CrossMark

Wasserstein and Kolmogorov Error Bounds for Variance-Gamma Approximation via Stein's Method I

Robert E. Gaunt[1]

Abstract

The variance-gamma (VG) distributions form a four-parameter family that includes as special and limiting cases the normal, gamma and Laplace distributions. Some of the numerous applications include financial modelling and approximation on Wiener space. Recently, Stein's method has been extended to the VG distribution. However, technical difficulties have meant that bounds for distributional approximations have only been given for smooth test functions (typically requiring at least two derivatives for the test function). In this paper, which deals with symmetric variance-gamma (SVG) distributions, and a companion paper (Gaunt 2018), which deals with the whole family of VG distributions, we address this issue. In this paper, we obtain new bounds for the derivatives of the solution of the SVG Stein equation, which allow for approximations to be made in the Kolmogorov and Wasserstein metrics, and also introduce a distributional transformation that is natural in the context of SVG approximation. We apply this theory to obtain Wasserstein or Kolmogorov error bounds for SVG approximation in four settings: comparison of VG and SVG distributions, SVG approximation of functionals of isonormal Gaussian processes, SVG approximation of a statistic for binary sequence comparison, and Laplace approximation of a random sum of independent mean zero random variables.

Keywords Stein's method · Variance-gamma approximation · Distributional transformation · Rate of convergence

Mathematics Subject Classification (2010) Primary 60F05 · 62E17

✉ Robert E. Gaunt
robert.gaunt@manchester.ac.uk

[1] School of Mathematics, The University of Manchester, Manchester M13 9PL, UK

Springer

1 Introduction

1.1 Overview of Stein's Method for Variance-Gamma Approximation

The variance-gamma (VG) distribution with parameters $r > 0, \theta \in \mathbb{R}, \sigma > 0, \mu \in \mathbb{R}$ has probability density function

$$p(x) = \frac{1}{\sigma\sqrt{\pi}\Gamma(\frac{r}{2})} e^{\frac{\theta}{\sigma^2}(x-\mu)} \left(\frac{|x-\mu|}{2\sqrt{\theta^2+\sigma^2}}\right)^{\frac{r-1}{2}} K_{\frac{r-1}{2}}\left(\frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x-\mu|\right), \quad (1.1)$$

where $x \in \mathbb{R}$ and the modified Bessel function of the second kind $K_\nu(x)$ is defined in "Appendix A". If the random variable Z has density (1.1), we write $Z \sim \mathrm{VG}(r, \theta, \sigma, \mu)$. The support of the VG distributions is $\mathbb{R}$ when $\sigma > 0$, but in the limit $\sigma \to 0$ the support is the region (μ, ∞) if $\theta > 0$, and is $(-\infty, \mu)$ if $\theta < 0$. Alternative parametrisations are given in [10] and [29] (in which they use the name-generalised Laplace distribution). Distributional properties are given in [16] and Chapter 4 of the book [29].

The VG distribution was introduced to the financial literature by [32]. Due to their semi-heavy tails, VG distributions are useful for modelling financial data [33]; see the book [29] and references therein for an overview of the many applications. The class of VG distributions contains many classical distributions as special or limiting cases, such as the normal, gamma, Laplace, product of zero mean normals and difference of gammas (see Proposition 1.2 of [16] for a list of further cases). Consequently, the VG distribution appears in many other settings beyond financial mathematics [29]; for example, in alignment-free sequence comparison [31,45]. In particular, starting with the works [15,16], Stein's method [50] has been developed for VG approximation. The theory of [15,16] and the Malliavin-Stein method (see [36]) was applied by [12] to obtain "six moment" theorems for the VG approximation of double Wiener-Itô integrals. Further VG approximations are given in [1] and [2], in which the limiting distribution is the difference of two centred gamma random variables.

Introduced in 1972, Stein's method [50] is a powerful tool for deriving distributional approximations with respect to a probability metric. The theory for normal and Poisson approximation is particularly well established with numerous application in probability and beyond; see the books [6] and [3]. There is active research into the development of Stein's method for other distributional limits (see [30] for an overview), and Stein's method for exponential and geometric approximation, for example, is now also well developed; see the survey [48]. In particular, [39] have developed a framework to obtain error bounds for the Kolmogorov and Wasserstein distance metrics for exponential approximation, and [40] developed a framework for total variation error bounds for geometric approximation.

This paper and its companion [23] focus on the development of Stein's method for VG approximation. At the heart of the method [16] is the Stein equation

$$\sigma^2(x-\mu)f''(x) + (\sigma^2 r + 2\theta(x-\mu))f'(x) + (r\theta - (x-\mu))f(x) = \tilde{h}(x),$$
$$(1.2)$$

where $\tilde{h}(x) = h(x) - \mathbb{E}h(Z)$ for $h : \mathbb{R} \to \mathbb{R}$ and $Z \sim \text{VG}(r, \theta, \sigma, \mu)$. Together with the Stein equations of [41] and [43], this was one of the first second-order Stein equations to appear in the literature. We now set $\mu = 0$; the general case follows from the translation property that if $Z \sim \text{VG}(r, \theta, \sigma, \mu)$ then $Z - \mu \sim \text{VG}(r, \theta, \sigma, 0)$. The solution to (1.2) is then

$$
f_h(x) = -\frac{e^{-\theta x/\sigma^2}}{\sigma^2 |x|^\nu} K_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|x|\right) \int_0^x e^{\theta t/\sigma^2} |t|^\nu I_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|t|\right) \tilde{h}(t) \, dt
$$
$$
- \frac{e^{-\theta x/\sigma^2}}{\sigma^2 |x|^\nu} I_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|x|\right) \int_x^\infty e^{\theta t/\sigma^2} |t|^\nu K_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|t|\right) \tilde{h}(t) \, dt,
$$
(1.3)

where $\nu = \frac{r-1}{2}$ and the modified Bessel function of the first kind $I_\nu(x)$ is defined in "Appendix A". If h is bounded, then $f_h(x)$ and $f_h'(x)$ are bounded for all $x \in \mathbb{R}$. Moreover, this is the unique bounded solution when $r \geq 1$.

To approximate a random variable of interest W by a VG random variable Z, one may evaluate both sides of (1.2) at W, take expectations and finally take the supremum of both sides over a class of functions $\mathcal{H}$ to obtain

$$
\sup_{h \in \mathcal{H}} |\mathbb{E}h(W) - \mathbb{E}h(Z)| = \sup_{h \in \mathcal{H}} \left| \mathbb{E}\left[\sigma^2 W f_h''(W) + (\sigma^2 r + 2\theta W) f_h'(W) \right.\right.
$$
$$
\left.\left. - (r\theta - W) f_h(W)\right]\right|.
$$
(1.4)

Many important probability metrics are of the form $\sup_{h \in \mathcal{H}} |\mathbb{E}h(W) - \mathbb{E}h(Z)|$. In particular, taking

$$
\mathcal{H}_K = \{\mathbf{1}(\cdot \leq z) \,|\, z \in \mathbb{R}\},
$$
$$
\mathcal{H}_W = \{h : \mathbb{R} \to \mathbb{R} \,|\, h \text{ is Lipschitz}, \|h'\| \leq 1\},
$$
$$
\mathcal{H}_{BW} = \{h : \mathbb{R} \to \mathbb{R} \,|\, h \text{ is Lipschitz}, \|h\| \leq 1 \text{ and } \|h'\| \leq 1\}
$$

gives the Kolmogorov, Wasserstein and bounded Wasserstein distances, which we denote by d_K, d_W and d_{BW}, respectively.

The problem of bounding $\sup_{h \in \mathcal{H}} |\mathbb{E}h(W) - \mathbb{E}h(Z)|$ is thus reduced to bounding the solution (1.3) and some of its lower order derivatives and bounding the expectation on the right-hand side of (1.4). To date, the only techniques for bounding this expectation for VG approximation are local couplings [15,16] and the integration by parts technique used to prove Theorem 4.1 of [12]. Other coupling techniques that are commonly found in the Stein's method literature, such as exchangeable pairs [51] and Stein couplings [7], have yet to be used in VG approximation, although one of the contributions of this paper is a new coupling technique for SVG approximation by Stein's method.

The presence of modified Bessel functions in the solution (1.3) together with the singularity at the origin in the Stein equation (1.2) makes bounding the solution and its derivatives technically challenging. Indeed, in spite of the introduction of new

inequalities for modified Bessel functions and their integrals [17,18] and extensive calculations ([15], Sect. 3.3 and Appendix D), the first bounds given in the literature [16] were only given for the case $\theta = 0$ and had a far from optimal dependence on the parameter r. Substantial progress was made by [9], in which their iterative approach reduced the problem of bounding the derivatives of any order to bounding just the solution and its first derivative. However, the bounds obtained in [9] have a dependence on the test function h which means that error bounds for VG approximation can only be given for smooth test functions.

1.2 Summary of Results and Outline of the Paper

In this paper and its companion [23], we obtain new bounds for the solution of the VG Stein equation that allow for Wasserstein and Kolmogorov error bounds for VG approximation via Stein's method. This paper focuses on the case $\theta = 0$ (symmetric variance-gamma (SVG) distributions), whilst [23] deals with the whole family of VG distributions. This organisation is due to the additional complexity of the $\theta \neq 0$ case. One of the difficulties is that when $\theta \neq 0$, the inequalities for expressions involving integrals of modified Bessel functions that we use to bound the solution take a more complicated form, meaning our main results need to be presented in parallel for the two cases. It should be noted, though, that, once the inequalities for modified Bessel functions have been established (which has now been done in [17,18,21]), the intrinsic difficulty of bounding the derivatives of the solution of the Stein equation in the two cases is similar. This organisation allows for a clear exposition with manageable calculations.

In Sect. 3, we obtain new bounds for the solution of the SVG Stein equation (Theorem 3.1 and Corollary 3.3) that have the correct dependence on the test function h to allow for Wasserstein ($\|h'\|$) and Kolmogorov ($\|\tilde{h}\|$) error bounds for SVG approximation via Stein's method. This task is arguably more technically demanding than for any other distribution for which this ingredient of Stein's method has been established. Indeed, Theorem 3.1 builds on the bounds of [15,16], the iterative technique of [9], and three papers on inequalities for integrals of modified Bessel functions [17,18,21] whose primary motivation was Stein's method for VG approximation. In Propositions 3.5 and 3.6, we note that higher-order derivatives of the solution cannot have a dependence on h of the form $\|\tilde{h}\|$ or $\|h'\|$.

In Sect. 4, we introduce (Definition 4.3) a distributional transformation, which we call the *centred equilibrium transformation of order r* that is natural in the context of SVG approximation via Stein's method. As our choice of name suggests, it generalises the centred equilibrium transformation [43], which is itself the natural analogue for Laplace approximation of the equilibrium transformation for exponential approximation [39]. In Theorem 4.10, we combine with the bounds of Sect. 3 to obtain general Wasserstein and Kolmogorov error bounds for SVG approximation. Our bounds are the SVG analogue of the general bounds of Theorem 3.1 of [39] that have been shown to be a useful tool for obtaining bounds for exponential approximation.

It should be noted that even with the new bounds of Sect. 3, with other coupling techniques, such as local couplings, more effort may be required to obtain Wasser-

stein and Kolmogorov bounds than would be the case for normal approximation, for example. This is due to the presence of the coefficient $\sigma^2 x$ in the leading derivative of the SVG Stein equation (1.2). This therefore provides motivation for introducing this distributional transformation.

In Sect. 5, we apply the results of Sects. 3 and 4 in four applications, these being: approximation of a general VG distribution by a SVG distribution; quantitative six moment theorems for SVG approximation of double Wiener-Itô integrals; SVG approximation of a statistic for binary sequence comparison (a special case of the D_2 statistic for alignment-free sequence comparison [4,31]); and Laplace approximation of a random sum of independent mean zero random variables. Our error bounds are given in the Wasserstein and Kolmogorov metrics, and in each case such bounds would not have been attainable by appealing to the present literature.

The rest of this paper is organised as follows. In Sect. 2, we introduce the class of SVG distributions and state some of their basic properties. Section 3 gives new bounds for the solution of the SVG Stein equation. In Sect. 4, we introduce a new distributional transformation, which we apply to give general bounds for SVG approximation in the Wasserstein and Kolmogorov metrics. In Sect. 5, we apply our results to obtain SVG approximations in several applications. Proofs of technical results are postponed until Sect. 6. Basic properties and inequalities for modified Bessel functions that are needed in this paper are collected in "Appendix A".

2 The Class of Symmetric Variance-Gamma Distributions

In this section, we introduce the class of symmetric variance-gamma (SVG) distributions and present some of their basic properties.

Definition 2.1 If $Z \sim \mathrm{VG}(r, 0, \sigma, \mu)$, for r, σ and μ defined as in (1.1), then Z is said to have a *symmetric variance-gamma* distribution. We write $Z \sim \mathrm{SVG}(r, \sigma, \mu)$.

Setting $\theta = 0$ in (1.1) gives the p.d.f. of $Z \sim \mathrm{SVG}(r, \sigma, \mu)$:

$$p(x) = \frac{1}{\sigma \sqrt{\pi} \Gamma(\frac{r}{2})} \left(\frac{|x - \mu|}{2\sigma} \right)^{\frac{r-1}{2}} K_{\frac{r-1}{2}} \left(\frac{|x - \mu|}{\sigma} \right), \quad x \in \mathbb{R}, \tag{2.1}$$

where $K_\nu(x)$ is a modified Bessel function of the second kind. The parameter r is known as the *scale* parameter. As r increases, the distribution becomes more rounded around its peak value μ (as can be seen from (2.3) below). The parameter σ is called the *tail* parameter. As σ decreases, the tails decay more quickly (see (2.2)). The parameter μ is the *location* parameter. Calculations can often be simplified by using the basic relation that if $Z \sim \mathrm{SVG}(r, 1, 0)$, then $\sigma Z + \mu \sim \mathrm{SVG}(r, \sigma, \mu)$. The $\mathrm{SVG}(r, 1, 0)$ distribution is in a sense the *standard* symmetric variance-gamma distribution.

The presence of the modified Bessel function makes (2.1) difficult to parse at first inspection. The following asymptotic formulas help in this regard. Applying (A.6) to (2.1) gives that, for all $r > 0$, $\sigma > 0$ and $\mu \in \mathbb{R}$,

$$p(x) \sim \frac{1}{(2\sigma)^{\frac{r}{2}}\Gamma(\frac{r}{2})}|x|^{\frac{r}{2}-1}e^{-|x-\mu|/\sigma}, \quad |x| \to \infty. \tag{2.2}$$

Similarly, applying (A.4) to (2.1) (see [15]) gives that

$$p(x) \sim \begin{cases} \dfrac{1}{\sigma\sqrt{\pi}}\dfrac{\Gamma(\frac{r-1}{2})}{\Gamma(\frac{r}{2})}, & x \to \mu, \ r > 1, \\[2ex] -\dfrac{1}{\pi\sigma}\log|x-\mu|, & x \to \mu, \ r = 1, \\[2ex] \dfrac{1}{(2\sigma)^r\sqrt{\pi}}\dfrac{\Gamma(\frac{1-r}{2})}{\Gamma(\frac{r}{2})}|x-\mu|^{r-1}, & x \to \mu, \ 0 < r < 1. \end{cases} \tag{2.3}$$

The density thus has a singularity at $x = \mu$ if $r \leq 1$. In fact, for any parameter, the SVG(r, σ, μ) distribution is unimodal with mode μ. This can be seen from the fact that the function $x^\nu K_\nu(x)$ is a decreasing function of x in the interval $(0, \infty)$ for all $\nu \in \mathbb{R}$ (see (A.8)).

The SVG distribution has a fundamental representation in terms of independent normal and gamma random variables ([29], Proposition 4.1.2). Let $X \sim \Gamma(\frac{r}{2}, \frac{1}{2})$ (with p.d.f. $\frac{1}{2^{r/2}\Gamma(r/2)}x^{r/2-1}e^{-x/2}$, $x > 0$) and $Y \sim N(0, 1)$ be independent. Then $\mu + \sigma\sqrt{X}Y \sim \text{SVG}(r, \sigma, \mu)$.

The SVG distribution has moment generating function $M(t) = (1 + \sigma^2 t^2)^{-r/2}$, $t \in \mathbb{R}$, and therefore has moments of arbitrary order. In particular, the mean and variance of $Z \sim \text{SVG}(r, \sigma, \mu)$ are given by

$$\mathbb{E}Z = \mu, \quad \text{Var}(Z) = r\sigma^2. \tag{2.4}$$

Perhaps surprisingly, this author could not find a formula for the absolute centred moments of the SVG(r, σ, μ) distribution in the literature. The result and its simple proof are given here.

Proposition 2.2 *Let $Z \sim \text{SVG}(r, \sigma, \mu)$. Then, for $k > 0$,*

$$\mathbb{E}|Z - \mu|^k = \frac{2^{\frac{k}{2}}\sigma^k}{\sqrt{\pi}}\frac{\Gamma(\frac{r+k}{2})\Gamma(\frac{k+1}{2})}{\Gamma(\frac{r}{2})}. \tag{2.5}$$

Proof We follow the approach given in Proposition 4.1.6 of [29] to obtain the moments of the SVG$(r, \sigma, 0)$ distribution. Recall that $Z - \mu =_d \sigma\sqrt{X}Y$, where $X \sim \Gamma(\frac{r}{2}, \frac{1}{2})$ and $Y \sim N(0, 1)$ are independent. Therefore

$$\mathbb{E}|Z - \mu|^k = \sigma^k \mathbb{E}[X^{\frac{k}{2}}]\mathbb{E}|Y|^k,$$

whence the result follows on using the standard formulas $\mathbb{E}X^s = \Gamma(\frac{r}{2} + s)/\Gamma(\frac{r}{2})$ and $\mathbb{E}|Y|^k = \frac{1}{\sqrt{\pi}}2^{\frac{k}{2}}\Gamma(\frac{k+1}{2})$. □

In interpreting Corollary 5.4, it will be useful to note the following formulas for the moments and cumulants of $Y \sim \mathrm{SVG}(r, \sigma, 0)$ ([12], Lemma 3.6):

$$\mathbb{E}Y^2 = r\sigma^2, \quad \mathbb{E}Y^4 = 3\sigma^4 r(r+2), \quad \mathbb{E}Y^6 = 15\sigma^6 r(r+2)(r+4),$$
$$\kappa_2(Y) = r\sigma^2, \quad \kappa_4(Y) = 6r\sigma^4, \quad \kappa_6(Y) = 120r\sigma^6,$$

with the odd order moments and cumulants all being equal to zero.

