

Variance-Gamma approximation via Stein’s method

Robert E. Gaunt*

Abstract

Variance-Gamma distributions are widely used in financial modelling and contain as special cases the normal, Gamma and Laplace distributions. In this paper we extend Stein’s method to this class of distributions. In particular, we obtain a Stein equation and smoothness estimates for its solution. This Stein equation has the attractive property of reducing to the known normal and Gamma Stein equations for certain parameter values. We apply these results and local couplings to bound the distance between sums of the form $\sum_{i,j,k=1}^{m,n,r} X_{ik} Y_{jk}$, where the X_{ik} and Y_{jk} are independent and identically distributed random variables with zero mean, by their limiting Variance-Gamma distribution. Through the use of novel symmetry arguments, we obtain a bound on the distance that is of order $m^{-1} + n^{-1}$ for smooth test functions. We end with a simple application to binary sequence comparison.

Keywords: Stein’s method; Variance-Gamma approximation; rates of convergence.

AMS MSC 2010: 60F05.

Submitted to EJP on September 17, 2013, final version accepted on March 26, 2014.

1 Introduction

In 1972, Stein [41] introduced a powerful method for deriving bounds for normal approximation. Since then, this method has been extended to many other distributions, such as the Poisson [10], Gamma [27], [29], Exponential [9], [31] and Laplace [13], [34]. Through the use of differential or difference equations, and various coupling techniques, Stein’s method enables many types of dependence structures to be treated, and also gives explicit bounds for distances between distributions.

At the heart of Stein’s method lies a characterisation of the target distribution and a corresponding characterising differential or difference equation. For example, Stein’s method for normal approximation rests on the following characterization of the normal distribution, which can be found in Stein [42], namely $Z \sim N(\mu, \sigma^2)$ if and only if

$$\mathbb{E}[\sigma^2 f'(Z) - (Z - \mu)f(Z)] = 0 \tag{1.1}$$

for all sufficiently smooth f . This gives rise to the following inhomogeneous differential equation, known as the Stein equation:

$$\sigma^2 f'(x) - (x - \mu)f(x) = h(x) - \mathbb{E}h(Z), \tag{1.2}$$

*University of Oxford, UK. E-mail: gaunt@stats.ox.ac.uk

where $Z \sim N(\mu, \sigma^2)$, and the test function h is a real-valued function. For any bounded test function, a solution f to (1.2) exists (see Lemma 2.4 of Chen et al. [11]). There are a number of techniques for obtaining Stein equations, such as the density approach of Stein et al. [43], the scope of which has recently been extended by Ley and Swan [24]. Another commonly used technique is a generator approach, introduced by Barbour [3]. This approach involves recognising the target the distribution as the stationary distribution of Markov process and then using the theory of generators of stochastic process to arrive at a Stein equation; for a detailed overview of this method see Reinert [36]. Luk [27] used this approach to obtain the following Stein equation for the $\Gamma(r, \lambda)$ distribution:

$$x f''(x) + (r - \lambda x) f'(x) = h(x) - \mathbb{E}h(X), \tag{1.3}$$

where $X \sim \Gamma(r, \lambda)$.

The next essential ingredient of Stein's method is smoothness estimates for the solution of the Stein equation. This can often be done by solving the Stein equation using standard solution methods for differential equations and then using direct calculations to bound the required derivatives of the solution (Stein [42] used the approach to bound the first two derivatives of the solution to the normal Stein equation (1.2)). The generator approach is also often used to obtain smoothness estimates. The use of probabilistic arguments to bound the derivatives of the solution often make it easier to arrive at smoothness estimates than through the use of analytical techniques. Luk [27] and Pickett [33] used the generator approach to bound k -th order derivatives of the solution of the $\Gamma(r, \lambda)$ Stein equation (1.3). Pickett's bounds are as follows

$$\|f^{(k)}\| \leq \left\{ \sqrt{\frac{2\pi}{r}} + \frac{2}{r} \right\} \|h^{(k-1)}\|, \quad k \geq 1, \tag{1.4}$$

where $\|f\| = \|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$ and $h^{(0)} \equiv h$.

In this paper we obtain the key ingredients required to extend Stein's method to the class of Variance-Gamma distributions. The Variance-Gamma distributions are defined as follows (this parametrisation is similar to that given in Finlay and Seneta [17]).

Definition 1.1 (Variance-Gamma distribution, first parametrisation). *The random variable X is said to have a Variance-Gamma distribution with parameters $r > 0$, $\theta \in \mathbb{R}$, $\sigma > 0$, $\mu \in \mathbb{R}$ if and only if it has probability density function given by*

$$p_{\text{VG}_1}(x; r, \theta, \sigma, \mu) = \frac{1}{\sigma \sqrt{\pi} \Gamma(\frac{r}{2})} e^{\frac{\theta}{\sigma^2}(x-\mu)} \left(\frac{|x-\mu|}{2\sqrt{\theta^2 + \sigma^2}} \right)^{\frac{r-1}{2}} K_{\frac{r-1}{2}} \left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2} |x-\mu| \right), \tag{1.5}$$

where $x \in \mathbb{R}$, and $K_\nu(x)$ is a modified Bessel function of the second kind; see Appendix B for a definition. If (1.5) holds then we write $X \sim \text{VG}_1(r, \theta, \sigma, \mu)$.

