}(T_n, Z) \leq \mathbb{E}|V| d_{\mathrm{W}}(U_n, U) + \mathbb{E}[U_n^2] d_2(V_n, V). \tag{4.20}$$

By standard formulas for the moments and absolute moments of the beta and normal distributions, we have that $\mathbb{E}[U_n^2] = 1$ and $\mathbb{E}|V| = \sigma\sqrt{2/\pi}$. Also, by a similar calculation to the one used to obtain the formula (4.15) we have, for $n \geq 2$,

$$\mathbb{E}[U_n] = \frac{\sqrt{\pi}\sqrt{n}\Gamma(n)}{2\Gamma(n+1/2)} \leq \frac{\sqrt{\pi}\sqrt{2}\Gamma(2)}{2\Gamma(5/2)} = \frac{2\sqrt{2}}{3},$$

where we used that $\sqrt{n}\Gamma(n)/\Gamma(n+1/2)$ is a decreasing function of n on $(0, \infty)$ [24]. Theorems 4.2–4.4 give bounds for $d_{\mathrm{K}}(V_n, V)$, $d_{\mathrm{W}}(V_n, V)$ and $d_2(V_n, V)$, respectively, and $d_{\mathrm{K}}(U_n, U)$ is bounded by inequality (4.16) and $d_{\mathrm{W}}(U_n, U)$ is bounded by inequality (4.11). Substituting all of these estimates into (4.18), (4.19) and (4.20) then yields the bounds as stated in Theorem 1.3. □

5. The Rayleigh Stein equation

Let $R \sim \text{Rayleigh}(\sigma)$, $\sigma > 0$, follow the Rayleigh distribution with density function

$$\rho_R(x) = \frac{x}{\sigma^2}\mathrm{e}^{-x^2/(2\sigma^2)}, \quad x > 0.$$

The Rayleigh distribution is a special case of the chi distribution (up to scaling). A random variable K following the chi distribution with $k>0$ degrees of freedom, denoted by $\chi_{(k)}$, has probability density function

$$\rho_k(x)=\frac{1}{2^{k/2-1}\Gamma(k/2)}x^{k-1}\mathrm{e}^{-x^2/2},\quad x>0.$$

We proceed by obtaining bounds for the solution of the chi distribution Stein equation, before specialising to the solution of the Rayleigh Stein equation.

We first note that the density ρ_k satisfies the differential equation

$$\big(s(x)\rho(x)\big)'=\tau(x)\rho(x), \tag{5.1}$$

where $s(x)=x$ and $\tau(x)=k-x^2$. It therefore follows from Theorem 1 of [42] that a Stein equation for the $\chi_{(k)}$ distribution is given by

$$xf'(x)+(k-x^2)f(x)=h(x)-\mathbb{E}[h(K)], \tag{5.2}$$

where $K\sim\chi_{(k)}$. It is straightforward to solve (5.2) (see Prop. 1 of [42]):

$$f(x)=\frac{1}{x\rho_k(x)}\int_0^x(h(t)-\mathbb{E}[h(K)])\rho_k(t)\,\mathrm{d}t, \tag{5.3}$$

$$=-\frac{1}{x\rho_k(x)}\int_x^\infty(h(t)-\mathbb{E}[h(K)])\rho_k(t)\,\mathrm{d}t. \tag{5.4}$$

In order to bound the solution (5.3) and its first derivative, it will be useful to note the following straightforward extension of Lemmas 1 and 3 of [41].

Lemma 5.1. *Let ρ be the probability density function of a random variable Y, supported on (a,b), which satisfies the differential equation (5.1), where $s(x)$ is a polynomial of degree no greater than two and $\tau(x)$ is monotonic in (a,b) with exactly one sign change at the point $m\in(a,b)$. Let $h:(a,b)\to\mathbb{R}$ be bounded. Then, the solution of the Stein equation $s(x)f'(x)+\tau(x)f(x)=h(x)-\mathbb{E}h(Y)$, as given by $f(x)=\frac{1}{s(x)\rho(x)}\int_a^x(h(t)-\mathbb{E}[h(Y)])\rho(t)\,\mathrm{d}t$, satisfies the bounds*

$$\|f\|\leq M\|h-\mathbb{E}[h(Y)]\|, \tag{5.5}$$

$$\|s(x)f'(x)\|\leq 2\|h-\mathbb{E}[h(Y)]\|, \tag{5.6}$$

where

$$M=\frac{1}{s(m)\rho(m)}\max\{F(m),1-F(m)\},$$

with F denoting the distribution function of Y.