Lastly, we note that the class of SVG distributions contains several classical distributions as special or limiting cases ([16], Proposition 1.2).

1. Let X_r have the $\mathrm{SVG}(r, \frac{\sigma}{\sqrt{r}}, \mu)$ distribution. Then X_r converges in distribution to a $N(\mu, \sigma^2)$ random variable in the limit $r \to \infty$.
2. A $\mathrm{SVG}(2, \sigma, \mu)$ random variable has the $\mathrm{Laplace}(\mu, \sigma)$ distribution with p.d.f. $p(x) = \frac{1}{2\sigma} e^{-|x-\mu|/\sigma}$, $x \in \mathbb{R}$.
3. Let $X_1, \ldots, X_r$ and $Y_1, \ldots, Y_r$ be independent standard normal random variables. Then $\sigma \sum_{k=1}^r X_k Y_k$ has the $\mathrm{SVG}(r, \sigma, 0)$ distribution.
4. Suppose that $X \sim \Gamma(r, \lambda)$ and $Y \sim \Gamma(r, \lambda)$ are independent. Then the random variable $X - Y$ has the $\mathrm{SVG}(2r, \lambda^{-1}, 0)$ distribution.

3 Bounds for the Solution of the Stein Equation

In this section, we obtain bounds for the solution of the SVG Stein equation (that is (1.2) with $\theta = 0$) which have the correct dependence on the test function h to allow for Wasserstein and Kolmogorov distance bounds for SVG approximation via Stein's method.

For ease of exposition, in our proofs, we shall analyse the solution of the $\mathrm{SVG}(r, 1, 0)$ Stein equation. The general case follows from that fact that $\mathrm{SVG}(r, \sigma, \mu) =_d \mu + \sigma \mathrm{SVG}(r, 1, 0)$ and a simple rescaling and translation. The solution of the $\mathrm{SVG}(r, 1, 0)$ Stein equation is then

$$f(x) = -\frac{K_\nu(|x|)}{|x|^\nu} \int_0^x |t|^\nu I_\nu(|t|) \tilde{h}(t) \, dt - \frac{I_\nu(|x|)}{|x|^\nu} \int_x^\infty |t|^\nu K_\nu(|t|) \tilde{h}(t) \, dt \qquad (3.1)$$

$$= -\frac{K_\nu(|x|)}{|x|^\nu} \int_0^x |t|^\nu I_\nu(|t|) \tilde{h}(t) \, dt + \frac{I_\nu(|x|)}{|x|^\nu} \int_{-\infty}^x |t|^\nu K_\nu(|t|) \tilde{h}(t) \, dt, \qquad (3.2)$$

where $\nu = \frac{r-1}{2}$ and $\tilde{h}(x) = h(x) - \mathbb{E}h(Z)$ for $Z \sim \mathrm{SVG}(r, 1, 0)$. The equality between (3.1) and (3.2) follows because $|t|^\nu K_\nu(|t|)$ is proportional to the $\mathrm{SVG}(r, 1, 0)$ density. The equality is very useful, because it means that we will be able to restrict our attention to bounding the solution in the region $x \geq 0$, from which a bound for all $x \in \mathbb{R}$ is immediate.

We now note two useful bounds due to [16] for the solution of the $\mathrm{SVG}(r, \sigma, \mu)$ Stein equation that will be used in the proof of Theorem 3.1 and some of the applications of Sect. 5. For bounded and measurable $h : \mathbb{R} \to \mathbb{R}$,

$$\|f\| \le \frac{1}{\sigma}\left(\frac{1}{r} + \frac{\pi\Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})}\right)\|\tilde{h}\|, \tag{3.3}$$

$$\|f'\| \le \frac{2}{\sigma^2 r}\|\tilde{h}\|. \tag{3.4}$$

Let us now state the main result of this section.

Theorem 3.1 *Suppose that* $h : \mathbb{R} \to \mathbb{R}$ *is bounded and measurable. Let* f *be the solution of the* SVG(r, σ, μ) *Stein equation. Then*

$$\|(x-\mu)f(x)\| \le \left(\frac{3}{2} + \frac{1}{2r}\right)\|\tilde{h}\|, \tag{3.5}$$

$$\|(x-\mu)f'(x)\| \le \frac{1}{\sigma}\left(1 + \frac{1}{2r}\right)\|\tilde{h}\|, \tag{3.6}$$

$$\|(x-\mu)f''(x)\| \le \frac{1}{2\sigma^2}\left(9 + \frac{1}{r}\right)\|\tilde{h}\|. \tag{3.7}$$

Now suppose that $h : \mathbb{R} \to \mathbb{R}$ *is Lipschitz. Then*

$$\|f\| \le \frac{7}{2}\|h'\|, \tag{3.8}$$

$$\|f'\| \le \frac{9}{2\sigma}\left(\frac{1}{r+1} + \frac{\pi\Gamma(\frac{r+1}{2})}{2\Gamma(\frac{r}{2}+1)}\right)\|h'\|, \tag{3.9}$$

$$\|f''\| \le \frac{9}{\sigma^2(r+1)}\|h'\|, \tag{3.10}$$

and also

$$\|(x-\mu)f'(x)\| \le \frac{9}{2}\left(\frac{3}{2} + \frac{1}{2(r+1)}\right)\|h'\|, \tag{3.11}$$

$$\|(x-\mu)f''(x)\| \le \frac{9}{2\sigma}\left(1 + \frac{1}{2(r+1)}\right)\|h'\|, \tag{3.12}$$

$$\|(x-\mu)f^{(3)}(x)\| \le \frac{9}{4\sigma^2}\left(9 + \frac{1}{r+1}\right)\|h'\|. \tag{3.13}$$

Proof As noted above, for ease of notation, we set $\sigma = 1$ and $\mu = 0$. The bounds for the general case, as stated in the theorem, follow from a simple change in variables; see the proof of Theorem 3.6 of [16]. We also recall that it suffices to obtain bounds in the region $x \ge 0$.

Let us first establish the bound for $\|f\|$, which we will need to obtain several of the other bounds. By the mean value theorem, $|\tilde{h}(x)| \le \|h'\|(|x| + \mathbb{E}|Z|)$, where $Z \sim$ SVG$(r, 1, 0)$. From (2.5) we have $\mathbb{E}|Z| = \frac{2}{\sqrt{\pi}}\Gamma(\frac{r+1}{2})/\Gamma(\frac{r}{2})$. Now, on using inequalities (A.16), (A.18), (A.17) and (A.19) to obtain the second inequality we have, for $x \ge 0$,

$$|f(x)| \le \|h'\| \left\{ \frac{K_\nu(x)}{x^\nu} \int_0^x (t + \mathbb{E}|Z|) t^\nu I_\nu(t)\,dt + \frac{I_\nu(x)}{x^\nu} \int_x^\infty (t + \mathbb{E}|Z|) t^\nu K_\nu(t)\,dt \right\}$$

$$\le \|h'\| \left\{ \frac{1}{2} + \mathbb{E}|Z| \cdot \frac{1}{2\nu + 1} + 1 + \mathbb{E}|Z| \cdot \frac{\sqrt{\pi}\,\Gamma(\nu + \frac{1}{2})}{2\Gamma(\nu + 1)} \right\}$$

$$= \|h'\| \left\{ \frac{3}{2} + \frac{2}{\sqrt{\pi}} \frac{\Gamma(\nu + 1)}{\Gamma(\nu + \frac{1}{2})} \left(\frac{1}{2\nu + 1} + \frac{\sqrt{\pi}\,\Gamma(\nu + \frac{1}{2})}{2\Gamma(\nu + 1)} \right) \right\}$$

$$= \|h'\| \left\{ \frac{5}{2} + \frac{\Gamma(\nu + 1)}{\sqrt{\pi}\,\Gamma(\nu + \frac{3}{2})} \right\},$$

where we used the standard formula $u\Gamma(u) = \Gamma(u + 1)$ to obtain the final equality. Now, $\frac{\Gamma(\nu+1)}{\Gamma(\nu+\frac{3}{2})}$ is a decreasing function of ν for $\nu > -\frac{1}{2}$ (see [26]), and so is bounded above by $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ for all $\nu > -\frac{1}{2}$. Hence, $|f(x)| \le \frac{7}{2}\|h'\|$ for all $x \ge 0$, which is sufficient to prove (3.8).

The bounds for $\|f'\|$ and $\|f''\|$ can be obtained through an application of the iterative technique of [9]. Differentiating both sides of the SVG$(r, 1, 0)$ Stein equation (1.2) gives

$$xf^{(3)}(x) + (r + 1)f''(x) - xf'(x) = h'(x) + f(x). \tag{3.14}$$

We recognise (3.14) as the SVG$(r + 1, 1, 0)$ Stein equation, applied to f', with test function $h'(x) + f(x)$. We note that the test function $h'(x) + f(x)$ has mean zero with respect to the random variable $Y \sim$ SVG$(r+1, 1, 0)$. This fact will be required in order to later apply inequalities (3.3) and (3.4). Since h is Lipschitz, by inequality (3.8), we have that $\mathbb{E}|h'(Y) + f(Y)| < \infty$, and in particular as (3.14) is the SVG$(r+1, 1, 0)$ Stein equation applied to f', we have that $\mathbb{E}[Yf^{(3)}(Y) + (r + 1)f''(Y) - Yf'(Y)] = 0$, and thus $\mathbb{E}[h'(Y) + f(Y)] = 0$. Therefore, applying inequalities (3.3) and (3.4), respectively, with r replaced by $r + 1$ and test function $h'(x) + f(x)$ gives

$$\|f'\| = \left(\frac{1}{r + 1} + \frac{\pi\Gamma\left(\frac{r+1}{2}\right)}{2\Gamma\left(\frac{r}{2} + 1\right)} \right) \|h'(x) + f(x)\|$$

$$\le \left(\frac{1}{r + 1} + \frac{\pi\Gamma\left(\frac{r+1}{2}\right)}{2\Gamma\left(\frac{r}{2} + 1\right)} \right) \left(\|h'\| + \|f\| \right) \le \frac{9}{2} \left(\frac{1}{r + 1} + \frac{\pi\Gamma\left(\frac{r+1}{2}\right)}{2\Gamma\left(\frac{r}{2} + 1\right)} \right) \|h'\|,$$

$$\|f''\| \le \frac{2}{r + 1} \left(\|h'\| + \|f\| \right) \le \frac{9}{r + 1} \|h'\|,$$

where we used (3.8) to bound $\|f\|$.

Let us now establish the bounds (3.5), (3.6) and (3.7). On using inequalities (A.20) and (A.21) to obtain the second inequality, we obtain, for $x \ge 0$,

$$|xf(x)| = \left| \frac{K_\nu(x)}{x^{\nu-1}} \int_0^x t^\nu I_\nu(t)\tilde{h}(t)\,\mathrm{d}t - \frac{I_\nu(x)}{x^{\nu-1}} \int_x^\infty t^\nu K_\nu(t)\tilde{h}(t)\,\mathrm{d}t \right|$$

$$\leq \|\tilde{h}\| \left\{ \frac{K_\nu(x)}{x^{\nu-1}} \int_0^x t^\nu I_\nu(t)\,\mathrm{d}t + \frac{I_\nu(x)}{x^{\nu-1}} \int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \right\}$$

$$\leq \|\tilde{h}\| \left(\frac{\nu+1}{2\nu+1} + 1 \right) = \frac{3\nu+2}{2\nu+1}\|\tilde{h}\| = \frac{3r+1}{2r}\|\tilde{h}\| = \left(\frac{3}{2} + \frac{1}{2r} \right)\|\tilde{h}\|.$$

On using the differentiation formulas (A.9) and (A.10) and inequalities (A.22) and (A.23), we obtain, for $x \geq 0$,

$$|xf'(x)| = \left| \frac{K_{\nu+1}(x)}{x^{\nu-1}} \int_0^x t^\nu I_\nu(t)\tilde{h}(t)\,\mathrm{d}t + \frac{I_{\nu+1}(x)}{x^{\nu-1}} \int_x^\infty t^\nu K_\nu(t)\tilde{h}(t)\,\mathrm{d}t \right|$$

$$\leq \|\tilde{h}\| \left\{ \frac{K_{\nu+1}(x)}{x^{\nu-1}} \int_0^x t^\nu I_\nu(t)\,\mathrm{d}t + \frac{I_{\nu+1}(x)}{x^{\nu-1}} \int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \right\}$$

$$\leq \|\tilde{h}\| \left(\frac{\nu+1}{2\nu+1} + \frac{1}{2} \right) = \frac{4\nu+3}{2(2\nu+1)}\|\tilde{h}\| = \frac{2r+1}{2r}\|\tilde{h}\| = \left(1 + \frac{1}{2r} \right)\|\tilde{h}\|.$$

Since it suffices to consider $x \geq 0$, we have proved (3.5) and (3.6). Now, from the SVG$(r, 1, 0)$ Stein equation we have that, for $x \in \mathbb{R}$,

$$|xf''(x)| = |\tilde{h}(x) - rf'(x) + xf(x)| \leq \|\tilde{h}\| + r\|f'\| + \|xf(x)\|.$$

Applying (3.4) to bound $\|f'\|$ and (3.5) to bound $\|xf(x)\|$ yields the bound

$$\|xf''(x)\| \leq \left(1 + r \cdot \frac{2}{r} + \left(\frac{3}{2} + \frac{1}{2r} \right) \right)\|\tilde{h}\| = \frac{1}{2}\left(9 + \frac{1}{r} \right)\|\tilde{h}\|.$$

We now bound $\|xf'(x)\|$, $\|xf''(x)\|$ and $\|xf^{(3)}(x)\|$ for Lipschitz h using the iterative technique of [9] similarly to how we obtained inequalities (3.9) and (3.10). We recall that (3.14) is the SVG$(r+1, 1, 0)$ Stein equation, applied to f', with test function $h'(x) + f(x)$, which we also recall has mean zero with respect to the SVG$(r+1, 1, 0)$ measure. Therefore, applying inequalities (3.5), (3.6) and (3.7), respectively, with r replaced by $r+1$ and test function $h'(x) + f(x)$ gives

$$\|xf'(x)\| \leq \left(\frac{3}{2} + \frac{1}{2(r+1)} \right)\|h'(x) + f(x)\|$$

$$\leq \left(\frac{3}{2} + \frac{1}{2(r+1)} \right)(\|h'\| + \|f\|) \leq \frac{9}{2}\left(\frac{3}{2} + \frac{1}{2(r+1)} \right)\|h'\|,$$

$$\|xf''(x)\| \leq \left(1 + \frac{1}{2(r+1)} \right)(\|h'\| + \|f\|) \leq \frac{9}{2}\left(1 + \frac{1}{2(r+1)} \right)\|h'\|,$$

$$\|xf^{(3)}(x)\| \leq \frac{1}{2}\left(9 + \frac{1}{r+1} \right)(\|h'\| + \|f\|) \leq \frac{9}{4}\left(9 + \frac{1}{r+1} \right)\|h'\|,$$

where we used inequality (3.8) to bound $\|f\|$. The proof is complete. $\qquad\square$

In the following corollary, we apply some of the estimates of Theorem 3.1 to bound some useful quantities. We shall make use of these bounds in Sect. 4. It will be convenient to define the operator T_r by $T_r f(x) = xf'(x) + rf(x)$.

Corollary 3.2 *Let f be the solution of the* SVG$(r, \sigma, 0)$ *Stein equation. Then for h :* $\mathbb{R} \to \mathbb{R}$ *bounded and measurable, and Lipschitz, respectively,*

$$\sigma^2 \|T_r f'\| \leq \left(\frac{5}{2} + \frac{1}{2r}\right)\|\tilde{h}\|, \tag{3.15}$$

$$\sigma^2 \|(T_r f')'\| \leq \frac{9}{4}\left(5 + \frac{1}{r+1}\right)\|h'\|. \tag{3.16}$$

Proof As f satisfies $\sigma^2 T_r f'(x) = \tilde{h}(x) + xf(x)$, we have $\sigma^2 (T_r f')'(x) = h'(x) + xf'(x) + f(x)$. From the triangle inequality and the estimates of Theorem 3.1,

$$\sigma^2 \|T_r f'\| \leq \|\tilde{h}\| + \|xf(x)\| \leq \left(\frac{5}{2} + \frac{1}{2r}\right)\|\tilde{h}\|,$$

$$\sigma^2 \|(T_r f')'\| \leq \|h'\| + \|xf'(x)\| + \|f\| \leq \left\{1 + \frac{9}{2}\left(\frac{3}{2} + \frac{1}{2(r+1)}\right) + \frac{7}{2}\right\}\|h'\|,$$

which proves the result. □

Corollary 3.3 *Let f_z denote the solution of the* SVG(r, σ, μ) *Stein equation with test function $h_z(x) = \mathbf{1}(x \leq z)$. Then*

$$\|f_z\| \leq \frac{1}{\sigma}\left(\frac{1}{r} + \frac{\pi\Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})}\right), \quad \|f_z'\| \leq \frac{2}{\sigma^2 r}, \quad \sigma^2 \|T_r f_z'\| \leq \frac{5}{2} + \frac{1}{2r},$$

$$\|(x - \mu)f_z(x)\| \leq \frac{3}{2} + \frac{1}{2r}, \quad \|(x - \mu)f_z'(x)\| \leq \frac{1}{\sigma}\left(1 + \frac{1}{2r}\right),$$

$$\|(x - \mu)f_z''(x)\| \leq \frac{1}{2\sigma^2}\left(9 + \frac{1}{r}\right).$$

Proof Apply the inequality $\|\tilde{h}_z\| \leq 1$ to the bounds (3.3), (3.4), (3.15), (3.5), (3.6) and (3.7), respectively. □

Remark 3.4 For the normal [6] and exponential [5] Stein equations, because the solution of the Stein equation with test function $h_z(x) = \mathbf{1}(x \leq z)$ can be expressed in terms of elementary functions, a detailed analysis of the solution yields bounds with smaller constants that would be obtained by first working with a general bounded test function h and then bounding $\|\tilde{h}_z\| \leq 1$. However, because of the presence of modified Bessel functions in the solution, such improvements would be more difficult to obtain here.