The density (1.5) may at first appear to be undefined in the limit $\sigma \rightarrow 0$, but this limit does in fact exist and this can easily be verified from the asymptotic properties of the modified Bessel function $K_\nu(x)$ (see formula (B.4) from Appendix B). As we shall see in Proposition 1.2 (below), taking the limit $\sigma \rightarrow 0$ and putting $\mu = 0$ gives the family of Gamma distributions. It is also worth noting that the support of the Variance-Gamma distributions is \mathbb{R} when $\sigma > 0$, but in the limit $\sigma \rightarrow 0$ the support is the region (μ, ∞) if $\theta > 0$, and is $(-\infty, \mu)$ if $\theta < 0$.

The Variance-Gamma distributions were introduced to the financial literature by Madan and Seneta [28]. For certain parameter values the Variance-Gamma distributions have semi heavy tails that decay slower than the tails of the normal distribution, and therefore are often appropriate for financial modelling.

The class of Variance-Gamma distributions includes the Laplace distribution as a special case and in the appropriate limits reduces to the normal and Gamma distributions. This family of distributions also contains many other distributions that are of interest, which we list in the following proposition (the proof is given in Appendix A). As far as the author is aware, this is the first list of characterisations of the Variance-Gamma distributions to appear in the literature.

Proposition 1.2. (i) Let $\sigma > 0$ and $\mu \in \mathbb{R}$ and suppose that Z_r has the $\text{VG}_1(r, 0, \sigma/\sqrt{r}, \mu)$ distribution. Then Z_r converges in distribution to a $N(\mu, \sigma^2)$ random variable in the limit $r \rightarrow \infty$.

(ii) Let $\sigma > 0$ and $\mu \in \mathbb{R}$, then a $\text{VG}_1(2, 0, \sigma, \mu)$ random variable has the Laplace(μ, σ) distribution with probability density function

$$p_{\text{VG}_1}(x; 2, 0, \sigma, \mu) = \frac{1}{2\sigma} \exp\left(-\frac{|x - \mu|}{\sigma}\right), \quad x \in \mathbb{R}. \tag{1.6}$$

(iii) Suppose that (X, Y) has the bivariate normal distribution with correlation ρ and marginals $X \sim N(0, \sigma_X^2)$ and $Y \sim N(0, \sigma_Y^2)$. Then the product XY follows the $\text{VG}_1(1, \rho\sigma_X\sigma_Y, \sigma_X\sigma_Y\sqrt{1 - \rho^2}, 0)$ distribution.

(iv) Let X_1, \dots, X_r and Y_1, \dots, Y_r be independent standard normal random variables. Then $\mu + \sigma \sum_{k=1}^r X_k Y_k$ has the $\text{VG}_1(r, 0, \sigma, \mu)$ distribution. As a special case we have that a Laplace random variable with density (A.1) has the representation $\mu + \sigma(X_1 Y_1 + X_2 Y_2)$.

(v) The Gamma distribution is a limiting case of the Variance-Gamma distributions: for $r > 0$ and $\lambda > 0$, the random variable $X_\sigma \sim \text{VG}_1(2r, (2\lambda)^{-1}, \sigma, 0)$ converges in distribution to a $\Gamma(r, \lambda)$ random variable in the limit $\sigma \downarrow 0$.

(vi) Suppose that (X, Y) follows a bivariate gamma distribution with correlation ρ and marginals $X \sim \Gamma(r, \lambda_1)$ and $Y \sim \Gamma(r, \lambda_2)$. Then the random variable $X - Y$ has the $\text{VG}_1(2r, (2\lambda_1)^{-1} - (2\lambda_2)^{-1}, (\lambda_1 \lambda_2)^{-1/2}(1 - \rho)^{1/2}, 0)$ distribution.

The representations of the Variance-Gamma distributions given in Proposition 1.2 enable us to determine a number of statistics that may have asymptotic Variance-Gamma distributions.

One of the main results of this paper (see Lemma 3.1) is the following Stein equation for the Variance-Gamma distributions:

$$\sigma^2(x - \mu)f''(x) + (\sigma^2 r + 2\theta(x - \mu))f'(x) + (r\theta - (x - \mu))f(x) = h(x) - \text{VG}_{\sigma, \mu}^{r, \theta} h, \tag{1.7}$$

where $\text{VG}_{\sigma, \mu}^{r, \theta} h$ denotes the quantity $\mathbb{E}h(X)$ for $X \sim \text{VG}_1(r, \theta, \sigma, \mu)$. We also obtain uniform bounds for the first four derivatives of the solution of the Stein equation for the case $\theta = 0$.