Remark 5.2. The bound (5.5) is a generalisation of the corresponding bound of Lemma 1 of [41], which is only given for the case that $\tau(x)=a(\mathbb{E}[Y]-x)$, where $a\neq 0$. The crucial feature of this function that is exploited in the proof of [41] is that $\tau(x)$ is monotonic with exactly one sign change at $x=\mathbb{E}[Y]$. As noted by [27], we can therefore extend the result of [41] to any $\tau(x)$ that is monotonic with only one change of sign.

Lemma 5.3. *Let $f:(0,\infty)\to\mathbb{R}$ denote the solution (5.3) of the Stein equation (5.2). Let $h:(0,\infty)\to\mathbb{R}$ be bounded. Then*

$$\|f\|\leq\frac{\Gamma(k/2)\mathrm{e}^{k/2}}{2(k/2)^{k/2}}\|h-\mathbb{E}[h(K)]\|, \tag{5.7}$$

$$\|xf'(x)\|\leq 2\|h-\mathbb{E}[h(K)]\|. \tag{5.8}$$

Proof. Bounds (5.7) and (5.8) follow easily from Lemma 5.1; note that $\tau(x)=k-x^2$ satisfies the assumption of the lemma. To apply the lemma, we note that here $m=\sqrt{k}$, being the positive solution to the equation $k-x^2=0$; $s(x)=x$; and we used the trivial bound $\max\{F(m),1-F(m)\}\leq 1$. □

We now specialise to the case $k=2$, which corresponds to the Rayleigh distribution.

Lemma 5.4. *Let f denote the solution of the Rayleigh Stein equation $\sigma^2xf'(x)+(2\sigma^2-x)f(x)=h(x)-\mathbb{E}[h(R)]$, where $R\sim\mathrm{Rayleigh}(\sigma)$. Let $h:(0,\infty)\to\mathbb{R}$ be bounded. Then*

$$\|f\|\leq\frac{\mathrm{e}}{2\sigma^2}\|h-\mathbb{E}[h(R)]\|, \tag{5.9}$$

$$\|xf'(x)\|\leq\frac{2}{\sigma^2}\|h-\mathbb{E}[h(R)]\|. \tag{5.10}$$

Now suppose that h is Lipschitz. Then

$$\|xf(x)\|\leq\frac{2.325}{\sigma}\|h'\|, \tag{5.11}$$

$$\|f'\|\leq\frac{6.11}{\sigma^3}\|h'\|, \tag{5.12}$$

$$\|xf''(x)\|\leq\frac{11.30}{\sigma^3}\|h'\|. \tag{5.13}$$

Proof. For ease of notation, we consider the case $\sigma=1$. The general case follows from rescaling. Bounds (5.9) and (5.10) follow immediately from Lemma 5.3.

Now we prove inequality (5.11). Let h be Lipschitz. By the mean value theorem, for $t>0$, $|h(t)-\mathbb{E}[h(R)]|\leq\|h'\|(t+\mathbb{E}[R])=\|h'\|(t+\sqrt{\pi/2})$. Therefore, for $x>0$,

$$|xf(x)|\leq\frac{\|h'\|}{\rho_R(x)}\int_0^x\big(\sqrt{\tfrac{\pi}{2}}+t\big)\rho_R(t)\,\mathrm{d}t=:\frac{\|h'\|}{\rho_R(x)}I_1(x),$$

and

$$|xf(x)|\leq\frac{\|h'\|}{\rho_R(x)}\int_x^\infty\big(\sqrt{\tfrac{\pi}{2}}+t\big)\rho_R(t)\,\mathrm{d}t=:\frac{\|h'\|}{\rho_R(x)}I_2(x).$$

By integration by parts, the integrals $I_1(x)$ and $I_2(x)$ can be evaluated in terms of the error function $\mathrm{erf}(x)=\frac{2}{\sqrt{\pi}}\int_0^x\mathrm{e}^{-t^2}\,\mathrm{d}t$:

$$I_1(x)=\sqrt{\frac{\pi}{2}}(1-\mathrm{e}^{-x^2/2})+\sqrt{\frac{\pi}{2}}\mathrm{erf}\Big(\frac{x}{\sqrt{2}}\Big)-x\mathrm{e}^{-x^2/2},$$

$$I_2(x)=\sqrt{\frac{\pi}{2}}(1+\mathrm{e}^{-x^2/2})-\sqrt{\frac{\pi}{2}}\mathrm{erf}\Big(\frac{x}{\sqrt{2}}\Big)+x\mathrm{e}^{-x^2/2}.$$