It is natural to ask whether, for all $z \in \mathbb{R}$, a bound of the form $\|f_z''\| \leq C_{r,\sigma}$ could be obtained for the solution f_z. The following proposition, which is proved in Sect. 6, shows that this is not possible.

Proposition 3.5 $f'_\mu(x)$ *has a discontinuity at* $x = \mu$.

Similarly, one may ask whether a bound of the form $\| f^{(3)} \| \leq C_{r,\sigma} \|h'\|$ could be obtained for all Lipschitz $h : \mathbb{R} \to \mathbb{R}$. The following proposition, which is proved in Sect. 6, again shows this is not possible (see [11] for similar results that apply to solutions of Stein equations for a wide class of distributions). Our approach differs from that of Proposition 3.5 in that we do not find a Lipschitz test function h for which f'' has a discontinuity. This would be more tedious to establish for f'' than for f' and instead we consider a highly oscillating test function and perform an asymptotic analysis.

Proposition 3.6 *Let f be the solution of the* SVG(r, σ, μ) *Stein equation. Then, there does not exist a constant $M_{r,\sigma} > 0$ such that $\| f^{(3)} \| \leq M_{r,\sigma} \|h'\|$ for all Lipschitz* $h : \mathbb{R} \to \mathbb{R}$.

Remark 3.7 (i) Throughout this remark, we set $\mu = 0$. The bounds (3.4) and (3.10) are of order r^{-1} as $r \to \infty$. This is indeed the optimal order, which can be seen by the following argument, which is similar to the one given in Remark 2.2 of [24] to show that the rate in their bound for solution of the gamma Stein equation was optimal.

Evaluating both sides of the SVG$(r, \sigma, 0)$ Stein equation at $x = 0$ gives $f'(0) = \frac{1}{\sigma^2 r} \tilde{h}(0)$. Also, evaluating both sides of (3.14) (with general σ) at $x = 0$ gives that $f''(0) = \frac{1}{\sigma^2(r+1)} \big(h'(0) + f(0) \big)$, from which we conclude that the $O(r^{-1})$ rate in (3.10) is also optimal.

(ii) The bound (3.3) for $\|f\|$ is of order $r^{-\frac{1}{2}}$ as $r \to \infty$. Indeed, for $r > 1$,

$$\sqrt{\frac{2}{r}} < \frac{\Gamma(\frac{r}{2})}{\Gamma(\frac{r+1}{2})} < \sqrt{\frac{2}{r - \frac{1}{2}}}, \tag{3.17}$$

which follows from the inequalities $\frac{\Gamma(x+\frac{1}{2})}{\Gamma(x+1)} > (x + \frac{1}{2})^{-\frac{1}{2}}$ for $x > 0$ (see [25]) and $\frac{\Gamma(x+\frac{1}{2})}{\Gamma(x+1)} < (x + \frac{1}{4})^{-\frac{1}{2}}$ for $x > -\frac{1}{4}$ (see [13]). The $O(r^{-\frac{1}{2}})$ rate is optimal, which can be seen as follows. Take h to be $h(x) = 1$ if $x \geq 0$ and $h(x) = -1$ if $x < 0$, so that $\tilde{h}(x) = h(x)$. Then

$$f(0+) = -\lim_{x \downarrow 0} \left\{ \frac{1}{\sigma^2 x^\nu} K_\nu \left(\frac{x}{\sigma} \right) \int_0^x t^\nu I_\nu \left(\frac{t}{\sigma} \right) dt \right\}$$

$$- \lim_{x \downarrow 0} \left\{ \frac{1}{\sigma^2 x^\nu} I_\nu \left(\frac{x}{\sigma} \right) \int_x^\infty t^\nu K_\nu \left(\frac{t}{\sigma} \right) dt \right\}$$

$$= \frac{\sqrt{\pi}\,\Gamma(\nu + \frac{1}{2})}{2\sigma^2 \Gamma(\nu + 1)} = \frac{\sqrt{\pi}\,\Gamma(\frac{r}{2})}{2\sigma^2 \Gamma(\frac{r+1}{2})}.$$

Here, we used that the first limit is equal to zero by the asymptotic formulas (A.3) and (A.4). We computed the second limit using the asymptotic formula (A.3) and that the integrand is proportional to the density of the SVG$(2\nu + 1, \sigma, 0)$ distribution. Therefore, by (3.17), we conclude that the optimal rate is order $r^{-\frac{1}{2}}$ as $r \to \infty$.

(iii) Arguing as we did in part (i), we have that $f(0) = \sigma^2(r+1)f''(0) - h'(0)$, which for a general Lipschitz test function h is $O(1)$ (see bound (3.10)), and so the bound (3.8) is of optimal order.

(iv) In the light of inequalities (3.8)–(3.10), one might expect inequalities (3.6) and (3.7) to be of lower than (3.5) as $r \to \infty$. However, this is not the case. A calculation involving L'Hôpital's rule (which is given in Sect. 6) shows that, for any bounded $h : \mathbb{R} \to \mathbb{R}$,

$$\lim_{x \to \infty} xf(x) = -\tilde{h}(\infty), \quad \lim_{x \to -\infty} xf(x) = \tilde{h}(-\infty), \tag{3.18}$$

and from the SVG$(r, \sigma, 0)$ Stein equation and inequality (3.9) we obtain

$$\lim_{x \to -\infty} xf''(x) = \frac{1}{\sigma^2}\Big[\tilde{h}(-\infty) + \lim_{x \to -\infty} xf(x)\Big] = \frac{2}{\sigma^2}\tilde{h}(-\infty).$$

Thus, inequalities (3.5) and (3.7) are of optimal order in r. We expect this to also be the case for inequalities (3.11)–(3.13), although verifying this would involve a more detailed analysis, which we omit for space reasons.

4 The Centred Equilibrium Transformation of Order r

In this section, we introduce a new distributional transformation and apply it to obtain general Wasserstein and Kolmogorov error bounds for SVG approximation.

We begin with the following proposition which relates the Kolmogorov and Wasserstein distances between a general distribution and a SVG distribution. This proposition is of interest, because Wasserstein distance bounds are often easier to obtain than Kolmogorov distance bounds through Stein's method. The proof is deferred until Sect. 6.

Proposition 4.1 *Let $Z \sim \mathrm{SVG}(r, \sigma, \mu)$. Then, for any random variable W:*
(i) If $r > 1$,

$$d_{\mathrm{K}}(W, Z) \leq \sqrt{\frac{1}{\sigma\sqrt{\pi}} \frac{\Gamma\left(\frac{r-1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)} d_{\mathrm{W}}(W, Z)}. \tag{4.1}$$

(ii) Suppose that $\sigma^{-1}d_{\mathrm{W}}(W, Z) < 0.676$. Then, if $r = 1$,

$$d_{\mathrm{K}}(W, Z) \leq \left\{2 + \log\left(\frac{2}{\sqrt{\pi}}\right) + \frac{1}{2}\log\left(\frac{\sigma}{d_{\mathrm{W}}(W, Z)}\right)\right\}\sqrt{\frac{d_{\mathrm{W}}(W, Z)}{\pi\sigma}}. \tag{4.2}$$

(iii) If $0 < r < 1$,

$$d_{\mathrm{K}}(W, Z) \leq 2\left(\frac{\Gamma\left(\frac{1-r}{2}\right)}{\sqrt{\pi}2^{r-1}\Gamma\left(\frac{r}{2}\right)}\right)^{\frac{1}{r+1}}\left(\sigma^{-1}d_{\mathrm{W}}(W, Z)\right)^{\frac{r}{r+1}}. \tag{4.3}$$

Remark 4.2 (i) If $\sigma^{-1}d_W(W, Z) = 0.676$, then the upper bound in part (ii) is equal to 1.075, and is therefore uninformative.

(ii) Recall that $N(\mu, \sigma^2) =_d \lim_{r \to \infty} \text{SVG}(r, \frac{\sigma}{\sqrt{r}}, \mu)$. Therefore from (4.1) and the limit $\lim_{x \to \infty} \frac{\sqrt{x}\Gamma(x - \frac{1}{2})}{\Gamma(x)} = 1$ (see (3.17)), we recover the inequality (with obvious abuse of notation)

$$d_K(W, N(\mu, \sigma^2)) \le \left(\frac{2}{\pi \sigma^2}\right)^{\frac{1}{4}} \sqrt{d_W(W, N(\mu, \sigma^2))},$$

which is a special case of part 2 of Proposition 1.2 of [48]. It is known (see [6], p. 48) that this bound gives the optimal rate under some conditions, but in other applications the rate is suboptimal. Proposition 5.1 gives an application in which the inequalities (4.1), (4.2) and (4.3) are not of optimal rate in $\delta = d_W(W, Z)$; see Remark 5.2.

As in Sect. 3, we define the operator T_r by $T_r f(x) = xf'(x) + rf(x)$. We also denote $D = \frac{d}{dx}$. From now until the end of this section, we set $\mu = 0$.

Definition 4.3 Let W be a random variable with mean zero and variance $0 < r\sigma^2 < \infty$. We say that W^{V_r} has the *W-centred equilibrium distribution of order r* if

$$\mathbb{E}Wf(W) = \sigma^2\mathbb{E}T_r f'(W^{V_r}) \tag{4.4}$$

for all twice-differentiable $f : \mathbb{R} \to \mathbb{R}$ such that the expectations in (4.4) exist.

As we shall see later, it is convenient to write $\text{Var}(W) = r\sigma^2$, because the variance of a $\text{SVG}(r, \sigma, 0)$ random variable is $r\sigma^2$. As the name suggests, the centred equilibrium distribution of order r generalises the centred equilibrium distribution of W, denoted by W^L, that was introduced by [43]. Its characterising equation is

$$\mathbb{E}f(W) - f(0) = \frac{1}{2}\mathbb{E}W^2\mathbb{E}f''(W^L). \tag{4.5}$$

We also refer the reader to [8] for a generalisation of (4.5) to all random variables W with finite second moment. The centred equilibrium distribution is itself the Laplace analogue of the equilibrium distribution that has been shown to be a useful tool in Stein's method for exponential approximation by [39]. We can see that $W^{V_2} = W^L$ by setting $f(x) = xg(x)$ in (4.5). For $r \ne 2$, a characterising equation of the form (4.5) is not useful. To see this, recall that the Stein operator for the $\text{SVG}(r, \sigma, 0)$ distribution is $\mathcal{A}f(x) = \sigma^2 xf''(x) + \sigma^2 rf'(x) - xf(x)$. Setting $f(x) = g(x)/x$ then gives

$$\mathcal{A}g(x) = \sigma^2 g''(x) + (r - 2)\sigma^2\left(\frac{g'(x)}{x} - \frac{g(x)}{x^2}\right) - g(x),$$

which has a singularity at $x = 0$ if $r \ne 2$.

We also note that $W^{V_1} = W^{*(2)}$, where $W^{*(2)}$ has the W-zero bias distribution of order 2 (see [19]). This distributional transformation is a natural generalisation of the zero bias transformation (defined below) to the setting of Stein's method for products

of independent standard normal random variables. We shall make use of this fact in Sect. 5.3.

We now obtain an inverse of the operator $T_r D$. This inverse operator will be used later in this section to establish properties of the centred equilibrium distribution of order r. Recall that the Beta$(r, 1)$ distribution has p.d.f. $p(x) = rx^{r-1}, 0 < x < 1$.

Lemma 4.4 *Let $B_r \sim$ Beta$(r, 1)$ and $U \sim U(0, 1)$ be independent, and define the operator G_r by $G_r f(x) = \frac{x}{r}\mathbb{E}f(xUB_r)$. Then, G_r is the right inverse of the operator $T_r D$ in the sense that*

$$T_r D G_r f(x) = f(x). \tag{4.6}$$

Suppose now that f is twice differentiable. Then, for any $r \geq 1$,

$$G_r T_r D f(x) = f(x) - f(0). \tag{4.7}$$

Therefore, G_r is the inverse of $T_r D$ when the domain of $T_r D$ is the space of all twice-differentiable functions f on $\mathbb{R}$ with $f(0) = 0$.

Proof We begin by obtaining a useful formula for $G_r f(x) = \frac{x}{r}\mathbb{E}f(xUB_r)$:

$$G_r f(x) = \frac{x}{r}\int_0^1\int_0^1 f(xub)rb^{r-1}\,\mathrm{d}b\,\mathrm{d}u = \int_0^x\int_0^t f(s)s^{r-1}t^{-r}\,\mathrm{d}s\,\mathrm{d}t. \tag{4.8}$$

We now use (4.8) to verify (4.6):

$$
\begin{aligned}
T_r D G_r f(x) &= T_r\left(x^{-r}\int_0^x f(s)s^{r-1}\,\mathrm{d}s\right)\\
&= x\left(-rx^{r-1}\int_0^x f(s)s^{r-1}\,\mathrm{d}s + x^{-r}\cdot f(x)x^{r-1}\right)\\
&\quad + rx^{-r}\int_0^x f(s)s^{r-1}\,\mathrm{d}s\\
&= f(x).
\end{aligned}
$$

Finally, we verify relation (4.7). We have

$$
\begin{aligned}
G_r T_r D f(x) &= \int_0^x\int_0^t \big(sf''(s) + rf'(s)\big)s^{r-1}t^{-r}\,\mathrm{d}s\,\mathrm{d}t\\
&= \int_0^x t^{-r}\int_0^t \big(s^r f'(s)\big)'\,\mathrm{d}s\,\mathrm{d}t = \int_0^x f'(t)\,\mathrm{d}t = f(x) - f(0),
\end{aligned}
$$

as required. $\qquad\square$

Before presenting some properties of the centred equilibrium distribution of order r, we recall two distributional transformations that are standard in the Stein's method literature. If W is a mean zero random variable with finite, nonzero variance σ^2, we

say that W^* has the W-zero biased distribution [27] if for all differentiable f for which $\mathbb{E}Wf(W)$ exists,

$$\mathbb{E}Wf(W) = \sigma^2 \mathbb{E}f'(W^*).$$

For any random variable W with finite second moment, we say that $W^{\square}$ has the W-square bias distribution ([6], pp. 34–35) if for all f such that $\mathbb{E}W^2 f(W)$ exists,

$$\mathbb{E}W^2 f(W) = \mathbb{E}W^2 \mathbb{E}f(W^{\square}).$$

When $EW = 0$, there is neat relationship between these distribution transformations: $W^* =_d UW^{\square}$, where $U \sim U(0, 1)$ is independent of $W^{\square}$ (this is a slight variant of Proposition 2.3 [6]; see [19], Proposition 3.2).

The following construction of W^{V_r} generalises Theorem 3.2 of [43]. Similar constructions for distributional transformations that are natural in the context in gamma and generalised gamma approximation can be found in [44] and [42].

Proposition 4.5 *Let W be a random variable with zero mean and finite, nonzero variance $r\sigma^2$, and let W^* have the W-zero bias distribution. Let $B_r \sim \text{Beta}(r, 1)$ be independent of W^*. Then, the random variable*

$$W^{V_r} =_d B_r W^*$$

has the centred equilibrium distribution of order r.

Proof Let $f \in C_c$, the collection of continuous functions with compact support. In Lemma 4.4 we defined the operator $G_r g(x) = \frac{x}{r}\mathbb{E}g(xUB_r)$ and showed that $T_r DG_r g(x) = g(x)$ for any g. We therefore have

$$\sigma^2 \mathbb{E}f(W^{V_r}) = \sigma^2 \mathbb{E}T_r DG_r f(W^{V_r}) = \mathbb{E}WG_r f(W) = \frac{1}{r}\mathbb{E}W^2 f(UB_r W)$$

$$= \frac{1}{r}\mathbb{E}W^2 \mathbb{E}f(UB_r W^{\square}) = \sigma^2 \mathbb{E}f(UB_r W^{\square}) = \sigma^2 \mathbb{E}f(B_r W^*).$$

Since the expectation of $f(W^{V_r})$ and $f(B_r W^*)$ is equal for all $f \in C_c$, the random variables W^{V_r} and $B_r W^*$ must be equal in distribution.   □

In the following proposition, we collect some useful properties of the centred equilibrium distribution of order r. As might be expected in the light of Proposition 4.5, some of these properties are quite similar to those given for the zero bias distribution in Lemma 2.1 of [27].