In Section 3, we analyse the Stein equation (1.7). In particular, we show that the normal Stein equation (1.2) and Gamma Stein equation (1.3) are special cases. As a Stein equation for a given distribution is not unique (see Barbour [2]), the fact that in the appropriate limit the Variance-Gamma Stein equation (1.7) reduces to the known normal and Gamma Stein equation is an attractive feature.

Stein's method has also recently been extended to the Laplace distribution (see Pike and Ren [34] and Döbler [13]), although the Laplace Stein equation obtained by [34] differs from the Laplace Stein equation that arises as a special case of (1.7); see Section 3.1.1 for a more detailed discussion. Another special case of the Stein equation (1.7) is a Stein equation for the product of two independent central normal random variables, which is in agreement with the Stein equation for products of independent central normal that was recently obtained by Gaunt [20]. Therefore, the results from this paper allow the existing literature for Stein's method for normal, Gamma, Laplace and product normal approximation to be considered in a more general framework.

More importantly, our development of Stein's method for the Variance-Gamma distributions allows a number of new situations to be treated by Stein's method. In Section 4, we illustrate our method by obtaining a bound for the distance between the statistic

$$W_r = \sum_{i,j,k=1}^{m,n,r} X_{ik}Y_{jk} = \sum_{k=1}^r \left(\frac{1}{\sqrt{m}} \sum_{i=1}^m X_{ik} \right) \left(\frac{1}{\sqrt{n}} \sum_{j=1}^n Y_{jk} \right), \tag{1.8}$$

where the X_{ik} and Y_{jk} are independent and identically distributed with zero mean, and its asymptotic distribution, which, by the central limit theorem and part (iv) of Proposition 1.2, is the $\text{VG}_1(r, 0, 1, 0)$ distribution. By using the $\text{VG}_1(r, 0, 1, 0)$ Stein equation

$$xf''(x) + rf'(x) - xf(x) = h(x) - \text{VG}_{1,0}^{r,0}h, \tag{1.9}$$

local approach couplings, and symmetry arguments, that were introduced by Pickett [33], we obtain a $O(m^{-1} + n^{-1})$ bound for smooth test functions. A similar phenomena was observed in chi-square approximation by Pickett, and also by Goldstein and Reinert [21] in which they obtained $O(n^{-1})$ convergence rates in normal approximation, for smooth test functions, under the assumption of vanishing third moments. For non-smooth test functions we would, however, expect a $O(m^{-1/2} + n^{-1/2})$ convergence rate (cf. Berry-Esséen Theorem (Berry [6] and Esséen [16]) to hold; see Remark 4.11.

The rest of this paper is organised as follows. In Section 2, we introduce the Variance-Gamma distributions and state some of their standard properties. In Section 3, we obtain a characterising lemma for the Variance-Gamma distributions and a corresponding Stein equation. We also obtain the unique bounded solution of the Stein equation, and present uniform bounds for the first four derivatives of the solution for the case $\theta = 0$. In Section 4, we use Stein's method for Variance-Gamma approximation to bound the distance between the statistic (1.8) and its limiting Variance-Gamma distribution. We then apply this bound to an application of binary sequence comparison, which is a simple special case of the more general problem of word sequence comparison. In Appendix A, we include the proofs of some technical lemmas that are required in this paper. Appendix B provides a list of some elementary properties of modified Bessel functions that we make use of in this paper.

2 The class of Variance-Gamma distributions

In this section we present the Variance-Gamma distributions and some of their standard properties. Throughout this paper we will make use of two different parametrisations of the Variance-Gamma distributions; the first parametrisation was given in Section 1, and making the change of variables

$$\nu = \frac{r-1}{2}, \quad \alpha = \frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}, \quad \beta = \frac{\theta}{\sigma^2}. \tag{2.1}$$

leads to another useful parametrisation. This parametrisation can be found in Eberlein and Hammerstein [15].

Definition 2.1 (Variance-Gamma distribution, second parametrisation). *The random variable X is said to have a Variance-Gamma distribution with parameters ν, α, β, μ , where $\nu > -1/2, \mu \in \mathbb{R}, \alpha > |\beta|$, if and only if its probability density function is given by*

$$p_{\text{VG}_2}(x; \nu, \alpha, \beta, \mu) = \frac{(\alpha^2 - \beta^2)^{\nu+1/2}}{\sqrt{\pi}\Gamma(\nu + \frac{1}{2})} \left(\frac{|x - \mu|}{2\alpha} \right)^\nu e^{\beta(x-\mu)} K_\nu(\alpha|x - \mu|), \quad x \in \mathbb{R}. \tag{2.2}$$

If (2.2) holds then we write $X \sim \text{VG}_2(\nu, \alpha, \beta, \mu)$.

Definition 2.2. If $X \sim \text{VG}_1(r, 0, \sigma, \mu)$, for r , σ , and μ defined as in Definition 1.5 (or equivalently $X \sim \text{VG}_2(\nu, \alpha, 0, \mu)$), then X is said to have a Symmetric Variance-Gamma distribution.