It can be seen that $I_1(x)/\rho_R(x)$ and $I_2(x)/\rho_R(x)$ are increasing and decreasing functions of x, respectively, and we used *Mathematica* to compute that the two functions intersect at the point $x^* = 1.360722\ldots$. Therefore, for all $x > 0$,

$$|xf(x)| \leq \frac{I_1(x^*)}{p(x^*)}\|h'\| = 2.325\|h'\|.$$

Lastly, we establish the bounds (5.12) and (5.13). Differentiating both sides of (5.2) and rearranging gives

$$xf''(x) + (3 - x^2)f'(x) = h'(x) + 2xf(x), \tag{5.14}$$

which we recognise as the $\chi_{(3)}$ Stein equation with test function $h'(x) + 2xf(x)$, applied to the function f'. It is important to note that the test function $h'(x) + 2xf(x)$ has mean zero with respect to the random variable $K_3 \sim \chi_{(3)}$. This follows because $xf''(x) + (3 - x^2)f'(x)$ is a Stein operator for the $\chi_{(3)}$ distribution, meaning that $\mathbb{E}[K_3 f''(Y) + (3 - K_3^2)f'(K_3)] = 0$, and therefore from (5.14) we have that $\mathbb{E}[h'(K_3) + 2K_3 f(K_3)] = 0$. We can therefore use the iterative technique of [8] to deduce bounds for $\|f'\|$ and $\|xf''(x)\|$ from our bounds (5.7) and (5.8) with $k = 3$ and (5.11). We have

$$\begin{aligned}\|f'\| &\leq \frac{2\Gamma(3/2)\mathrm{e}^{3/2}}{(3/2)^{3/2}}\|h'(x) + 2xf(x)\| \leq \frac{\Gamma(3/2)\mathrm{e}^{3/2}}{2(3/2)^{3/2}}\big(\|h'\| + 2\|xf(x)\|\big) \\ &\leq \frac{\Gamma(3/2)\mathrm{e}^{3/2}}{2(3/2)^{3/2}}(1 + 2\cdot 2.325)\|h'\| = 6.11\|h'\|,\end{aligned}$$

and

$$\|xf''(x)\| \leq 2\|h'(x) + 2xf(x)\| \leq 2(1 + 2\cdot 2.325)\|h'\| = 11.30\|h'\|,$$

which completes the proof. □

Acknowledgements. The author is supported by a Dame Kathleen Ollerenshaw Research Fellowship. The author would like to thank Yvik Swan for generously sharing some results that will be added to the supplementary material of the arXiv version of his preprint [11], which are stated in Remark 4.9. The author is grateful to the referees for their careful reading of the manuscript and for identifying several typos and errors in the displayed equations.

References

[1] B. Arras and C. Houdré, On Stein's method for infinitely divisible laws with finite first moment. Springer Briefs in Probability and Mathematical Statistics. Springer, Cham (2019).
[2] A.C. Berry, The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Am. Math. Soc.* **49** (1941) 122–136.
[3] A. Braverman and J.G. Dai, High order steady-state diffusion approximation of the Erlang-C system. Preprint arXiv: 1602.02866 (2016).
[4] L.H.Y. Chen and Q.-M. Shao, Stein's method for normal approximation, in Vol. 4 of *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* Singapore Univ. Press, Singapore (2005) 1–59.
[5] F. Daly, Upper bounds for Stein-type operators. *Electon. J. Probab.* **13** (2008) 566–587.
[6] C. Döbler, Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electr. J. Probab.* **20** (2015) 1–34.
[7] C. Döbler, Distributional transformations without orthogonality relations. *J. Theor. Probab.* **30** (2017) 85–116.
[8] C. Döbler, R.E. Gaunt and S.J. Vollmer, An iterative technique for bounding derivatives of solutions of Stein equations. *Electr. J. Probab.* **22** (2017) 1–39.
[9] L. Döbler, R. Samworth and J. Wellner, Bounding distributional errors via density ratios. Preprint arXiv: 1905.03009 (2019).
[10] M. Ernst, G. Reinert and Y. Swan, First order covariance inequalities via Stein's method. To appear in *Bernoulli* (2020).

[11] M. Ernst and Y. Swan, Distances between distributions via Stein's method. Preprint arXiv: 1909.11518 (2019).

[12] C.G. Esseen, On the Liapunoff limit of error in the theory of probability. *Ark. Mat. Astron. Fys.* **A28** (1942) 1–19.

[13] M. Fathi, Higher-order Stein kernels for Gaussian approximation. To appear in *Stud. Math.* (2020).

[14] W. Feller, On the Berry-Esseen Theorem. *Z. Wahrscheinlichkeit* **10** (1968) 261–268.

[15] R.E. Gaunt, Variance-Gamma approximation via Stein's method. *Electr. J. Probab.* **19** (2014) 1–33.

[16] R.E. Gaunt, Rates of convergence in normal approximation under moment conditions via new bounds on solutions of the Stein equation. *J. Theor. Probab.* **29** (2016) 231–247.

[17] R.E. Gaunt, Products of normal, beta and gamma random variables: Stein operators and distributional theory. *Braz. J. Probab. Stat.* **32** (2018) 437–466.