Proposition 4.6 *Let W be a mean zero variable with finite, nonzero variance $r\sigma^2$, and let W^{V_r} have the W-centred equilibrium distribution of order r in accordance with Definition 4.3.*

(i) The $\text{SVG}(r, \sigma, 0)$ distribution is the unique fixed point of the centred equilibrium transformation of order r.

(ii) The distribution of W^{V_r} is unimodal about zero and absolutely continuous with density

$$f_{W^{V_r}}(w) = \frac{1}{\sigma^2} \int_0^1 t^{r-2} \mathbb{E}[W\mathbf{1}(W > w/t)]\, \mathrm{d}t. \tag{4.9}$$

It follows that the support of W^{V_r} is the closed convex hull of the support of W and that W^{V_r} is bounded whenever W is bounded.

(iii) The centred equilibrium transformation of order r preserves symmetry.

(iv) For $p \geq 0$,

$$\mathbb{E}[(W^{V_r})^p] = \frac{\mathbb{E}W^{p+2}}{r\sigma^2(p+1)(p+r)} \quad and \quad \mathbb{E}|W^{V_r}|^p = \frac{\mathbb{E}|W|^{p+2}}{r\sigma^2(p+1)(p+r)}.$$

(v) For $c \in \mathbb{R}$, cW^{V_r} has the cW-centred equilibrium distribution of order r.

Proof (i) This is immediate from Definition 4.3 and the Stein characterisation for the SVG$(r, \sigma, 0)$ distribution given in Lemma 3.1 of [16].

(ii) Firstly, we note that, for fixed $t \in (0, 1)$, the expectation $\mathbb{E}[W\mathbf{1}(W > w/t)]$ is increasing for $w < 0$ and decreasing for $w > 0$. We therefore deduce that $p(w)$ is increasing for $w < 0$ and decreasing for $w > 0$. Now, from Proposition 4.5, we have that $W^{V_r} =_d B_r W^*$. Formula (4.9) then follows from the fact that X^* is absolutely continuous with density $f_{W^*}(w) = \mathbb{E}[W\mathbf{1}(W > w)]/\mathrm{Var}(W)$ (part (ii) of Lemma 2.1 of [27]) and the standard formula for computing the density of a product.

(iii) We follow the argument of part (iii) of Lemma 2.1 of [27]. Let w be a continuity point of a symmetric random variable W. Then, for fixed $t \in (0, 1)$, $\mathbb{E}[W\mathbf{1}(W > w/t)] = \mathbb{E}[-W\mathbf{1}(-W > w/t)] = -\mathbb{E}[W\mathbf{1}(W < -w/t)] = \mathbb{E}[W\mathbf{1}(W > -w/t)]$, using $\mathbb{E}W = 0$. It is now evident from (4.9) that $f_{W^{V_r}}(w) = f_{W^{V_r}}(-w)$ for almost all w. Therefore, there is a version of the $\mathrm{d}w$ density of W^{V_r} which is the same at w and $-w$ for almost all $w[\mathrm{d}w]$, and so W^{V_r} is symmetric.

(iv) Substitute w^{p+1} and $|w|^{p+1}$ for $f(w)$ in the characterising equation (4.4).

(v) Let g be a function such that $\mathbb{E}Wg(W)$ exists, and define $\tilde{g}(x) = cg(cx)$. Then $\tilde{g}^{(k)}(x) = c^{k+1}g^{(k)}(cx)$. As W^{V_r} has the W-centred equilibrium distribution of order r,

$$\mathbb{E}cWg(cW) = \mathbb{E}W\tilde{g}(W) = \sigma^2 \mathbb{E}T_r D\tilde{g}(W^{V_r}) = (c\sigma)^2 \mathbb{E}T_r Dg(cW^{V_r}).$$

Hence, cW^{V_r} has the cW-centred equilibrium distribution of order r. $\square$

We end this section by proving Theorem 4.10 below, which formalises the notion that if $\mathcal{L}(W)$ and $\mathcal{L}(W^{V_r})$ are approximately equal then W has an approximation SVG distribution. This theorem is the SVG analogue of Theorem 2.1 of [39], in which the Wasserstein and Kolmogorov error bounds are given in terms of the difference in absolute expectation between the random variable of interest W and its W-equilibrium transformation. We follow the approach of [39] and begin by stating three lemmas.

Lemma 4.7 *Let $Z \sim \mathrm{SVG}(r, \sigma, 0)$. Then, for any random variable W,*

$$\mathbb{P}(a \le W \le b) \le C_{r,\sigma,b-a} + 2d_{\mathrm{K}}(W, Z), \tag{4.10}$$

where

$$C_{r,\sigma,\alpha} = \begin{cases} \dfrac{\alpha}{2\sigma\sqrt{\pi}} \dfrac{\Gamma\left(\frac{r-1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)}, & r > 1, \\[3ex] \dfrac{\alpha}{\pi\sigma}\left[1 + \log\left(\dfrac{2\sigma}{\alpha}\right)\right], & r = 1, \\[3ex] \dfrac{\Gamma\left(\frac{1-r}{2}\right)}{\sqrt{\pi}\,2^r\,\Gamma\left(\frac{r}{2}+1\right)}\left(\dfrac{\alpha}{\sigma}\right)^r, & 0 < r < 1. \end{cases}$$

Proof Clearly,

$$\mathbb{P}(a \le W \le b) \le \mathbb{P}(a \le Z \le b) + 2d_{\mathrm{K}}(W, Z).$$

Since, for all $r > 0$ and $\sigma > 0$, the $\mathrm{SVG}(r, \sigma, 0)$ density $p(x)$ is an increasing function of x for $x < 0$ and a decreasing function of x for $x > 0$, we have that

$$\mathbb{P}(a \le Z \le b) \le \int_{-(b-a)/2}^{(b-a)/2} p(x)\,\mathrm{d}x = 2\int_0^{(b-a)/2} p(x)\,\mathrm{d}x. \tag{4.11}$$

To obtain (4.10), we bound the integral on the right-hand side of (4.11), treating the cases $r > 1$, $r = 1$ and $0 < r < 1$ separately. For $r > 1$ we bound the density $p(x)$ by $\frac{1}{2\sigma\sqrt{\pi}}\Gamma\left(\frac{r-1}{2}\right)/\Gamma\left(\frac{r}{2}\right)$ using (2.3) and then compute the trivial integral; for $r = 1$ we use inequality (6.7); and for $0 < r < 1$ we use inequality (6.8). This yields (4.10), as required. $\square$

The next lemma follows immediately from the estimates of Theorem 3.1 and Corollary 3.2, and the subsequent lemma is straightforward and we hence omit the proof.

Lemma 4.8 *For any $a \in \mathbb{R}$ and any $\epsilon > 0$, let*

$$h_{a,\epsilon}(x) := \epsilon^{-1}\int_0^\epsilon \mathbf{1}(x + s \le a)\,\mathrm{d}s. \tag{4.12}$$

Let $f_{a,\epsilon}$ be the solution of the $\mathrm{SVG}(r, \sigma, 0)$ Stein equation with test function $h_{a,\epsilon}$. Define $h_{a,0}(x) = \mathbf{1}(x \le a)$ and $f_{a,0}$ accordingly. Then

$$\|f_{a,\epsilon}\| \leq \frac{1}{\sigma}\left(\frac{1}{r} + \frac{\pi\Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})}\right), \tag{4.13}$$

$$\|xf_{a,\epsilon}(x)\| \leq \frac{3}{2} + \frac{1}{2r}, \tag{4.14}$$

$$\|xf'_{a,\epsilon}(x)\| \leq \frac{1}{\sigma}\left(1 + \frac{1}{2r}\right), \tag{4.15}$$

$$\sigma^2\|T_r f'_{a,\epsilon}\| \leq \frac{5}{2} + \frac{1}{2r}. \tag{4.16}$$

Lemma 4.9 *Let $Z \sim \mathrm{SVG}(r, \sigma, 0)$ and W be a real-valued random variable. Then, for any $\epsilon > 0$,*

$$d_{\mathrm{K}}(W, Z) \leq C_{r,\sigma,\epsilon} + \sup_{a\in\mathbb{R}} |\mathbb{E}h_{a,\epsilon}(W) - \mathbb{E}h_{a,\epsilon}(Z)|,$$

where $C_{r,\sigma,\epsilon}$ is defined as in Lemma 4.7 and $h_{a,\epsilon}$ is defined as in Lemma 4.8.

Theorem 4.10 *Let W be a mean zero random variable with variance $0 < r\sigma^2 < \infty$. Suppose that (W, W^{V_r}) is given on a joint probability space so that W^{V_r} has the W-centred equilibrium distribution of order r. Then*

$$d_{\mathrm{K}}(W, Z) \leq \left(2 + \frac{3}{r} + \frac{\pi\Gamma(\frac{r}{2})}{\Gamma(\frac{r+1}{2})}\right)\frac{\beta}{\sigma} + \frac{5}{2}C_{r,\sigma,4\beta} + \left(10 + \frac{2}{r}\right)\mathbb{P}(|W - W^{V_r}| > \beta), \tag{4.17}$$

where $C_{r,\sigma,4\beta}$ is defined as in Lemma 4.7. Also,

$$d_{\mathrm{K}}(W^{V_r}, Z) \leq \left(1 + \frac{3}{2r} + \frac{\pi\Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})}\right)\frac{\beta}{\sigma} + \left(3 + \frac{1}{r}\right)\mathbb{P}(|W - W^{V_r}| > \beta). \tag{4.18}$$

Suppose in addition that $\mathbb{E}|W|^3 < \infty$. Then

$$d_{\mathrm{W}}(W, Z) \leq \frac{9}{4}\left(5 + \frac{1}{r+1}\right)\mathbb{E}|W - W^{V_r}|, \tag{4.19}$$

$$d_{\mathrm{W}}(W^{V_r}, Z) \leq \frac{1}{4}\left(41 + \frac{9}{r+1}\right)\mathbb{E}|W - W^{V_r}|, \tag{4.20}$$

$$d_{\mathrm{K}}(W^{V_r}, Z) \leq \frac{1}{\sigma}\left(1 + \frac{3}{2r} + \frac{\pi\Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})}\right)\mathbb{E}|W - W^{V_r}|. \tag{4.21}$$

Proof For this proof, we shall write $\kappa = d_{\mathrm{K}}(W, Z)$. Let $\Delta := W - W^{V_r}$. Define $I_1 := \mathbf{1}(|\Delta| \leq \beta)$; note that W^{V_r} may not have finite second moment. Let f be the solution of the $\mathrm{SVG}(r, \sigma, 0)$ Stein equation with test function $h_{a,\epsilon}$, as defined in (4.12).

Then $\mathbb{E}T_r f'(W^{V_r})$ is well defined, because $\|T_r f'\| < \infty$ (see Lemma 4.8), and we have

$$\mathbb{E}[\sigma^2 T_r f'(W) - Wf(W)] = \sigma^2 \mathbb{E}[I_1(T_r f'(W) - T_r f'(W^{V_r}))]$$
$$+ \sigma^2 \mathbb{E}[(1 - I_1)(T_r f'(W) - T_r f'(W^{V_r}))]$$
$$=: J_1 + J_2.$$

Using (4.16) gives $|J_2| \leq 2 \times \left(\frac{5}{2} + \frac{1}{2r}\right)\mathbb{P}(|\Delta| > \beta)$. Arguing as we did at the start of the proof of Corollary 3.2 to obtain the second equality, and then using inequalities (4.13) and (4.15) and Lemma 4.7 in the last step gives

$$J_1 = \sigma^2 \mathbb{E}\left[I_1 \int_0^\Delta (T_r f')'(W + t)\, \mathrm{d}t\right]$$
$$= \mathbb{E}\left[I_1 \int_0^\Delta \left\{(W + t)f'(W + t) + f(W + t) - \epsilon^{-1}\mathbf{1}(a - \epsilon \leq W + t \leq a)\right\} \mathrm{d}t\right]$$
$$\leq \left(\|xf'(x)\| + \|f\|\right)\mathbb{E}|I_1\Delta| + \epsilon^{-1}\int_{-\beta}^0 \mathbb{P}(a - \epsilon \leq W + t \leq a)\, \mathrm{d}t$$
$$\leq \left(1 + \frac{3}{2r} + \frac{\pi\Gamma\left(\frac{r}{2}\right)}{2\Gamma\left(\frac{r+1}{2}\right)}\right)\frac{\beta}{\sigma} + \beta\epsilon^{-1}C_{r,\sigma,\epsilon} + 2\beta\epsilon^{-1}\kappa.$$

Similarly,

$$J_1 \geq -\left(\|xf'(x)\| + \|f\|\right)\mathbb{E}|I_1\Delta| - \epsilon^{-1}\int_{-\beta}^0 \mathbb{P}(a - \epsilon \leq W + t \leq a)\, \mathrm{d}t$$
$$\geq -\left(1 + \frac{3}{2r} + \frac{\pi\Gamma\left(\frac{r}{2}\right)}{2\Gamma\left(\frac{r+1}{2}\right)}\right)\frac{\beta}{\sigma} - \beta\epsilon^{-1}C_{r,\sigma,\epsilon} - 2\beta\epsilon^{-1}\kappa,$$

and so we conclude that

$$|J_1| \leq \left(1 + \frac{3}{2r} + \frac{\pi\Gamma\left(\frac{r}{2}\right)}{2\Gamma\left(\frac{r+1}{2}\right)}\right)\frac{\beta}{\sigma} + \beta\epsilon^{-1}C_{r,\sigma,\epsilon} + 2\beta\epsilon^{-1}\kappa.$$

Using Lemma 4.9 and taking $\epsilon = 4\beta$ now gives

$$\kappa \leq \left(5 + \frac{1}{r}\right)\mathbb{P}(|\Delta| > \beta) + \left(1 + \frac{3}{2r} + \frac{\pi\Gamma\left(\frac{r}{2}\right)}{2\Gamma\left(\frac{r+1}{2}\right)}\right)\frac{\beta}{\sigma} + (1 + \beta\epsilon^{-1})C_{r,\sigma,\epsilon}$$
$$+ 2\beta\epsilon^{-1}\kappa$$
$$\leq \left(5 + \frac{1}{r}\right)\mathbb{P}(|\Delta| > \beta) + \left(1 + \frac{3}{2r} + \frac{\pi\Gamma\left(\frac{r}{2}\right)}{2\Gamma\left(\frac{r+1}{2}\right)}\right)\frac{\beta}{\sigma} + \frac{5}{4}C_{r,\sigma,4\beta} + \frac{1}{2}\kappa,$$

whence on solving for κ yields (4.17).

Now let us prove (4.18). We can write

$$\mathbb{E}[\sigma^2 T_r f'(W^{V_r}) - W^{V_r} f(W^{V_r})] = \mathbb{E}[Wf(W) - W^{V_r} f(W^{V_r})]$$
$$= \mathbb{E}[I_1(Wf(W) - W^{V_r} f(W^{V_r}))] + \mathbb{E}[(1 - I_1)(Wf(W) - W^{V_r} f(W^{V_r}))].$$

Taylor expanding, applying the triangle inequality to $\|xf'(x) + f(x)\|$, and using the estimates (4.13), (4.14) and (4.15) then gives that

$$\mathbb{E}[\sigma^2 T_r f'(W^{V_r}) - W^{V_r} f(W^{V_r})]$$
$$\leq \|xf'(x) + f(x)\|\mathbb{E}|I_1\Delta| + 2\|xf(x)\|\mathbb{P}(|\Delta| > \beta)$$
$$\leq \frac{1}{\sigma}\left(1 + \frac{3}{2r} + \frac{\pi\Gamma\left(\frac{r}{2}\right)}{2\Gamma\left(\frac{r+1}{2}\right)}\right)\beta + \left(3 + \frac{1}{r}\right)\mathbb{P}(|\Delta| > \beta),$$

which gives (4.18).

Suppose now that $\mathbb{E}|W|^3 < \infty$, which, by part (iv) of Proposition 4.6, ensures that $\mathbb{E}|W^{V_r}| < \infty$. Let $h \in \mathcal{H}_W$. Then

$$\mathbb{E}h(W) - \mathbb{E}h(Z) = \mathbb{E}[\sigma^2 T_r f'(W) - Wf(W)] = \sigma^2\mathbb{E}[T_r f'(W) - T_r f'(W^{V_r})],$$

and by Taylor expansion, we have

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \leq \sigma^2\|(T_r f')'\|\mathbb{E}|W - W^{V_r}|.$$

On using the estimate (3.16) we obtain (4.19), as required. Also,

$$\left|\sigma^2\mathbb{E}[(T_r f')(W^{V_r}) - W^{V_r} f(W^{V_r})]\right| = \left|\mathbb{E}Wf(W) - \mathbb{E}W^{V_r} f(W^{V_r})\right|$$
$$\leq \|xf'(x) + f(x)\|\mathbb{E}|W - W^{V_r}|. \quad (4.22)$$

Applying the estimates (3.11) and (3.8) to (4.22) yields (4.20), whilst applying the estimates (3.6) and (3.3) yields (4.21). □

5 Applications

5.1 Comparison of Variance-Gamma Distributions

The following proposition quantifies the error in approximating a general VG distribution by a SVG distribution. We refer the reader to [30] for a number of similar bounds for comparison of univariate distributions. The proof provides an example under which the bounds on $\|(x - \mu)f^{(k)}(x)\|$, $k = 0, 1, 2, 3$, for the solution of the SVG Stein equation that were given in Theorem 3.1 prove useful. This application also serves as a simple example in which the inequalities of Proposition 4.1 are suboptimal.