The first parametrisation leads to simple characterisations of the Variance-Gamma distributions in terms of normal and Gamma distributions, and therefore in many cases it allows us to recognise statistics that will have an asymptotic Variance-Gamma distribution. For this reason, we state our main results in terms of this parametrisation. However, the second parametrisation proves to be very useful in simplifying the calculations of Section 3, as the solution of the Variance-Gamma Stein equation has a simpler representation for this parametrisation. We can then state the results in terms of the first parametrisation by using (2.1).

The Variance-Gamma distributions have moments of arbitrary order (see Eberlein and Hammerstein [15]), in particular the mean and variance (for both parametrisations) of a random variable X with a Variance-Gamma distribution are given by

$$\begin{aligned} \mathbb{E}X &= \mu + \frac{(2\nu + 1)\beta}{\alpha^2 - \beta^2} = \mu + r\theta, \\ \text{Var}X &= \frac{2\nu + 1}{\alpha^2 - \beta^2} \left(1 + \frac{2\beta^2}{\alpha^2 - \beta^2} \right) = r(\sigma^2 + 2\theta^2). \end{aligned} \tag{2.3}$$

The following proposition, which can be found in Bibby and Sørensen [7], shows that the class of Variance-Gamma distributions is closed under convolution, provided that the random variables have common values of θ and σ (or, equivalently, common values of α and β in the second parametrisation).

Proposition 2.3. Let X_1 and X_2 be independent random variables such that $X_i \sim \text{VG}_1(r_i, \theta, \sigma, \mu_i)$, $i = 1, 2$, then we have that

$$X_1 + X_2 \sim \text{VG}_1(r_1 + r_2, \theta, \sigma, \mu_1 + \mu_2).$$

Variance-Gamma random variables can be characterised in terms of independent normal and Gamma random variables. This characterisation is given in the following proposition, which can be found in Barndorff-Nielsen et al. [5].

Proposition 2.4. Let $r > 0$, $\theta \in \mathbb{R}$, $\sigma > 0$ and $\mu \in \mathbb{R}$. Suppose that $U \sim N(0, 1)$ and $V \sim \Gamma(r/2, 1/2)$ are independent random variables and let $Z \sim \text{VG}_1(r, \theta, \sigma, \mu)$, then

$$Z \stackrel{\mathcal{D}}{=} \mu + \theta V + \sigma \sqrt{V} U.$$

Using Proposition 2.4 we can establish the following useful representation of the Variance-Gamma distributions, which appears to be a new result. Indeed, the representation allows us to see that the statistic (1.8) has a limiting Variance-Gamma distribution.

Corollary 2.5. Let $\theta \in \mathbb{R}$, $\sigma > 0$, $\mu \in \mathbb{R}$, and r be a positive integer. Let X_1, X_2, \dots, X_r and Y_1, Y_2, \dots, Y_r be independent standard normal random variables and let Z be a $\text{VG}_1(r, \theta, \sigma, \mu)$ random variable, then

$$Z \stackrel{\mathcal{D}}{=} \mu + \theta \sum_{i=1}^r X_i^2 + \sigma \sum_{i=1}^r X_i Y_i.$$

Proof. Let X_1, X_2, \dots, X_r and Y_1, Y_2, \dots, Y_r be sequences of independent standard normal random variables. Then X_i^2 , $i = 1, 2, \dots, m$, has a $\chi_{(1)}^2$ distribution, that is a $\Gamma(1/2, 1/2)$ distribution. Define

$$Z_1 = \mu + \theta X_1^2 + \sigma X_1 Y_1, \quad Z_i = \theta X_i^2 + \sigma X_i Y_i, \quad i = 2, 3, \dots, r.$$

Note that $X_i Y_i \stackrel{D}{=} |X_i| Y_i$. Hence, by Proposition 2.4, we have that Z_1 is a $\text{VG}_1(1, \theta, \sigma, \mu)$ random variable and $Z_i, i = 2, \dots, r$, are $\text{VG}_1(1, \theta, \sigma, 0)$ random variables. It therefore follows from Proposition 2.3 that the sum $Z = \sum_{i=1}^r Z_i$ follows the $\text{VG}_1(r, \theta, \sigma, \mu)$ distribution. \square

3 Stein's method for Variance-Gamma distributions

3.1 A Stein equation for the Variance-Gamma distributions

The following lemma, which characterises the Variance-Gamma distributions, will lead to a Stein equation for the Variance-Gamma distributions. Before stating the lemma, we note that an application of the asymptotic formula (B.4) to the density function (2.2) allows us to deduce the tail behaviour of the $\text{VG}_2(\nu, \alpha, \beta, \mu)$ distribution:

$$p_{\text{VG}_2}(x; \nu, \alpha, \beta, \mu) \sim \begin{cases} \frac{1}{\pi\Gamma(\nu + \frac{1}{2})} \left(\frac{\alpha^2 - \beta^2}{2\alpha}\right)^{\nu + \frac{1}{2}} x^{\nu - \frac{1}{2}} e^{-(\alpha - \beta)(x - \mu)}, & x \rightarrow \infty, \\ \frac{1}{\pi\Gamma(\nu + \frac{1}{2})} \left(\frac{\alpha^2 - \beta^2}{2\alpha}\right)^{\nu + \frac{1}{2}} (-x)^{\nu - \frac{1}{2}} e^{(\alpha + \beta)(x - \mu)}, & x \rightarrow -\infty. \end{cases} \quad (3.1)$$

Note that the tails are in general not symmetric.