[18] R.E. Gaunt Wasserstein and Kolmogorov error bounds for variance-gamma approximation via Stein's method I. *J. Theor. Probab.* **33** (2020) 465–505.

[19] R.E. Gaunt Stein's method for functions of multivariate normal random variables. *Ann. Inst. Henri Poincaré Prob. Stat.* **56** (2020) 1484–1513.

[20] R.E. Gaunt, G. Mijoule, and Y. Swan, An algebra of Stein operators. *J. Math. Anal. Appl.* **469** (2019) 260–279.

[21] R.E. Gaunt, A.M. Pickett and G. Reinert, Chi-square approximation by Stein's method with application to Pearson's statistic. *Ann. Appl. Probab.* **27** (2017) 720–756.

[22] L. Goldstein and G. Reinert, Stein's Method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** (1997) 935–952.

[23] L. Goldstein and G. Reinert, Stein's method for the Beta distribution and the Pólya-Eggenberger Urn. *J. Appl. Probab.* **50** (2013) 1187–1205.

[24] M.E.H. Ismail, L. Lorch and M.E. Muldoon, Completely monotonic functions associated with the gamma function and its q-analogues. *J. Math. Anal. Appl.* **116** (1986) 1–9.

[25] A.V. Kakosyan, L.B. Klebanov and I.A. Melamed, Characterization of Distributions by the Method of Intensively Monotone Operators, Vol. 1088 of *Lecture Notes in Math.* Springer, Berlin (1984).

[26] V. Kalashnikov, Geometric Sums: Bounds for Rare Events with Applications. Risk Analysis, Reliability, Queueing. Kluwer Academic Publishers Group, Dordrecht (1997).

[27] E. Konzou and A. Koudou, About the Stein equation for the generalized inverse Gaussian and Kummer distributions. *ESAIM: PS* **24** (2020) 607–626.

[28] S. Kotz, T.J. Kozubowski and K. Podgórski, The Laplace Distribution and Generalizations: A Revisit with New Applications. Springer (2001).

[29] C. Lefèvre and S. Utev, Exact norms of a Stein-type operator and associated stochastic orderings. *Probab. Theory Rel.* **127** (2003) 353–366.

[30] C. Ley, G. Reinert and Y. Swan, Stein's method for comparison of univariate distributions. *Probab. Surv.* **14** (2017) 1–52.

[31] H. Luk, *Stein's Method for the Gamma Distribution and Related Statistical Applications.* Ph.D. thesis, University of Southern California (1994).

[32] F.W.J. Olver, D.W. Lozier, R.F. Boisvert and C.W. Clark, NIST Handbook of Mathematical Functions. Cambridge University Press (2010).

[33] A.G. Pakes, A characterization of gamma mixtures of stable laws motivated by limit theorems. *Stat. Neerl.* **46** (1992) 209–218.

[34] A.G. Pakes, On characterizations through mixed sums. *Aust. J. Stat.* **34** (1992) 323–339.

[35] E. Peköz and A. Röllin, New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** (2011) 587–608.

[36] E. Peköz, A. Röllin and N. Ross, Total variation error bounds for geometric approximation. *Bernoulli* **19** (2013) 610–632.

[37] E. Peköz, A. Röllin and N. Ross, Exponential and Laplace approximation for occupation statistics of branching random walk. *Electr. J. Probab.* **25** (2020) 1–22.

[38] J. Pike and H. Ren, Stein's method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** (2014) 571–587.

[39] G. Reinert, Couplings for normal approximations with Stein's method, in Microsurveys in Discrete Probability, volume of *DIMACS series AMS* (1998) 193–207.

[40] A. Rényi, A characterization of Poisson processes. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **1** (1957) 519–527.

[41] W. Schoutens, Orthogonal Polynomials in Steins Method. EURANDOM Report 99-041, EURANDOM, 1999.

[42] W. Schoutens, Orthogonal polynomials in Stein's method. *J. Math. Anal. Appl.* **253** (2001) 515–531.

[43] I. Shevtsova, An improvement of convergence rate estimates in the Lyapunov theorem. *Dokl. Math.* **253** (2010) 862–864.

[44] I. Shevtsova, On the absolute constants in the Berry Esseen type inequalities for identically distributed summands. Preprint arXiv: 1111.6554 (2011).

[45] C. Stein, A bound for the error in the normal approximation to the the distribution of a sum of dependent random variables. Vol. 2 of *Proc. Sixth Berkeley Symp. Math. Statis. Prob.* Univ. California Press, Berkeley (1972) 583–602.

[46] C. Stein, Approximate Computation of Expectations. IMS, Hayward, California (1986).

[47] A.A. Toda, Weak limit of the geometric sum of independent but not identically distributed random variables. Preprint arXiv: 1111.1786 (2011).