Proposition 5.1 *Let $X \sim VG(r_1, \theta_1, \sigma_1, \mu_1)$ and $Y \sim SVG(r_2, \sigma_2, \mu_2)$. Then*

$$
d_W(X, Y) \leq \frac{9}{2}\left(1 + \frac{1}{2(r_2 + 1)}\right)\frac{|\sigma_1^2 - \sigma_2^2|}{\sigma_2}
$$
$$
+ \frac{9}{2\sigma_2}\left(\frac{1}{r_2 + 1} + \frac{\pi\Gamma\left(\frac{r_2+1}{2}\right)}{2\Gamma\left(\frac{r_2}{2} + 1\right)}\right)\left(|\sigma_1^2 r_1 - \sigma_2^2 r_2| + 2|\theta_1(\mu_1 - \mu_2)|\right)
$$
$$
+ \left(\frac{7}{2} + \frac{9\sigma_1^2}{\sigma_2^2(r_2 + 1)}\right)|\mu_1 - \mu_2| + \left(\frac{7r_1}{2} + \frac{27}{2} + \frac{9}{2(r_2 + 1)}\right)|\theta_1|.
$$
$$(5.1)$$

Suppose now that $\mu_1 = \mu_2$. Then

$$
d_K(X, Y) \leq \frac{1}{2}\left(9 + \frac{1}{r_2}\right)\left|1 - \frac{\sigma_1^2}{\sigma_2^2}\right| + 2\left|1 - \frac{\sigma_1^2 r_1}{\sigma_2^2 r_2}\right| + \frac{|\theta_1|}{\sigma_2}\left(2 + \frac{r_1 + 1}{r_2} + \frac{\pi r_1 \Gamma\left(\frac{r_2}{2}\right)}{2\Gamma\left(\frac{r_2+1}{2}\right)}\right).
$$
$$(5.2)$$

Remark 5.2 The function $h(x) = x$ is in the class $\mathcal{H}_W$. Therefore

$$
d_W(X, Y) \geq |\mathbb{E}X - \mathbb{E}Y| = |r_1\theta_1 + \mu_1 - \mu_2|.
$$

When $\mu_1 = \mu_2$, this lower bound is equal to $r_1|\theta_1|$, and so there exist constants $c > 0$ and $C > 0$ independent of θ_1 such that $c|\theta_1| \leq d_W(X, Y) \leq C|\theta_1|$, if in addition $r_1 = r_2$ and $\sigma_1 = \sigma_2$. Comparing with the Kolmogorov bound (5.2), we see that the inequalities of Proposition 4.1 are suboptimal in this application.

Proof Let $\mathcal{A}_{r,\theta,\sigma,\mu}$ denote the differential operator on the left-hand side of the VG Stein equation (1.2). Suppose that $h : \mathbb{R} \to \mathbb{R}$ is either bounded or Lipschitz. Let f be the solution of the $SVG(r_2, \sigma_2, \mu_2)$ Stein equation. Then

$$
\mathbb{E}h(X) - \mathbb{E}h(Y) = \mathbb{E}[\mathcal{A}_{r_2,0,\sigma_2,\mu_2}f(X)]
$$
$$
= \mathbb{E}[\mathcal{A}_{r_2,0,\sigma_2,\mu_2}f(X) - \mathcal{A}_{r_1,\theta_1,\sigma_1,\mu_1}f(X)].
$$
$$(5.3)$$

That $\mathbb{E}[\mathcal{A}_{r_1,\theta_1,\sigma_1,\mu_1}f(X)] = 0$ follows from the assumptions on h, the estimates of Theorem 3.1, and Lemma 3.1 of [16]. Firstly, we prove (5.1). Suppose $h \in \mathcal{H}_W$. Then, from (5.3),

$$
|\mathbb{E}h(X) - \mathbb{E}h(Y)|
$$
$$
= \left|\mathbb{E}[\sigma_1^2(X - \mu_1)f''(X) + (\sigma_1^2 r_1 + 2\theta_1(X - \mu_1)f'(X) + (r_1\theta_1 - (X - \mu_1))f(X)\right.
$$
$$
\left. - \sigma_2^2(X - \mu_2)f''(X) - \sigma_2^2 r_2 f'(X) + (X - \mu_2)f(X)]\right|
$$
$$
= \left|\mathbb{E}[(\sigma_1^2 - \sigma_2^2)(X - \mu_2)f''(X) + \sigma_1^2(\mu_2 - \mu_1)f''(X) + (\sigma_1^2 r_1 - \sigma_2^2 r_2)f'(X)\right.
$$
$$
\left. + 2\theta_1(X - \mu_2)f'(X) + 2\theta_1(\mu_2 - \mu_1)f'(X) + r_1\theta_1 f(X) + (\mu_1 - \mu_2)f(X)]\right|
$$
$$
\leq |\sigma_1^2 - \sigma_2^2|\|(x - \mu_2)f''(x)\| + \sigma_1^2|\mu_1 - \mu_2|\|f''\|
$$

$$+ \left(|\sigma_1^2 r_1 - \sigma_2^2 r_2| + 2|\theta_1(\mu_1 - \mu_2)|\right)\|f'\|$$
$$+ 2|\theta_1|\|(x - \mu_2)f'(x)\| + (r_1|\theta_1| + |\mu_1 - \mu_2|)\|f\|. \tag{5.4}$$

Using the estimates of Theorem 3.1 (with $\|h'\| \leq 1$) to bound (5.4) yields (5.1).

Now suppose that $\mu_1 = \mu_2$. Take $h_z(x) = \mathbf{1}(x \leq z)$. On using the estimates of Corollary 3.3 to bound (5.4), we obtain (5.2), as required. □

5.2 Malliavin-Stein Method for Symmetric Variance-Gamma Approximation

In recent years, one of the most significant applications of Stein's method has been to Gaussian analysis on Wiener space. This body of research was initiated by [34], in which Stein's method and Malliavin calculus are combined to derive a quantitative "fourth moment" theorem for the normal approximation of a sequence of random variables living in a fixed Wiener chaos.

In a recent work [12], the Malliavin-Stein method was extended to the VG distribution. Here, we obtain explicit constants in some of the main results (in the SVG case) of [12], these being six moment theorems for the SVG approximation of double Wiener-Itô integrals. Our results also fix a technical issue in that the Wasserstein distance bounds stated in [12] had only been proven in the weaker bounded Wasserstein distance (at the time of [12] the bounds for the solution of the Stein equation in the literature [15,16] had a dependence on the test function h such that this was the best that could be achieved).

Let us first introduce some notation; see the book [36] for a more detailed discussion. Let $\mathbb{D}^{p,q}$ be the Banach space of all functions in $L^q(\gamma)$, where γ is the standard Gaussian measure, whose Malliavin derivatives up to order p also belong to $L^q(\gamma)$. Let $\mathbb{D}^\infty$ be the class of infinitely many times Malliavin differentiable random variables. We introduce the so-called Γ-operators Γ_j [35]. For a random variable $F \in \mathbb{D}^\infty$, we define $\Gamma_1(F) = F$ and, for every $j \geq 2$,

$$\Gamma_j(F) = \langle DF, -DL^{-1}\Gamma_{j-1}(F)\rangle_{\mathfrak{H}}.$$

Here D is the Malliavin derivative, L^{-1} is the pseudo-inverse of the infinitesimal generator of the Ornstein-Uhlenbeck semi-group, and $\mathfrak{H}$ is a real separable Hilbert space. Finally, for $f \in \mathfrak{H}^{\odot 2}$, we write $I_2(f)$ for the double Wiener-Itô integral of f.

Theorem 5.3 *Let $F \in \mathbb{D}^{2,4}$ be such that $\mathbb{E}F = 0$ and let $Z \sim \text{SVG}(r, \sigma, 0)$. Then*

$$d_{\mathrm{W}}(F, Z) \leq \frac{9}{\sigma^2(r+1)}\mathbb{E}|\sigma^2 F - \Gamma_3(F)|$$
$$+ \frac{9}{2\sigma}\left(\frac{1}{r+1} + \frac{\pi\Gamma\left(\frac{r+1}{2}\right)}{2\Gamma\left(\frac{r}{2}+1\right)}\right)|r\sigma^2 - \mathbb{E}[\Gamma_2(F)]|. \tag{5.5}$$

If in addition $F \in \mathbb{D}^{3,8}$, then $\Gamma_3(F)$ is square-integrable and

$$\mathbb{E}|\sigma^2 F - \Gamma_3(F)| \leq \left(\mathbb{E}[(\sigma^2 F - \Gamma_3(F))^2]\right)^{\frac{1}{2}}. \tag{5.6}$$

Proof Let $f : \mathbb{R} \to \mathbb{R}$ be twice differentiable with bounded first and second derivative. Then it was shown in the proof of Theorem 4.1 of [12] that

$$
\begin{aligned}
&\big|\mathbb{E}\big[\sigma^2 F f''(F) + \sigma^2 r f'(F) - F f(F)\big]\big| \\
&= \big|\mathbb{E}\big[f''(F)(\sigma^2 F - \Gamma_3(F)) + f'(F)(r\sigma^2 - \mathbb{E}[\Gamma_2(F)])\big]\big| \qquad (5.7) \\
&\leq \|f''\|\mathbb{E}|\sigma^2 F - \Gamma_3(F)| + \|f'\|\mathbb{E}|r\sigma^2 - \mathbb{E}[\Gamma_2(F)]|.
\end{aligned}
$$

If $h \in \mathcal{H}_W$, then the solution f of the SVG$(r, \sigma, 0)$ Stein equation is twice differentiable with bounded first and second derivatives. Using the estimates (3.10) and (3.9) of Theorem 3.1 to bound $\|f''\|$ and $\|f'\|$ then yields (5.5). Inequality (5.6) is justified in [12]. $\qquad\square$

Corollary 5.4 *Let $F_n = I_2(f_n)$ with $f_n \in \mathfrak{H}^{\odot 2}$, $n \geq 1$. Also, let $Z \sim$ SVG$(r, \sigma, 0)$ and assume that $\mathbb{E}[F_n^2] = r\sigma^2$. Then*

$$
d_W(F_n, Z) \leq \frac{9}{\sigma^2(r+1)} \left(\frac{1}{120} \kappa_6(F_n) - \frac{\sigma^2}{3} \kappa_4(F_n) + \frac{1}{4}(\kappa_3(F_n))^2 + \sigma^4 \kappa_2(F_n) \right)^{\frac{1}{2}}.
$$
$$(5.8)$$

Proof It is a standard result that $\mathbb{E}[\Gamma_2(F_n)] = \kappa_2(F_n)$ (see Lemma 4.2 and Theorem 4.3 of [35]), and it was shown in the proof of Theorem 5.8 of [12] that

$$
\mathbb{E}[(\sigma^2 F_n - \Gamma_3(F_n))^2] = \frac{1}{120} \kappa_6(F_n) - \frac{\sigma^2}{3} \kappa_4(F_n) + \frac{1}{4}(\kappa_3(F_n))^2 + \sigma^4 \kappa_2(F_n).
$$

Inserting these formulas into (5.5) yields (5.8), as required. $\qquad\square$

Remark 5.5 One can obtain Kolmogorov distance bounds by applying Proposition 4.1 to the bound (5.8). However, these bounds are unlikely to be of optimal order. Unlike for normal approximation, for which an optimal rate of convergence in Kolmogorov distance has been obtained [37], there is a technical difficulty for SVG approximation because the first derivative of the solution f_z of the SVG$(r, \sigma, 0)$ Stein equation with test function $h_z(x) = \mathbf{1}(x \leq z)$ has a discontinuity at the origin when $z = 0$ (see Proposition 3.5). We can, however, bound the expression using the inequalities (3.7) for $\|x f''(x)\|$ and (3.4) for $\|f'\|$ to obtain the bound

$$
\begin{aligned}
d_K(F, Z) &\leq \frac{1}{2\sigma^2}\left(9 + \frac{1}{r}\right)\mathbb{E}\left|\sigma^2 - \frac{\Gamma_3(F)}{F}\right| + \frac{2}{\sigma^2 r}|r\sigma^2 - \mathbb{E}[\Gamma_2(F)]| \\
&\leq \frac{1}{2\sigma^2}\left(9 + \frac{1}{r}\right)\left\{\mathbb{E}\left[\left(\sigma^2 - \frac{\Gamma_3(F)}{F}\right)^2\right]\right\}^{\frac{1}{2}} + \frac{2}{\sigma^2 r}|r\sigma^2 - \mathbb{E}[\Gamma_2(F)]|,
\end{aligned}
$$

provided the expectations exist. However, there are no formulas in the literature for the expectations $\mathbb{E}[\Gamma_3(F)/F]$ and $\mathbb{E}[(\Gamma_3(F))^2/F^2]$ (when they exist), and it is unlikely they could be expressed solely in terms of lower order cumulants of F.

5.3 Binary Sequence Comparison

Here we consider an application of Theorem 4.10 to binary sequence comparison. This a special case of a more general problem of word sequence comparison, which is of importance to biological sequence comparison. One way of comparing sequences uses k-tuples (a sequence of letters of length k). If two sequences are closely related, we would expect their k-tuple content to be similar. A statistic for sequence comparison based on k-tuple content, known as the D_2 statistic, was suggested by [4] (see [45] for further statistics based on k-tuple content). Letting $\mathcal{A}$ denote an alphabet of size d, and $X_{\mathbf{w}}$ and $Y_{\mathbf{w}}$ the number of occurrences of the word $\mathbf{w} \in \mathcal{A}^k$ in the first and second sequences, respectively, then the D_2 statistic is defined by

$$D_2 = \sum_{\mathbf{w} \in \mathcal{A}^k} X_{\mathbf{w}} Y_{\mathbf{w}}.$$

Due to the complicated dependence structure (for a detailed account see [46]), approximating the asymptotic distribution of D_2 is a difficult problem. However, for certain parameter regimes D_2 has been shown to be asymptotically normal and Poisson [31].

We now consider the case of an alphabet of size 2 with comparison based on the content of 1-tuple. We suppose that the sequences are of length m and n, the alphabet is $\{0, 1\}$, and $\mathbb{P}(0 \text{ appears}) = \mathbb{P}(1 \text{ appears}) = \frac{1}{2}$. Denoting the number of occurrences of 0 in the two sequences by X and Y, then

$$D_2 = XY + (m - X)(n - Y).$$

Clearly, X and Y are independent binomial variables with expectations $\frac{m}{2}$ and $\frac{n}{2}$. Straightforward calculations (see [31]) show that $\mathbb{E}D_2 = \frac{mn}{2}$ and $\text{Var}(D_2) = \frac{mn}{4}$ and the standardised D_2 statistic can be written as

$$W = \frac{D_2 - \mathbb{E}D_2}{\sqrt{\text{Var}(D_2)}} = \left(\frac{X - \frac{m}{2}}{\sqrt{\frac{m}{4}}} \right) \left(\frac{Y - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right). \tag{5.9}$$

By the central limit theorem, $(X - \frac{m}{2})/\sqrt{\frac{m}{4}}$ and $(Y - \frac{n}{2})/\sqrt{\frac{n}{4}}$ are approximately $N(0, 1)$ distributed, and so W has an approximate SVG$(1, 1, 0)$ distribution. In [16], a $O(m^{-1} + n^{-1})$ bound for the rate of convergence was given in a smooth test function metric (which requires the test function to be three times differentiable). In Theorem 5.8 below, we use Theorem 4.10 to obtain bounds in the more usual Wasserstein and Kolmogorov metrics. Our rate of convergence is slower, but we do quantify the approximation in stronger metrics. We will first need to prove the following theorem.

Theorem 5.6 *Suppose $X_1, \ldots, X_m$ are i.i.d. and $Y_1, \ldots, Y_n$ are i.i.d., with $\mathbb{E}X_1 = \mathbb{E}Y_1 = 0$, $\mathbb{E}X_1^2 = \mathbb{E}Y_1^2$ and $\mathbb{E}|X_1|^3 < \infty$ and $\mathbb{E}|Y_1|^3 < \infty$. Let $W_1 = \frac{1}{\sqrt{m}} \sum_{i=1}^m X_i$ and $W_2 = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$ and set $W = W_1 W_2$. Let $Z \sim \text{SVG}(1, 1, 0)$. Then*

$$d_W(W, Z) \leq 20.11 \left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right) \mathbb{E}|X_1|^3 \mathbb{E}|Y_1|^3. \tag{5.10}$$

If in addition $\mathbb{E}X_1^3 = \mathbb{E}Y_1^3 = 0$ *and* $\mathbb{E}X_1^4 < \infty$ *and* $\mathbb{E}Y_1^4 < \infty$, *then*

$$d_K(W, Z) \leq \left\{ 44.33 + 2.02 \left[\log \left(\frac{1}{\mathbb{E}X_1^4 \mathbb{E}Y_1^4} \right) + \log \left(\frac{mn}{m+n} \right) \right] \right\} \left(\frac{1}{m} + \frac{1}{n} \right)^{\frac{1}{3}}$$
$$\left(\mathbb{E}X_1^4 \mathbb{E}Y_1^4 \right)^{\frac{1}{3}}. \tag{5.11}$$

Remark 5.7 The rate of convergence in Kolmogorov distance bound (5.11) is unlikely to be of optimal order, but is better than the $O\left(m^{-\frac{1}{4}} \log(m) + n^{-\frac{1}{4}} \log(n)\right)$ rate that would result from simply applying Proposition 4.1 to (5.10). A reasonable conjecture is that the optimal rate is $O(m^{-\frac{1}{2}} + n^{-\frac{1}{2}})$.