Lemma 3.1. *Let W be a real-valued random variable. Then W follows the $\text{VG}_2(\nu, \alpha, \beta, \mu)$ distribution if and only if*

$$\mathbb{E}[(W - \mu)f''(W) + (2\nu + 1 + 2\beta(W - \mu))f'(W) + ((2\nu + 1)\beta - (\alpha^2 - \beta^2)(W - \mu))f(W)] = 0 \quad (3.2)$$

for all piecewise twice continuously differentiable functions $f : \mathbb{R} \rightarrow \mathbb{R}$ that satisfy

$$\lim_{x \rightarrow \infty} f^{(k)}(x)x^{\nu + 3/2}e^{-(\alpha - \beta)x} = 0 \quad \text{and} \quad \lim_{x \rightarrow -\infty} f^{(k)}(x)(-x)^{\nu + 3/2}e^{(\alpha + \beta)x} = 0 \quad (3.3)$$

for $k = 0, 1, 2$, where $f^{(0)} \equiv f$.

Proof. To simplify the calculations, we prove the result for the special case $\mu = 0$, $\alpha = 1$, $-1 < \beta < 1$. For $W = \alpha(Z - \mu)$ we have that $W \sim \text{VG}_2(\nu, 1, \beta, 0)$ if and only if $Z \sim \text{VG}_2(\nu, \alpha, \alpha\beta, \mu)$, and so we can deduce the general case by applying a simple linear transformation.

Necessity. Suppose that $W \sim \text{VG}_2(\nu, 1, \beta, 0)$. We split the range of integration to obtain

$$\mathbb{E}[Wf''(W) + (2\nu + 1 + 2\beta W)f'(W) + ((2\nu + 1)\beta - (1 - \beta^2)W)f(W)] = I_1 + I_2,$$

where

$$I_1 = \int_0^\infty \{xf''(x) + (2\nu + 1 + 2\beta x)f'(x) + ((2\nu + 1)\beta - (1 - \beta^2)x)f(x)\}p(x) dx,$$

$$I_2 = \int_{-\infty}^0 \{xf''(x) + (2\nu + 1 + 2\beta x)f'(x) + ((2\nu + 1)\beta - (1 - \beta^2)x)f(x)\}p(x) dx,$$

and $p(x) = \kappa_{\nu, \beta} x^\nu e^{\beta x} K_\nu(x)$, where $\kappa_{\nu, \beta}$ is the normalising constant, is the density of W . The integrals I_1 and I_2 exist because f is piecewise twice continuously differentiable that satisfies the conditions of (3.3), which on recalling the tail behaviour of $p(x)$ given in (3.1) ensures that, for $k = 0, 1, 2$, we have that $xp(x)f^{(k)}(x) = o(|x|^{-1})$ as $|x| \rightarrow \infty$.

Firstly, we consider I_1 . Let $A(x) = x$, $B(x) = 2\nu + 1 + 2\beta x$ and $C(x) = (2\nu + 1)\beta - (1 - \beta^2)x$. Then applying integration by parts twice gives

$$\begin{aligned} I_1 &= \int_0^\infty \{A(x)p''(x) + (2A'(x) - B(x))p'(x) + (A''(x) - B'(x) + C(x))p(x)\}f(x) \, dx \\ &\quad + \left[A(x)p(x)f'(x) \right]_0^\infty + \left[\{B(x)p(x) - (A(x)p(x))'\}f(x) \right]_0^\infty \\ &= \int_0^\infty \{xp''(x) + (-2\nu + 1 - 2\beta x)p'(x) + ((2\nu - 1)\beta - (1 - \beta^2)x)p(x)\}f(x) \, dx \\ &\quad + \left[xp(x)f'(x) \right]_0^\infty + \left[(2\nu + 2\beta x)p(x)f(x) - xp'(x)f(x) \right]_0^\infty. \end{aligned}$$

Straightforward differentiation of the function $p(x) = \kappa_{\nu,\beta}x^\nu e^{\beta x}K_\nu(x)$ shows that the integrand in the above display is equal to

$$\kappa_{\nu,\beta}x^{\nu-1}e^{\beta x}\{x^2K''_\nu(x) + xK'_\nu(x) - (x^2 + \nu^2)K_\nu(x)\}f(x) = 0,$$

as $K_\nu(x)$ is a solution of the modified Bessel differential equation (see (B.10)).

We now note that $p(x) = O(x^{\nu-1/2}e^{-(1-\beta)x})$ as $x \rightarrow \infty$ (see (3.1)), and by differentiating $p(x)$ and using the asymptotic formula (B.4) for $K_\nu(x)$ we can see that $p'(x)$ is also of order $x^{\nu-1/2}e^{-(1-\beta)x}$ as $x \rightarrow \infty$. Hence, $xp(x)f'(x)$, $p(x)f(x)$, $xp(x)f(x)$ and $xp'(x)f(x)$ are equal to 0 in the limit $x \rightarrow \infty$. The terms $xp(x)f'(x)$ and $xp(x)f(x)$ are also equal to 0 at the origin, because f and f' are continuous and thus bounded at the origin. Hence, I_1 simplifies to

$$I_1 = -\lim_{x \downarrow 0} \{(2\nu + 2\beta x)p(x) - xp'(x)\}f(x). \tag{3.4}$$

Using formula (B.7) to differentiate $K_\nu(x)$ gives

$$\begin{aligned} I_1 &= -\kappa_{\nu,\beta} \lim_{x \downarrow 0} f(x)\{(2\nu + 2\beta x)x^\nu e^{\beta x}K_\nu(x) - x^\nu e^{\beta x}(xK'_\nu(x) + \nu K_\nu(x) + \beta xK_\nu(x))\} \\ &= -\kappa_{\nu,\beta} \lim_{x \downarrow 0} x^\nu f(x)\{-xK'_\nu(x) + (\nu + \beta x)K_\nu(x)\} \\ &= -\kappa_{\nu,\beta} \lim_{x \downarrow 0} x^\nu f(x)\{\frac{1}{2}x(K_{\nu+1}(x) - K_{\nu-1}(x)) + (\nu + \beta x)K_\nu(x)\}. \end{aligned}$$

We now calculate the limit in the above expression. We first consider the case $\nu > 0$. Applying the asymptotic formula (B.2) gives

$$I_1 = -\kappa_{\nu,\beta} \lim_{x \downarrow 0} \{2^{\nu-1}\Gamma(\nu + 1) + 2^{\nu-1}\nu\Gamma(\nu)\}f(x) = -\kappa_{\nu,\beta} \lim_{x \rightarrow 0^+} 2^\nu \Gamma(\nu + 1)f(x),$$

since $\nu\Gamma(\nu) = \Gamma(\nu + 1)$. Now consider the case $\nu = 0$. We use the fact that $K_1(x) = K_{-1}(x)$ to obtain

$$I_1 = -\kappa_{0,\beta} \lim_{x \downarrow 0} f(x)xK_1(x) = -\kappa_{0,\beta} \lim_{x \rightarrow 0^+} \Gamma(1)f(x) = -\kappa_{0,\beta} \lim_{x \rightarrow 0^+} 2^0\Gamma(1 + 1)f(x),$$

since $\Gamma(1) = \Gamma(2)$. Therefore we have

$$I_1 = -\kappa_{\nu,\beta} \lim_{x \downarrow 0} 2^\nu \Gamma(\nu + 1)f(x) \quad \text{for all } \nu \geq 0.$$

Finally, we consider the case $-1/2 < \nu < 0$. We use the fact that $K_{-\lambda}(x) = K_\lambda(x)$ to obtain

$$\begin{aligned} I_1 &= -\kappa_{\nu,\beta} \lim_{x \downarrow 0} \{\frac{1}{2}x^{\nu+1}(K_{\nu+1}(x) + K_{1-\nu}(x)) + \nu x^\nu K_{-\nu}(x)\}f(x) \\ &= -\kappa_{\nu,\beta} \lim_{x \downarrow 0} \{2^{\nu-1}\Gamma(\nu + 1) + 2^{\nu-1}(\Gamma(1 - \nu) - (-\nu)\Gamma(-\nu))x^{2\nu}\}f(x) \\ &= -\kappa_{\nu,\beta} \lim_{x \downarrow 0} 2^{\nu-1}\Gamma(\nu + 1)f(x) \quad \text{for } -1/2 < \nu < 0. \end{aligned}$$

A similar argument (with the difference being that here $p(x) = \kappa_{\nu,\beta}(-x)^\nu e^{\beta x} K_\nu(-x)$) shows that

$$I_2 = \begin{cases} \kappa_{\nu,\beta} \lim_{x \uparrow 0} 2^\nu \Gamma(\nu + 1) f(x), & \nu \geq 0, \\ \kappa_{\nu,\beta} \lim_{x \uparrow 0} 2^{\nu-1} \Gamma(\nu + 1) f(x), & -1/2 < \nu < 0. \end{cases}$$

As f is continuous, it follows that $I_1 = -I_2$ (and so $I_1 + I_2 = 0$), which completes the proof of necessity.