Proof Since $Z \sim SVG(1, 1, 0)$, we will apply Theorem 4.10 with $r = 1$, for which $W^{V_1} = W^{*(2)}$, the W-zero bias transformation of order 2. We begin by collecting some useful properties of this distributional transformation. In [19], the following construction is given: $W^{*(2)} = \frac{1}{\sqrt{mn}} W_1^* W_2^*$. Since W_1 and W_2 are sums of independent random variables, we have by part (v) of Lemma 2.1 of [27] that $W_1^* = W_1 - \frac{X_I}{\sqrt{m}} + \frac{X_I^*}{\sqrt{m}}$ and $W_2^* = W_2 - \frac{Y_J}{\sqrt{n}} + \frac{Y_J^*}{\sqrt{n}}$, where I and J are chosen uniformly from $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$, respectively. It was shown in the proofs of Corollaries 4.1 and 4.2 of [19] that

$$\mathbb{E}|W - W^{*(2)}| \leq \frac{13}{8} \left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right) \mathbb{E}|X_1|^3 \mathbb{E}|Y_1|^3 \tag{5.12}$$

and, if $\mathbb{E}X_1^3 = \mathbb{E}Y_1^3 = 0$,

$$\mathbb{E}[(W - W^{*(2)})^2] \leq \frac{20}{3} \left(\frac{1}{m} + \frac{1}{n} \right) \mathbb{E}X_1^4 \mathbb{E}Y_1^4. \tag{5.13}$$

The assumption $\mathbb{E}X_1^3 = \mathbb{E}Y_1^3 = 0$ implies $\mathbb{E}X_1^* = \mathbb{E}Y_1^* = 0$ ([27], part (iv) of Lemma 2.1), which allowed [19] to obtain the $O\left(m^{-1} + n^{-1}\right)$ rate in (5.13).

The bound (5.10) is immediate from (5.12) and (4.19):

$$d_W(W, Z) \leq \frac{99}{8} \mathbb{E}|W - W^{*(2)}| \leq \frac{1287}{64} \left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right) \mathbb{E}|X_1|^3 \mathbb{E}|Y_1|^3,$$

and $\frac{1287}{64} = 20.11$.

Now we prove (5.11). We begin by setting $r = \sigma = 1$ in (4.17), using that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, and applying Markov's inequality to obtain

$$d_K(W, Z)$$

$$\leq \left\{5 + \pi^{3/2} + \frac{10}{\pi}\left[1 + \log\left(\frac{1}{2}\right)\right]\right\}\beta + \frac{10}{\pi}\log\left(\frac{1}{\beta}\right) + 12\mathbb{P}(|W - W^{*(2)}| > \beta)$$

$$\leq \left\{11.55 + 3.19\log\left(\frac{1}{\beta}\right)\right\}\beta + 12\frac{\mathbb{E}[(W - W^{*(2)})^2]}{\beta^2}.$$

Setting $\beta = \left(\mathbb{E}[(W - W^{*(2)})^2]\right)^{\frac{1}{3}}$ gives

$$d_K(W, Z) \leq \left\{23.55 + 1.07\log\left(\frac{1}{\mathbb{E}[(W - W^{*(2)})^2]}\right)\right\}\left(\mathbb{E}[(W - W^{*(2)})^2]\right)^{\frac{1}{3}}.$$

$$(5.14)$$

Substituting (5.13) into (5.14) and simplifying then yields (5.11). □

Theorem 5.8 *Let W be the standardised D_2 statistic, as defined in (5.9), based on 1-tuple content, for uniform i.i.d. binary sequences of lengths m and n. Let $Z \sim$ SVG(1, 1, 0). Then*

$$d_W(W, Z) \leq 20.11\left(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}}\right),$$

$$d_K(W, Z) \leq \left\{44.33 + 2.02\log\left(\frac{mn}{m+n}\right)\right\}\left(\frac{1}{m} + \frac{1}{n}\right)^{\frac{1}{3}}.$$

Proof Let $\mathbb{I}_i$ and $\mathbb{J}_i$ be the indicator random variables that letter 0 occurs at position i in the first and second sequences, respectively. Then $X = \sum_{i=1}^m \mathbb{I}_i$ and $Y = \sum_{j=1}^n \mathbb{J}_j$. We may then write

$$W = \left(\frac{X - \frac{m}{2}}{\sqrt{\frac{m}{4}}}\right)\left(\frac{Y - \frac{n}{2}}{\sqrt{\frac{n}{4}}}\right) = \left(\frac{1}{\sqrt{m}}\sum_{i=1}^m X_i\right)\left(\frac{1}{\sqrt{n}}\sum_{j=1}^n Y_j\right),$$

where $X_i = 2(\mathbb{I}_i - \frac{1}{2})$ and $Y_j = 2(\mathbb{J}_j - \frac{1}{2})$. The X_i and Y_j are all independent with zero mean and unit variance. Also, $\mathbb{E}X_1^3 = \mathbb{E}Y_1^3 = 0$, $\mathbb{E}|X_1|^3 = \mathbb{E}|Y_1|^3 = 1$ and $\mathbb{E}X_1^4 = \mathbb{E}Y_1^4 = 1$, and the result now follows from Theorem 5.6. □

5.4 Random Sums

Let $X_1, X_2, \ldots$ be i.i.d., positive, non-degenerate random variables with unit mean. Let N_p be a Geo(p) random variable with $\mathbb{P}(N_p = k) = p(1 - p)^{k-1}$, $k \geq 1$, that is independent of the X_i. Then, a well-known result of [47] states that $p\sum_{i=1}^{N_p} X_i$

converges in distribution to an exponential distribution with parameter 1 as $p \to 0$. Geometric summation does indeed arise in a variety of settings; see [28]. Stein's method was used by [39] to obtain a quantitative generalisation of the result of [47]. If we alter the assumptions so that the X_i have mean zero and finite nonzero variance, then $p^{\frac{1}{2}} \sum_{i=1}^{N_p} X_i$ converges to a Laplace distribution as $p \to 0$; see [52] and [43]. Recently, [43], through the use of the centred equilibrium transformation, mirrored the approach of [39] to obtain an explicit error bound in the bounded Wasserstein metric.

In this section, we use Theorem 4.10 to obtain Wasserstein and Kolmogorov error bounds for the theorems of [43]. Indeed, Theorems 5.9 and 5.10 below give Wasserstein and Kolmogorov distance bounds for the approximations of Theorems 1.3 and 4.4 of [43], respectively. The results of [39] are also given in these metrics, and we follow their approach to obtain our Kolmogorov bounds. For a random variable X, we denote by distribution function by F_X and its generalised inverse by F_X^{-1}.

Theorem 5.9 *Let N be a positive, integer-valued random variable with $\mu = \mathbb{E}N < \infty$ and let $X_1, X_2, \ldots$ be a sequence of independent random variables, independent of N, with $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i^2 = \sigma_i^2 \in (0, \infty)$. Set $\sigma^2 = \frac{1}{\mu} \mathbb{E}\left[\left(\sum_{i=1}^N X_i\right)^2\right] = \frac{1}{\mu} \mathbb{E}\left[\sum_{i=1}^N \sigma_i^2\right]$. Also, let M be any positive, integer-valued random variable, independent of the X_i, satisfying*

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

Let $Z \sim \text{Laplace}(0, \frac{\sigma}{\sqrt{2}})$. Then, with $W = \frac{1}{\sqrt{\mu}} \sum_{i=1}^N X_i$, we have

$$d_{\mathrm{W}}(W, Z) \leq 12 \mu^{-\frac{1}{2}} \left\{ \mathbb{E}|X_M - X_M^L| + \sup_{i \geq 1} \sigma_i \mathbb{E}\left[|N - M|^{\frac{1}{2}}\right] \right\}. \tag{5.15}$$

Suppose further that $|X_i| \leq C$ for all i and $|N - M| \leq K$. Then

$$d_{\mathrm{K}}(W, Z) \leq \frac{17.04}{\sigma \sqrt{\mu}} \left\{ \sup_{i \geq 1} \|F_{X_i}^{-1} - F_{X_i^L}^{-1}\| + CK \right\}; \tag{5.16}$$

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

Proof Since $Z \sim \text{Laplace}(0, \frac{\sigma}{\sqrt{2}}) =_d \text{SVG}(2, \frac{\sigma}{\sqrt{2}}, 0)$, we will apply Theorem 4.10 with $r = 2$, for which $W^{V_1} = W^L$, the W-centred equilibrium distribution. For W as defined in the statement of the theorem, it was shown in the proof of Theorem 4.4 of [43] that $W^L = \mu^{-\frac{1}{2}}\left(\sum_{i=1}^{M-1} + X_M^L\right)$. Then

$$W^L - W = \mu^{-\frac{1}{2}} \left\{ (X_M^L - X_M) + \mathrm{sgn}(M - N) \sum_{i=(M \wedge N)+1}^{N \vee M} X_i \right\}.$$

Plugging this into (4.19) (with $r = 2$) and bounding $\mathbb{E}\left|\sum_{i=(M\wedge N)+1}^{N\vee M} X_i\right| \leq \sup_{i\geq 1}\sigma_i\mathbb{E}\left[|N-M|^{\frac{1}{2}}\right]$ (see the proof of Theorem 4.4 of [43]) yields (5.15). Now, using (4.17) and the formulas $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ and $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$ gives that

$$
d_K(W, Z) \leq \left(\frac{7}{2} + 2\sqrt{\pi}\right)\frac{\beta\sqrt{2}}{\sigma} + \frac{5}{2}\cdot\frac{2\sqrt{2}\beta}{\sigma} + 11\mathbb{P}(|W-W^L| > \beta)
$$
$$
= 17.04\frac{\beta}{\sigma} + 11\mathbb{P}(|W-W^L| > \beta). \tag{5.17}
$$

Letting $\beta = \mu^{-\frac{1}{2}}\left\{\sup_{i\geq 1}\|F_{X_i}^{-1} - F_{X_i^L}^{-1}\| + CK\right\}$, and using Strassen's theorem we obtain (5.16) from (5.17), and the remark after (5.16) follows similarly. $\quad\square$

Theorem 5.10 *Let $X_1, X_2, \ldots$ be a sequence of independent random variables with $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = \sigma^2$, and let $N \sim \mathrm{Geo}(p)$ be independent of the X_i. Let $W = p^{\frac{1}{2}}\sum_{i=1}^{N} X_i$ and let $Z \sim \mathrm{Laplace}(0, \frac{\sigma}{\sqrt{2}})$. Then*

$$
d_K(W, Z) \leq 17.04\frac{p^{\frac{1}{2}}}{\sigma}\sup_{i\geq 1}\|F_{X_i}^{-1} - F_{X_i^L}^{-1}\|. \tag{5.18}
$$

If in addition $\rho = \sup_{i\geq 1}\mathbb{E}|X_i|^3 < \infty$, then

$$
d_W(W, Z) \leq 12p^{\frac{1}{2}}\left(\sigma + \frac{\rho}{3\sigma^2}\right). \tag{5.19}
$$

The $O(p^{\frac{1}{2}})$ rate in (5.19) is optimal.

Remark 5.11 (i) Theorem 1.3 of [43] gives the bound

$$
d_{BW}(W, Z) \leq p^{\frac{1}{2}}(2\sqrt{2} + \sigma)\left(\sigma + \frac{\rho}{3\sigma^2}\right). \tag{5.20}
$$

which holds under the same conditions as (5.19). Aside from being given in a stronger metric, the bound (5.19) has a theoretical advantage of having a multiplicative constant, 12, which is independent of σ, whereas (5.20) has a multiplicative constant $2\sqrt{2} + \sigma$. The bound (5.20) has a smaller constant than (5.19) when $\sigma < 12 - 2\sqrt{2}$, whilst the constant is larger when $\sigma > 12 - 2\sqrt{2}$.

(ii) The argument used to prove the final assertion of Theorem 5.10 also shows that the $O(p^{\frac{1}{2}})$ rate in (5.20) is optimal.

(iii) Suppose now that $\tau = \sup_{i\geq 1}\mathbb{E}X_i^4 < \infty$. Then arguing as we did in the proof of Theorem 5.6 would result in the alternative bound

$$
d_K(W, Z) \leq Cp^{\frac{1}{3}}(1 + \tau), \tag{5.21}
$$

where $C > 0$ does not depend on p. Thus, the dependence on p is worse than in (5.18), but (5.21) may be preferable if $\sup_{i \geq 1} \|F_{X_i}^{-1} - F_{X_i^L}^{-1}\|$ is difficult to compute or large. The same remark applies to Theorem 5.9.

The quantity $\sup_{i \geq 1} \|F_{X_i}^{-1} - F_{X_i^L}^{-1}\|$ can be easily bounded if the X_i have finite support. To see this, suppose that $X_1, X_2, \ldots$ are supported on a subset of the finite interval $[a, b] \subset \mathbb{R}$. Theorem 3.2 of [43] (see also Proposition 4.5) gives that $X^L =_d B_2 X^*$, where $B_2 \sim \text{Beta}(2, 1)$ and X^*, the X-zero bias distribution, are independent. But part (ii) of Lemma 2.1 of [27] tells us that the support of X^* is the closed convex hull of the support of X, and since B_r is supported on $[0, 1]$ it follows that X^L is supported on $[a, b]$. We therefore have the bound $\sup_{i \geq 1} \|F_{X_i}^{-1} - F_{X_i^L}^{-1}\| \leq b - a$.

Proof As noted by [43], the assumptions on N and the X_i imply that $\mathcal{L}(M) = \mathcal{L}(N)$, so we can take $M = N$. Inequality (5.18) is now immediate from (5.16). To obtain (5.19), we note the inequality (see [43])

$$\mathbb{E}|X_N - X_N^L| \leq \mathbb{E}|X_1| + \sup_{i \geq 1} \mathbb{E}|X_i^L| = \mathbb{E}|X_1| + \sup_{i \geq 1} \frac{\mathbb{E}|X_i|^3}{3\sigma^2} \leq \sigma + \frac{\rho}{3\sigma^2},$$

where we used the Cauchy-Schwarz inequality. Inequality (5.19) now follows from (5.15).

Finally, we prove that the $O(p^{\frac{1}{2}})$ rate in (5.19) is optimal. Suppose, in addition to the assumptions in the statement of the theorem, that $X_1, X_2, \ldots$ are i.i.d. with moments of all order and $\mathbb{E}X_1^3 \neq 0$. Consider the test function $h(x) = \sin(tx)$, $|t| \leq 1$, which is in the class $\mathcal{H}_W$. We have $\mathbb{E}\sin(tZ) = 0$. We now consider the characteristic function $\varphi_W(t) = \mathbb{E}[e^{itW}]$, and note the relation $\mathbb{E}\sin(tW) = \text{Im}[\varphi_W(t)]$. From the above, we have that $d_W(W, Z) \geq |\text{Im}[\varphi_W(t)]|$. Recall that the probability generating function of $N \sim \text{Geo}(p)$ is given by $G_N(s) = \frac{ps}{1-(1-p)s}$, $s < -\log(1 - p)$. Then

$$\varphi_W(t) = G_N\big(\varphi_{X_1}(p^{\frac{1}{2}}t)\big) = \frac{p\varphi_{X_1}(p^{\frac{1}{2}}t)}{1 - (1 - p)\varphi_{X_1}(p^{\frac{1}{2}}t)}. \tag{5.22}$$

Now, since $\mathbb{E}X_1 = 0$ and $\mathbb{E}X_1^2 = \sigma^2$, as $p \to 0$,

$$\varphi_{X_1}(p^{\frac{1}{2}}t) = 1 - \frac{1}{2}pt^2\sigma^2 - \frac{1}{6}ip^{\frac{3}{2}}t^3\mathbb{E}X_1^3 + O(p^2). \tag{5.23}$$

Plugging (5.23) into (5.22) and performing a simple asymptotic analysis using the formula $\frac{1}{1+z} = 1 - z + O(|z|^2)$, $|z| \to 0$, gives that $\text{Im}[\varphi_W(t)] = -\frac{\frac{1}{6}p^{1/2}t^3\mathbb{E}X_1^3}{1+\sigma^2t^2/2} + O(p)$, and so the $O(p^{\frac{1}{2}})$ rate cannot be improved. $\qquad\square$

6 Further Proofs

Proof of Proposition 3.5. As usual, we set $\sigma = 1$ and $\mu = 0$. The solution of the SVG$(r, 1, 0)$ Stein equation with test function $h_z(x) = \mathbf{1}(x \leq z)$ is then

$$
f_z(x) = -\frac{K_\nu(|x|)}{|x|^\nu} \int_0^x |t|^\nu I_\nu(|t|)[\mathbf{1}(t \leq z) - \mathbb{P}(Z \leq z)] \, dt
$$
$$
- \frac{I_\nu(|x|)}{|x|^\nu} \int_x^\infty |t|^\nu K_\nu(|t|)[\mathbf{1}(t \leq z) - \mathbb{P}(Z \leq z)] \, dt. \qquad (6.1)
$$

Setting $z = 0$ and differentiating (6.1) using (A.9) and (A.10) gives that

$$
f_0'(x) = \frac{K_{\nu+1}(|x|)}{|x|^\nu} \text{sgn}(x) \int_0^x |t|^\nu I_\nu(|t|)[\mathbf{1}(t \leq 0) - \tfrac{1}{2}] \, dt
$$
$$
- \frac{I_{\nu+1}(|x|)}{|x|^\nu} \text{sgn}(x) \int_x^\infty |t|^\nu K_\nu(|t|)[\mathbf{1}(t \leq 0) - \tfrac{1}{2}] \, dt.
$$

We now note that, for all $\nu > -\frac{1}{2}$,

$$
\lim_{x \to 0} \left[\frac{I_{\nu+1}(|x|)}{|x|^\nu} \int_x^\infty |t|^\nu K_\nu(|t|)[\mathbf{1}(t \leq 0) - \tfrac{1}{2}] \, dt \right] = 0,
$$

due to the asymptotic formula (A.3) and the fact that $|t|^\nu K_\nu(|t|)$ is a constant multiple of the SVG$(r, 1, 0)$ density meaning that the integral is bounded for all $x \in \mathbb{R}$. Then

$$
f_0'(0+) = -\lim_{x \downarrow 0} \left[\frac{K_{\nu+1}(x)}{2x^\nu} \int_0^x t^\nu I_\nu(t) \, dt \right],
$$
$$
f_0'(0-) = -\lim_{x \uparrow 0} \left[\frac{K_{\nu+1}(-x)}{2(-x)^\nu} \int_0^x (-t)^\nu I_\nu(-t) \, dt \right]
$$
$$
= \lim_{x \uparrow 0} \left[\frac{K_{\nu+1}(-x)}{2(-x)^\nu} \int_0^{-x} u^\nu I_\nu(u) \, du \right].
$$

On using the asymptotic formulas (A.3) and (A.4), we obtain $f_0'(0+) = -\frac{1}{2(2\nu+1)}$ and $f_0'(0-) = \frac{1}{2(2\nu+1)}$, which proves the assertion. $\square$

Proof of Proposition 3.6. As usual, we set $\sigma = 1$ and $\mu = 0$. Consider the test function $h(x) = \frac{\sin(ax)}{a}$, which is in the class $\mathcal{H}_W$. Therefore, if there was a general bound of the form $\|f^{(3)}\| \leq M_r \|h'\|$, then we would be able to find a constant $N_r > 0$, independent of a, such that $\|f^{(3)}\| \leq N_r$. We shall show that $f^{(3)}(x)$ blows up as $x \to 0$ for a such that $ax \ll 1 \ll a^2x$, meaning that such a bound cannot be obtained for $\|f^{(3)}\|$ which proves the proposition. Before performing this analysis, we note that the second derivative $h''(x) = -a\sin(ax)$ blows up if $ax \ll 1 \ll a^2x$ (consider the expansion $\sin(t) = t + O(t^3)$, $t \to 0$). A bound of the form

$\|f^{(3)}\| \leq M_{r,0}\|\tilde{h}\| + M_{r,1}\|h'\| + M_{r,2}\|h''\|$ is therefore still possible, and we know from Sect. 3.1.7 of [9] that this is indeed the case.