Sufficiency. For fixed $z \in \mathbb{R}$, let $f(x) := f_z(x)$ be a bounded solution to the differential equation

$$x f''(x) + (2\nu + 1 + 2\beta x) f'(x) + ((2\nu + 1)\beta - (1 - \beta^2)x) f(x) = \chi_{(-\infty, z]}(x) - K_{\nu,\beta}(z), \quad (3.5)$$

where $K_{\nu,\beta}(z)$ is the cumulative distribution function of the $\text{VG}_2(\nu, 1, \beta, 0)$ distribution. Using Lemma 3.3 (below) with $h(x) = \chi_{(-\infty, z]}(x)$ we see that a solution to (3.5) is given by

$$f_z(x) = -\frac{e^{-\beta x} K_\nu(|x|)}{|x|^\nu} \int_0^x e^{\beta y} |y|^\nu I_\nu(|y|) [\chi_{(-\infty, z]}(x) - K_{\nu,\beta}(z)] dy \\ - \frac{e^{-\beta x} I_\nu(|x|)}{|x|^\nu} \int_x^\infty e^{\beta y} |y|^\nu K_\nu(|y|) [\chi_{(-\infty, z]}(x) - K_{\nu,\beta}(z)] dy.$$

This solution and its first derivative are bounded (see Lemma 3.3) and is piecewise twice differentiable. As f_z and f'_z are bounded, they satisfy the condition (3.3) (with $\alpha = 1$) and f''_z must also satisfy the condition because, from (3.5),

$$|x f''_z(x)| \leq |(2\nu + 1 + 2\beta x) f'_z(x)| + |((2\nu + 1)\beta - (1 - \beta^2)x) f_z(x)| + 2 \leq A + B|x|$$

for some constants A and B . Hence, if (3.2) holds for all piecewise twice continuously differentiable functions satisfying (3.3) (with $\alpha = 1$), then by (3.5),

$$0 = \mathbb{E}[W f''_z(W) + (2\nu + 1 + 2\beta W) f'_z(W) + ((2\nu + 1)\beta - (1 - \beta^2)W) f_z(W)] \\ = \mathbb{E}[\chi_{(-\infty, z]}(W) - K_{\nu,\beta}(z)] \\ = \mathbb{P}(W \leq z) - K_{\nu,\beta}(z).$$

Therefore W has the $\text{VG}_2(\nu, 1, \beta, 0)$ distribution. □

Lemma 3.1 suggests the following Stein equation for the $\text{VG}_2(\nu, \alpha, \beta, \mu)$ distribution:

$$(x - \mu) f''(x) + (2\nu + 1 + 2\beta(x - \mu)) f'(x) + ((2\nu + 1)\beta - (\alpha^2 - \beta^2)(x - \mu)) f(x) = h(x) - \widetilde{\text{VG}}_{\beta,\mu}^{\nu,\alpha} h, \quad (3.6)$$

where $\widetilde{\text{VG}}_{\beta,\mu}^{\nu,\alpha} h$ denotes the quantity $\mathbb{E}(h(X))$ for $X \sim \text{VG}_2(\nu, \alpha, \beta, \mu)$.

In order to simplify the calculations of Section 3.2, we will make use of the Stein equation for the $\text{VG}_2(\nu, 1, \beta, 0)$ distribution, where $-1 < \beta < 1$. Results for the full parametrisation can then be recovered by making a simple linear transformation. For the $\text{VG}_2(\nu, 1, \beta, 0)$ distribution, the Stein equation (3.6) reduces to

$$x f''(x) + (2\nu + 1 + 2\beta x) f'(x) + ((2\nu + 1)\beta - (1 - \beta^2)x) f(x) = h(x) - \widetilde{\text{VG}}_{\beta,0}^{\nu,1} h. \quad (3.7)$$

Changing parametrisation in (3.6) via (2.1) and multiplying through by σ^2 gives the $\text{VG}_1(r, \theta, \sigma, \mu)$ Stein equation (1.7), which we presented in the introduction.

Remark 3.2. *The $\text{VG}_1(r, \theta, \sigma, \mu)$ Stein equation has the interesting property of being a (true) second order linear differential equation. Such Stein equations are uncommon in the literature, although Peköz et al. [32], Gaunt [20] and Pike and Ren [34] have*

obtained similar operators for the Kummer densities, the product of two mean zero normals, and the Laplace distribution, respectively. Gaunt [20] and Pike and Ren [34] used the method of variation of parameters (see Collins [12] for an account of the method) to solve their equations, whereas Peköz et al. used a substitution to turn their second order operator into a first order operator, which leads to a double integral solution. We attempted to follow this approach but the double integral solution we obtained was rather complicated. However, solving using variation of parameters lead to a representation of the solution (see Lemma 3.3) that enabled us to obtain uniform bounds for the solution and its first four derivatives (see Lemma 3.5 and Theorem 3.6).

We could have obtained a first order Stein operator for the $\text{VG}_1(r, \theta, \sigma, \mu)$ distributions using the density approach of Stein et al. [43]. However, this approach would lead to an operator involving the modified Bessel function $K_\nu(x)$. Using such a Stein equation to prove approximation results with standard coupling techniques would be difficult. In contrast, our $\text{VG}_1(r, \theta, \sigma, \mu)$ Stein equation is much more amenable to the use of couplings, as we shall see in Section 4. Peköz et al. [32] encountered a similar situation (the density approach would lead to an operator involving the Kummer function) and proceeded as we did by instead considering a second order operator with simple coefficients.