Let $x > 0$. We first obtain a formula for $f^{(3)}(x)$. To this end, we note that twice differentiating the representation (3.1) of the solution and then simplifying using the differentiation formulas (A.9) and (A.10) followed by the Wronskian formula $I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x) = \frac{1}{x}$ [38] gives that

$$f''(x) = \frac{\tilde{h}(x)}{x} - \left[\frac{d^2}{dx^2}\left(\frac{K_\nu(x)}{x^\nu}\right)\right]\int_0^x t^\nu I_\nu(t)\tilde{h}(t)\,dt - \left[\frac{d^2}{dx^2}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]$$
$$\times \int_x^\infty t^\nu K_\nu(t)\tilde{h}(t)\,dt.$$

Differentiating this formulas then gives

$$f^{(3)}(x) = \frac{h'(x)}{x} - \frac{\tilde{h}(x)}{x^2} - \left[\frac{d^3}{dx^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right]\int_0^x t^\nu I_\nu(t)\tilde{h}(t)\,dt + R_1$$
$$+ \tilde{h}(x)\left\{-x^\nu I_\nu(x)\frac{d^2}{dx^2}\left(\frac{K_\nu(x)}{x^\nu}\right) + x^\nu K_\nu(x)\frac{d^2}{dx^2}\left(\frac{I_\nu(x)}{x^\nu}\right)\right\}$$
$$= \frac{h'(x)}{x} - \frac{(2\nu+2)\tilde{h}(x)}{x^2} - \left[\frac{d^3}{dx^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right]\int_0^x t^\nu I_\nu(t)\tilde{h}(t)\,dt + R_1,$$

$$(6.2)$$

where

$$R_1 = -\left[\frac{d^3}{dx^3}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(t)\tilde{h}(t)\,dt.$$

Here, to obtain equality (6.2) we used differentiation formulas (A.9) and (A.10) followed again by the Wronskian formula. For all $\nu > -\frac{1}{2}$ and $x > 0$, we can use inequalities (A.14) and (A.19) to bound R_1:

$$|R_1| \leq \|\tilde{h}\|\left[\frac{d^3}{dx^3}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(t)\,dt \leq \|\tilde{h}\|\frac{I_\nu(x)}{x^\nu}$$
$$\times \int_x^\infty t^\nu K_\nu(t)\,dt \leq \|\tilde{h}\|\frac{\sqrt{\pi}\Gamma(\nu+\frac{1}{2})}{2\Gamma(\nu+1)}.$$

As $\|\tilde{h}\| \leq 2\|h\| = \frac{2}{a}$, the term R_1 does not explode when $a \to \infty$.

Applying integration by parts to (6.2) we obtain

$$f^{(3)}(x) = \frac{h'(x)}{x} + \left[\frac{d^3}{dx^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right]\int_0^x h'(u)\int_0^u t^\nu I_\nu(t)\,dt\,du + R_1 + R_2,$$

where

$$R_2 = -\tilde{h}(x)\left\{\frac{2v+2}{x^2} + \left[\frac{d^3}{dx^3}\left(\frac{K_v(x)}{x^v}\right)\right]\int_0^x t^v I_v(t)\, dt\right\} =: -\tilde{h}(x)A_v(x). \quad (6.3)$$

For all $v > -\frac{1}{2}$, we show that there exists a constant $C_v > 0$ independent of x such that $A_v(x) \le C_v$ for all $x > 0$. To see this, it suffices to consider the behaviour in the limits $x \downarrow 0$ and $x \to \infty$. We first note that $A_v(x) \to 0$ as $x \to \infty$, which follows from using the differentiation formula (A.13) followed by (A.6) and the following limiting form (see [22]):

$$\int_0^x t^v I_v(t)\, dt \sim \frac{1}{\sqrt{2\pi}} x^{v-1/2} e^x, \quad x \to \infty,\ v > -\tfrac{1}{2}.$$

Also, using the differentiation formula (A.13) followed by the limiting forms (A.3) and (A.4) gives that, for $v > -\frac{1}{2}$, as $x \downarrow 0$,

$$\left[\frac{d^3}{dx^3}\left(\frac{K_v(x)}{x^v}\right)\right]\int_0^x t^v I_v(t)\, dt$$
$$= -\left(\frac{(2v+1)K_v(x)}{x^v} + \left(1 + \frac{(2v+1)(2v+2)}{x^2}\right)\frac{K_{v+1}(x)}{x^v}\right)\int_0^x t^v I_v(t)\, dt$$
$$= -\left(\frac{(2v+1)(2v+2)\cdot 2^v\Gamma(v+1)}{x^{2v+3}} + O(x^{-2v-1})\right)\int_0^x \frac{t^{2v}}{2^v\Gamma(v+1)}\, dt$$
$$= -\frac{2v+2}{x^2} + O(1),$$

and therefore $A_v(x)$ is bounded as $x \downarrow 0$, as required. We conclude that R_2 does not explode when $a \to \infty$.

Now, we use the differentiation formula (A.13) to obtain

$$f^{(3)}(x) = \frac{h'(x)}{x} - \frac{(2v+1)(2v+2)K_{v+1}(x)}{x^{v+2}}\int_0^x h'(u)\int_0^u t^v I_v(t)\, dt\, du$$
$$+ R_1 + R_2 + R_3,$$

where

$$|R_3| = \left|\left(\frac{(2v+1)K_v(x)}{x^v} + \frac{K_{v+1}(x)}{x^v}\right)\int_0^x h'(u)\int_0^u t^v I_v(t)\, dt\, du\right|$$
$$\le \left(\frac{(2v+1)K_v(x)}{x^v} + \frac{K_{v+1}(x)}{x^v}\right)\cdot\frac{2(v+2)}{2v+1}x^v I_{v+2}(x), \quad (6.4)$$

where we used (A.15) and that $\|h'\| = 1$ to obtain the second inequality. For $v > -\frac{1}{2}$, the expression involving modified Bessel functions in (6.4) is uniformly bounded for all $x \ge 0$, which can be seen from a straightforward analysis involving the asymptotic formulas (A.3) – (A.6). Therefore, the term R_3 does not explode when $a \to \infty$.

We now analyse the behaviour of $f^{(3)}(x)$ in a neighbourhood of $x = 0$ when $a \to \infty$. For all $x \geq 0$, the terms R_1, R_2 and R_3 are $O(1)$ as $a \to \infty$. Therefore using the asymptotic formulas (A.3) and (A.4), we obtain

$$
\begin{aligned}
f^{(3)}(x) &= -\frac{\cos(ax)}{x} + \frac{(2\nu + 1)(2\nu + 2)}{x^{\nu+2}} \cdot \frac{2^\nu \Gamma(\nu + 1)}{x^{\nu+1}} \\
&\quad \times \int_0^x \cos(au) \int_0^u \frac{t^{2\nu}}{2^\nu \Gamma(\nu + 1)} \, dt \, du + O(1) \\
&= -\frac{\cos(ax)}{x} + \frac{(2\nu + 1)(2\nu + 2)}{x^{2\nu+3}} \int_0^x u^{2\nu+1} \cos(au) \, du + O(1), \quad x \downarrow 0.
\end{aligned}
$$

In addition to $x \downarrow 0$ and $a \to \infty$ we let $ax \downarrow 0$. Therefore on using that $\cos(t) = 1 - \frac{1}{2}t^2 + O(t^4)$ as $t \downarrow 0$, we have that, in this regime,

$$
\begin{aligned}
f^{(3)}(x) &= -\frac{1}{x}\left(1 - \frac{a^2 x^2}{2}\right) + \frac{2\nu + 2}{x^{2\nu+3}} \int_0^x u^{2\nu+1}\left(1 - \frac{a^2 u^2}{2}\right) du + O(1) \\
&= \frac{a^2 x}{2} - \frac{(\nu + 1)a^2 x}{2\nu + 4} + O(1) = \frac{a^2 x}{2(\nu + 2)} + O(1).
\end{aligned}
$$

If we take choose a such that $ax \ll 1 \ll a^2 x$, then $f^{(3)}(x)$ blows up in a neighbourhood of the origin, which proves the assertion. $\qquad\square$

Proof of (3.18). As usual, we set $\sigma = 1$. From the formula (3.1) for the solution of the SVG$(r, 1, 0)$ Stein equation we have

$$
\begin{aligned}
\lim_{x \to \infty} x f(x) &= -\lim_{x \to \infty} \left\{ \frac{K_\nu(x)}{x^\nu} \int_0^x t^\nu I_\nu(t) \tilde{h}(t) \, dt - \frac{I_\nu(x)}{x^\nu} \int_x^\infty t^\nu K_\nu(t) \tilde{h}(t) \, dt \right\} \\
&=: I_1 + I_2.
\end{aligned}
$$

We shall use L'Hôpital's rule to calculate I_1 and I_2. In anticipation of this we note that

$$
\frac{d}{dx}\left(\frac{x^{\nu-1}}{K_\nu(x)}\right) = \frac{d}{dx}\left(\frac{1}{x} \Big/ \frac{K_\nu(x)}{x^\nu}\right) = -\frac{x^{\nu-2}}{K_\nu(x)} + \frac{x^{\nu-1} K_{\nu+1}(x)}{K_\nu(x)^2},
$$

where we used the quotient rule and (A.10) in the final step. Similarly, on using (A.9) we obtain

$$
\frac{d}{dx}\left(\frac{x^{\nu-1}}{I_\nu(x)}\right) = -\frac{x^{\nu-2}}{I_\nu(x)} - \frac{x^{\nu-1} I_{\nu+1}(x)}{I_\nu(x)^2}.
$$

Therefore, by L'Hôpital's rule,

$$I_1 = -\lim_{x\to\infty} \left\{ \frac{x^\nu I_\nu(x)\tilde{h}(x)}{-\dfrac{x^{\nu-2}}{K_\nu(x)} + \dfrac{x^{\nu-1}K_{\nu+1}(x)}{K_\nu(x)^2}} \right\} = -\frac{1}{2}\tilde{h}(\infty),$$

$$I_2 = -\lim_{x\to\infty} \left\{ \frac{-x^\nu K_\nu(x)\tilde{h}(x)}{-\dfrac{x^{\nu-2}}{I_\nu(x)} - \dfrac{x^{\nu-1}I_{\nu+1}(x)}{I_\nu(x)^2}} \right\} = -\frac{1}{2}\tilde{h}(\infty),$$

where we used the asymptotic formulas (A.5) and (A.6) to compute the limits. Thus, $\lim_{x\to\infty} xf(x) = -\tilde{h}(\infty)$. Similarly, by considering (3.2) instead of (3.1), we obtain $\lim_{x\to-\infty} xf(x) = \tilde{h}(-\infty)$. $\qquad\square$

The following lemma will be used in the proof of Proposition 4.1.

Lemma 6.1 *(i) Let $\nu > 0$. Then $x^\nu K_\nu(x) \le 2^{\nu-1}\Gamma(\nu)$ for all $x > 0$.*
(ii) Suppose $0 < x < 0.729$. Then $K_0(x) < -2\log(x)$.

Proof (i) We have that $\frac{\mathrm{d}}{\mathrm{d}x}\big(x^\nu K_\nu(x)\big) = -x^\nu K_{\nu-1}(x) < 0$ (see (A.8)), which implies that $x^\nu K_\nu(x)$ is a decreasing function of x. From (A.4) we have $\lim_{x\downarrow 0} x^\nu K_\nu(x) = 2^{\nu-1}\Gamma(\nu)$, and we thus deduce the inequality.

(ii) From the differentiation formula (A.7), for all $x > 0$, $\frac{\mathrm{d}}{\mathrm{d}x}\big(-2\log(x) - K_0(x)\big) = -\frac{2}{x} + K_1(x) < 0$, where the inequality follows from part (i). Therefore $-2\log(x) - K_0(x)$ is a decreasing function of x. But one can check numerically using Mathematica that $-2\log(0.729) - K_0(0.729) = 0.00121$, and the conclusion follows. $\qquad\square$

Proof of Proposition 4.1. For ease of notation, we shall set $\mu = 0$; the extension to general $\mu \in \mathbb{R}$ is obvious. Throughout this proof, Z will denote a SVG$(r, \sigma, 0)$ random variable.

(i) Let $r > 1$. Proposition 1.2 of [48] states that if a random variable Y has Lebesgue density bounded by C, then for any random variable W,

$$d_{\mathrm{K}}(W, Y) \le \sqrt{2C d_{\mathrm{W}}(W, Y)}. \tag{6.5}$$

Since the SVG$(r, \sigma, 0)$ distribution is unimodal about 0, it follows from (2.3) that the density is bounded above by $C = \frac{1}{2\sigma\sqrt{\pi}}\Gamma(\frac{r-1}{2})/\Gamma(\frac{r}{2})$, which on substituting into (6.5) yields the desired bound.

(ii) Here we consider the case $r = 1$. We begin by following the approach used in the proof of Proposition 1.2 of [48], but we need to alter the argument because the SVG$(1, \sigma, 0)$ density $p(x) = \frac{1}{\pi\sigma}K_0\big(\frac{|x|}{\sigma}\big)$ is unbounded as $x \to 0$. Consider the functions $h_z(x) = \mathbf{1}(x \le z)$, and the 'smoothed' $h_{z,\alpha}(x)$ defined to be one for $x \le z + 2\alpha$, zero for $x > z$, and linear between. Then

$$\mathbb{P}(W \leq z) - \mathbb{P}(Z \leq z) = \mathbb{E}h_z(W) - \mathbb{E}h_{z,\alpha}(Z) + \mathbb{E}h_{z,\alpha}(Z) - \mathbb{E}h_z(Z)$$

$$\leq \mathbb{E}h_{z,\alpha}(W) - \mathbb{E}h_{z,\alpha}(Z) + \frac{1}{2}\mathbb{P}(z \leq Z \leq z + 2\alpha)$$

$$\leq \frac{1}{2\alpha}d_W(W, Z) + \frac{1}{2}\mathbb{P}(z \leq Z \leq z + 2\alpha)$$

$$\leq \frac{1}{2\alpha}d_W(W, Z) + \mathbb{P}(0 \leq Z \leq \alpha), \tag{6.6}$$

where the last inequality follows because the SVG$(1, \sigma, 0)$ density is a decreasing function of x for $x > 0$ and an increasing function for $x < 0$, and so $\mathbb{P}(z \leq Z \leq z+2\alpha)$ is maximised for $z = -\alpha$. Suppose that $\frac{\alpha}{\sigma} < 0.729$. Then we can use Lemma 6.1 to obtain

$$\mathbb{P}(0 \leq Z < \alpha) = \int_0^\alpha \frac{1}{\pi\sigma}K_0\left(\frac{t}{\sigma}\right)dt = \frac{1}{\pi}\int_0^{\frac{\alpha}{\sigma}}K_0(y)\,dy$$

$$\leq \frac{1}{\pi}\int_0^{\frac{\alpha}{\sigma}} -2\log(y)\,dy = \frac{2\alpha}{\pi\sigma}\left[1 + \log\left(\frac{\sigma}{\alpha}\right)\right]. \tag{6.7}$$

Substituting into (6.6) gives that, for any $z \in \mathbb{R}$ and $\alpha > 0$,

$$\mathbb{P}(W \leq z) - \mathbb{P}(Z \leq z) \leq \frac{1}{2\alpha}d_W(W, Z) + 1\frac{2\alpha}{\pi\sigma}\left[1 + \log\left(\frac{\sigma}{\alpha}\right)\right].$$

We take $\alpha = \frac{1}{2}\sqrt{\pi\sigma d_W(W, Z)}$, which, as we assumed that $\sigma^{-1}d_W(W, Z) < 0.676$, ensures that $\frac{\alpha}{\sigma} < 0.729$. This leads to the upper bound

$$\mathbb{P}(W \leq z) - \mathbb{P}(Z \leq z) \leq \left\{2 + \log\left(\frac{2}{\sqrt{\pi}}\right) + \frac{1}{2}\log\left(\frac{\sigma}{d_W(W, Z)}\right)\right\}\sqrt{\frac{d_W(W, Z)}{\pi\sigma}}.$$

Similarly, we can show that

$$\mathbb{P}(W \leq z) - \mathbb{P}(Z \leq z) \geq -\left\{2 + \log\left(\frac{2}{\sqrt{\pi}}\right) + \frac{1}{2}\log\left(\frac{\sigma}{d_W(W, Z)}\right)\right\}\sqrt{\frac{d_W(W, Z)}{\pi\sigma}}.$$

Combining these bounds proves (4.2).

(iii) Let $0 < r < 1$. Then the SVG$(r, \sigma, 0)$ density is unbounded as $x \to 0$ and is a decreasing function of x for $x > 0$ and an increasing function for $x < 0$. Therefore, we argue as we did in part (ii) and bound $\mathbb{P}(0 \leq Z \leq \alpha)$ and then substitute into (6.6). Let $\nu = \frac{r-1}{2}$, so that $-\frac{1}{2} < \nu < 0$. We have

$$\mathbb{P}(0 \leq Z \leq \alpha) = \frac{1}{\sigma\sqrt{\pi}2^{\nu}\Gamma(\nu + \frac{1}{2})} \int_0^\alpha \left(\frac{t}{\sigma}\right)^\nu K_\nu\left(\frac{t}{\sigma}\right) dt$$

$$= \frac{1}{\sqrt{\pi}2^{\nu}\Gamma(\nu + \frac{1}{2})} \int_0^{\frac{\alpha}{\sigma}} y^{2\nu} \cdot y^{-\nu} K_{-\nu}(y) \, dy$$

$$\leq \frac{1}{\sqrt{\pi}2^{\nu}\Gamma(\nu + \frac{1}{2})} \int_0^{\frac{\alpha}{\sigma}} 2^{-\nu-1}\Gamma(-\nu)y^{2\nu} \, dy$$

$$= \frac{\Gamma(-\nu)}{\sqrt{\pi}2^{2\nu+1}\Gamma(\nu + \frac{1}{2})} \frac{1}{2\nu + 1}\left(\frac{\alpha}{\sigma}\right)^{2\nu+1} = C_{\nu,\sigma}\alpha^{2\nu+1}, \qquad (6.8)$$

where we used a change in variables and (A.2) in the second step and Lemma 6.1 in the third. We therefore have that, for any $z \in \mathbb{R}$ and $\alpha > 0$,

$$\mathbb{P}(W \leq z) - \mathbb{P}(Z \leq z) \leq \frac{1}{2\alpha}d_{\mathrm{W}}(W, Z) + C_{\nu,\sigma}\alpha^{2\nu+1}.$$

To optimise, we take $\alpha = \left(\frac{d_{\mathrm{W}}(W,Z)}{2(2\nu+1)C_{\nu,\sigma}}\right)^{\frac{1}{2(\nu+1)}}$, which results in the bound

$$\mathbb{P}(W \leq z) - \mathbb{P}(Z \leq z) \leq 2\big(2(2\nu + 1)C_{\nu,\sigma}\big)^{\frac{1}{2(\nu+1)}}\big(d_{\mathrm{W}}(W, Z)\big)^{\frac{2\nu+1}{2(\nu+1)}}$$

$$= 2\left(\frac{2\Gamma(-\nu)}{\sqrt{\pi}(2\sigma)^{2\nu+1}\Gamma(\nu + \frac{1}{2})}\right)^{\frac{1}{2(\nu+1)}}\big(d_{\mathrm{W}}(W, Z)\big)^{\frac{2\nu+1}{2(\nu+1)}}.$$

As in part (ii), we can similarly obtain a lower bound, and on substituting $\nu = \frac{r-1}{2}$ we obtain (4.3), which completes the proof. □

Appendix A: Properties of Modified Bessel Functions

Here, we list standard properties and inequalities for modified Bessel functions that are used throughout this paper. All formulas can be found in [38], except for the differentiation formulas (A.12)–(A.13), which can be found in [15] and [20], and the inequalities.

The *modified Bessel function of the first kind* of order $\nu \in \mathbb{R}$ is defined, for $x \in \mathbb{R}$, by

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{\Gamma(\nu + k + 1)k!} \left(\frac{x}{2}\right)^{\nu+2k}.$$

The *modified Bessel function of the second kind* of order $\nu \in \mathbb{R}$ is defined, for $x > 0$, by

$$K_\nu(x) = \int_0^{\infty} e^{-x \cosh(t)} \cosh(\nu t) \, \mathrm{d}t. \tag{A.1}$$

It is clear from (A.1) that

$$K_{-\nu}(x) = K_\nu(x). \tag{A.2}$$

The modified Bessel functions have the following asymptotic behaviour:

$$I_\nu(x) \sim \frac{1}{\Gamma(\nu + 1)} \left(\frac{x}{2}\right)^\nu \left(1 + O(x^2)\right), \qquad x \downarrow 0, \tag{A.3}$$

$$K_\nu(x) \sim \begin{cases} 2^{|\nu|-1}\Gamma(|\nu|)x^{-|\nu|}\left(1 + O(x^2)\right), & x \downarrow 0, \ \nu \neq 0, \\ -\log x, & x \downarrow 0, \ \nu = 0, \end{cases} \tag{A.4}$$

$$I_\nu(x) \sim \frac{e^x}{\sqrt{2\pi x}}, \qquad x \to \infty, \tag{A.5}$$

$$K_\nu(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x}, \qquad x \to \infty. \tag{A.6}$$

The following differentiation formulas hold:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(K_0(x)\right) = -K_1(x), \tag{A.7}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x^\nu K_\nu(x)\right) = -x^\nu K_{\nu-1}(x), \tag{A.8}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{I_\nu(x)}{x^\nu}\right) = \frac{I_{\nu+1}(x)}{x^\nu}, \tag{A.9}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{K_\nu(x)}{x^\nu}\right) = -\frac{K_{\nu+1}(x)}{x^\nu}, \tag{A.10}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{I_\nu(x)}{x^\nu}\right) = \frac{I_\nu(x)}{x^\nu} - \frac{(2\nu + 1)I_{\nu+1}(x)}{x^{\nu+1}}, \tag{A.11}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{K_\nu(x)}{x^\nu}\right) = \frac{K_\nu(x)}{x^\nu} + \frac{(2\nu + 1)K_{\nu+1}(x)}{x^{\nu+1}}, \tag{A.12}$$

$$\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right) = -\frac{(2\nu + 1)K_\nu(x)}{x^\nu} - \left(1 + \frac{(2\nu + 1)(2\nu + 2)}{x^2}\right)\frac{K_{\nu+1}(x)}{x^\nu}. \tag{A.13}$$

Applying the inequality $I_{\mu+1}(x) < I_{\mu}(x)$, $x > 0$, $\mu > -\frac{1}{2}$ [49] to the sixth differentiation formula of Corollary 1 of [20] gives the inequality

$$\frac{d^3}{dx^3}\left(\frac{I_{\nu}(x)}{x^{\nu}}\right) < \frac{I_{\nu}(x)}{x^{\nu}}, \quad x > 0, \ \nu > -\frac{1}{2}. \tag{A.14}$$

The next inequality follows from two applications of inequality (2.6) of [17]. For $x \geq 0$,

$$\int_0^x \int_0^u t^{\nu} I_{\nu}(t) \, dt \, du \leq \frac{2(\nu+2)}{2\nu+1} x^{\nu} I_{\nu+2}(x), \quad \nu > -\frac{1}{2}. \tag{A.15}$$

The following bounds, which can be found in [18,21], are used to bound the solution to the SVG Stein equation. Let $\nu > -\frac{1}{2}$. Then, for all $x \geq 0$,

$$\frac{K_{\nu}(x)}{x^{\nu}} \int_0^x t^{\nu+1} I_{\nu}(t) \, dt < \frac{1}{2}, \tag{A.16}$$

$$\frac{I_{\nu}(x)}{x^{\nu}} \int_x^{\infty} t^{\nu+1} K_{\nu}(t) \, dt < 1, \tag{A.17}$$

$$\frac{K_{\nu}(x)}{x^{\nu}} \int_0^x t^{\nu} I_{\nu}(t) \, dt \leq \frac{1}{2\nu+1}, \tag{A.18}$$

$$\frac{I_{\nu}(x)}{x^{\nu}} \int_x^{\infty} t^{\nu} K_{\nu}(t) \, dt \leq \frac{\sqrt{\pi}\,\Gamma(\nu+\frac{1}{2})}{2\Gamma(\nu+1)}, \tag{A.19}$$

$$\frac{K_{\nu}(x)}{x^{\nu-1}} \int_0^x t^{\nu} I_{\nu}(t) \, dt < \frac{\nu+1}{2\nu+1}, \tag{A.20}$$

$$\frac{I_{\nu}(x)}{x^{\nu-1}} \int_x^{\infty} t^{\nu} K_{\nu}(t) \, dt < 1, \tag{A.21}$$

$$\frac{K_{\nu+1}(x)}{x^{\nu-1}} \int_0^x t^{\nu} I_{\nu}(t) \, dt < \frac{\nu+1}{2\nu+1}, \tag{A.22}$$

$$\frac{I_{\nu+1}(x)}{x^{\nu-1}} \int_x^{\infty} t^{\nu} K_{\nu}(t) \, dt < \frac{1}{2}. \tag{A.23}$$

References

1. Arras, B., Azmoodeh, E., Poly, G. Swan, Y.A.: bound on the 2-Wasserstein distance between linear combinations of independent random variables. To appear in Stoch. Proc. Appl. (2018)+
2. Bai, S., Taqqu, M.: Behavior of the generalized Rosenblatt process at extreme critical exponent values. Ann. Probab. **45**, 1278–1324 (2017)
3. Barbour, A.D., Holst, L., Janson, S.: Poisson Approximation. Oxford University Press, Oxford (1992)
4. Blaisdell, B.: A measure of the similarity of sets of sequences not requiring sequence alignment. Proc. Natl. Acad. Sci. USA **83**, 5155–5159 (1986)
5. Chatterjee, S., Fulman, J., Röllin, A.: Exponential approximation by Stein's method and spectral graph theory. ALEA Lat. Am. J. Probab. Math. Stat **8**, 197–223 (2011)
6. Chen, L.H.Y., Goldstein, L., Shao, Q.-M.: Normal Approximation by Stein's Method. Springer, Berlin (2011)

7. Chen, L. H. Y., Röllin, A.: Stein couplings for normal approximation (2010). arXiv:1003.6039
8. Döbler, C.: Distributional transformations without orthogonality relations. J. Theor. Probab. **30**, 85–116 (2017)
9. Döbler, C., Gaunt, R.E., Vollmer, S.J.: An iterative technique for bounding derivatives of solutions of Stein equations. Electron. J. Probab. **22**(96), 1–39 (2017)
10. Eberlein, E., Hammerstein, E.: Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes. In: Dalang, R. C. Dozzi, M., Russo, F. (Eds.), Seminar on Stochastic Analysis, Random Fields and Applications IV, in: *Progress in Probability* **58** Birkhäuser Verlag, (2004), pp. 105–153
11. Eden, R., Viquez, J.: Nourdin–Peccati analysis on Wiener and Wiener–Poisson space for general distributions. Stoch. Proc. Appl. **125**, 182–216 (2015)
12. Eichelsbacher, P., Thäle, C.: Malliavin–Stein method for Variance-Gamma approximation on Wiener space. Electron. J. Probab. **20**(123), 1–28 (2015)
13. Elezović, N., Giordano, C., Pečarić, J.: The best bounds in Gautschi's inequality. Math. Inequal. Appl. **3**, 239–252 (2000)
14. Finlay, R., Seneta, E.: Option pricing with VG-like models. Int. J. Theor. Appl. Finan. **11**, 943–955 (2008)
15. Gaunt, R. E.: Rates of Convergence of Variance-Gamma Approximations via Stein's Method. D.Phil. thesis, University of Oxford (2013)
16. Gaunt, R.E.: Variance-Gamma approximation via Stein's method. Electron. J. Probab. **19**(38), 1–33 (2014)
17. Gaunt, R.E.: Inequalities for modified Bessel functions and their integrals. J. Math. Anal. Appl. **420**, 373–386 (2014)
18. Gaunt, R.E.: Uniform bounds for expressions involving modified Bessel functions. Math. Inequal. Appl. **19**, 1003–1012 (2016)
19. Gaunt, R.E.: On Stein's method for products of normal random variables and zero bias couplings. Bernoulli **23**, 3311–3345 (2017)
20. Gaunt, R.E.: Derivative formulas for Bessel, Struve and Anger–Weber functions. J. Class. Anal. **11**, 69–78 (2017)
21. Gaunt, R.E.: Inequalities for integrals of modified Bessel functions and expressions involving them. J. Math. Anal. Appl. **462**, 172–190 (2018)
22. Gaunt, R.E.: Inequalities for some integrals involving modified Bessel functions. To appear in P. Am. Math. Soc. (2018)+
23. Gaunt, R. E.: Wasserstein and Kolmogorov error bounds for variance-gamma approximation via Stein's method II (2018, in preparation)
24. Gaunt, R.E., Pickett, A.M., Reinert, G.: Chi-square approximation by Stein's method with application to Pearson's statistic. Ann. Appl. Probab. **27**, 720–756 (2017)
25. Gautschi, W.: Some elementary inequalities relating to the gamma and incomplete gamma function. J. Math. Phys. **38**, 77–81 (1959)
26. Giordanoa, C., Laforgia, A.: Inequalities and monotonicity properties for the gamma function. J. Comput. Appl. Math. **133**, 387–396 (2001)
27. Goldstein, L., Reinert, G.: Stein's method and the zero bias transformation with application to simple random sampling. Ann. Appl. Probab. **7**, 935–952 (1997)
28. Kalashnikov, V.: Geometric Sums: Bounds for Rare Events with Applications: Risk Analysis, Reliability, Queueing. Kluwer Academic Publishers Group, Dordrecht (1997)
29. Kotz, S., Kozubowski, T.J., Podgórski, K.: A Revisit with New Applications. The Laplace Distribution and Generalizations. Springer, Berlin (2001)
30. Ley, C., Reinert, G., Swan, Y.: Stein's method for comparison of univariate distributions. Probab. Surv. **14**, 1–52 (2017)
31. Lippert, R.A., Huang, H., Waterman, M.S.: Distributional regimes for the number of k-word matches between two random sequences. P. Natl. Acad. Sci. USA **99**, 13980–13989 (2002)
32. Madan, D.B., Seneta, E.: The Variance Gamma (V.G.) model for share market returns. J. Bus **63**, 511–524 (1990)
33. Madan, D.B., Carr, P., Chang, E.C.: The variance gamma process and option pricing. Eur. Finance Rev. **2**, 74–105 (1998)
34. Nourdin, I., Peccati, G.: Stein's method on Wiener chaos. Probab. Theory Rel. **145**, 75–118 (2009)
35. Nourdin, I., Peccati, G.: Cumulants on the Wiener space. J. Funct. Anal. **258**, 3775–3791 (2010)

36. Nourdin, I., Peccati, G.: Normal Approximations with Malliavin Calculus: from Stein's Method to Universality, vol. 192. Cambridge University Press, Cambridge (2012)
37. Nourdin, I., Peccati, G.: The optimal fourth moment theorem. P. Am. Math. Soc. **143**, 3123–3133 (2015)
38. Olver, F.W.J., Lozier, D.W., Boisvert, R.F., Clark, C.W.: NIST Handbook of Mathematical Functions. Cambridge University Press, Cambridge (2010)
39. Peköz, E., Röllin, A.: New rates for exponential approximation and the theorems of Rényi and Yaglom. Ann. Probab. **39**, 587–608 (2011)
40. Peköz, E., Röllin, A., Ross, N.: Total variation error bounds for geometric approximation. Bernoulli **19**, 610–632 (2013)
41. Peköz, E., Röllin, A., Ross, N.: Degree asymptotics with rates for preferential attachment random graphs. Ann. Appl. Probab. **23**, 1188–1218 (2013)
42. Peköz, E., Röllin, A., Ross, N.: Generalized gamma approximation with rates for urns, walks and trees. Ann. Probab. **44**, 1776–1816 (2016)
43. Pike, J., Ren, H.: Stein's method and the Laplace distribution. ALEA Lat. Am. J. Probab. Math. Stat **11**, 571–587 (2014)
44. Pitman, J., Ross, N.: Archimedes, gauss, and stein. Not. Am. Math. Soc. **59**, 1416–1421 (2012)
45. Reinert, G., Chew, D., Sun, F., Waterman, M.S.: Alignment free sequence comparison (I): statistics and power. J. Comput. Biol. **16**, 1615–1634 (2009)
46. Reinert, G., Schbath, S., Waterman, M.S., Lothaire, M.: Applied Combinatorics on Words. Cambridge University Press, Cambridge (2005)
47. Rényi, A.: A characterization of Poisson processes. Magyar Tud. Akad. Mat. Kutató Int. Közl. **1**, 519–527 (1957)
48. Ross, N.: Fundamentals of Stein's method. Probab. Surv. **8**, 210–293 (2011)
49. Soni, R.P.: On an inequality for modified Bessel functions. J. Math. Phys. Camb. **44**, 406–407 (1965)
50. Stein, C.: A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, vol. 2, pp. 583–602. University California Press, Berkeley (1972)
51. Stein, C.: Approximate Computation of Expectations. IMS, Hayward, California (1986)
52. Toda, A.: A. Weak limit of the geometric sum of independent but not identically distributed random variables (2011). arXiv:1111.1786