3.1.1 Special cases of the Variance-Gamma Stein equation

Here we note a number of interesting special cases of the $\text{VG}_1(r, \theta, \sigma, \mu)$ Stein equation. Whilst the Gamma distribution is not covered by Lemma 3.1, we note that letting $r = 2s$, $\theta = (2\lambda)^{-1}$, $\mu = 0$ and taking the limit $\sigma \rightarrow 0$ in (1.7) gives the Stein equation

$$\lambda^{-1}(xf'(x) + (s - \lambda x)f(x)) = h(x) - \text{VG}_{0,0}^{2s,(2\lambda)^{-1}}h,$$

which, recalling (1.3), we recognise as the $\Gamma(s, \lambda)$ Stein equation (1.3) of Luk [27] (up to a multiplicative factor).

We also note that a Stein equation for the $\text{VG}_1(r, 0, \sigma/\sqrt{r}, \mu)$ distribution is

$$\frac{\sigma^2}{r}(x - \mu)f''(x) + \sigma^2 f'(x) - (x - \mu)f(x) = h(x) - \text{VG}_{\sigma/\sqrt{r},\mu}^{r,0}h,$$

which in the limit $r \rightarrow \infty$ is the classical $N(\mu, \sigma^2)$ Stein equation.

Taking $r = 1$, $\sigma = \sigma_X \sigma_Y$ and $\mu = 0$ in (1.7) gives the following Stein equation for distribution of the product of independent $N(0, \sigma_X^2)$ and $N(0, \sigma_Y^2)$ random variables (see part (iii) of Proposition 1.2):

$$\sigma_X^2 \sigma_Y^2 x f''(x) + \sigma_X^2 \sigma_Y^2 f'(x) - x f(x) = h(x) - \text{VG}_{\sigma_X \sigma_Y, 0}^{1,0}h.$$

This Stein equation is in agreement with the Stein equation for the product of two independent, zero mean normal random variables that was obtained by Gaunt [20].

Finally, we deduce a Stein equation for the Laplace distribution. Recalling part (ii) of Proposition 1.2, we have that $\text{Laplace}(0, \sigma) = \text{VG}_1(2, 0, \sigma, 0)$. Thus, we deduce the following Stein equation for the Laplace distribution:

$$\sigma^2 x f''(x) + 2\sigma^2 f'(x) - x f(x) = h(x) - \mathbb{E}h(X), \tag{3.8}$$

where $X \sim \text{Laplace}(0, \sigma)$. Pike and Ren [34] have obtained an alternative Stein characterisation of the Laplace distribution, which leads to the initial value problem

$$f(x) - \sigma^2 f''(x) = h(x) - \mathbb{E}h(X), \quad f(0) = 0. \tag{3.9}$$

They have also solved (3.9) and have obtained uniform bounds for the solution and its first three derivatives. Their characterisation was obtained by a repeated application of the density method, and is similar to the characterisation for the Exponential distribution that results from the density method (see Stein et al. [43], Example 1.6), which leads to the Stein equation

$$f'(x) - \lambda f(x) + \lambda f(0+) = h(x) - \mathbb{E}h(Y), \tag{3.10}$$

where $Y \sim \text{Exp}(\lambda)$. Since $\text{Exp}(\lambda) = \Gamma(1, \lambda)$, equation (3.10) and the Gamma Stein equation (1.3) (with $r = 1$) give a choice of Stein equations for applications involving the Exponential distribution. Both equations have been shown to be effective in the study of Exponential approximation, but in certain situations one equation may prove to be more useful than the other; see, for example, Pickett [33] for a utilisation of (1.3), and Peköz and Röllin [31] for an application involving (3.10). We would expect a similar situation to occur with the Laplace Stein equations (3.8) and (3.9), although we do not further investigate the use of these Stein equations in Laplace approximation.

3.1.2 Applications of Lemma 3.1

The main application of Lemma 3.1 that is considered in this paper involves the use of the resulting Stein equation in the proofs of the limit theorems of Section 4. There are, however, other interesting results that follow from Lemma 3.1. We consider a couple here.

Suppose $W \sim \text{VG}_2(\nu, \alpha, \beta, 0)$. Then taking $f(x) = e^{tx}$, where $|t + \beta| < \alpha$ (which ensures that condition (3.3) is satisfied), in the charactering equation (3.2) and setting $M(t) = \mathbb{E}[e^{tW}]$, we deduce that $M(t)$ satisfies the differential equation

$$(t^2 + 2\beta t - (\alpha^2 - \beta^2))M'(t) + (2\nu + 1)(t + \beta)M(t) = 0.$$

Solving this equation subject to the condition $M(0) = 1$ then gives that the moment generating function of the Variance-Gamma distribution with $\mu = 0$ is

$$M(t) = \left(\frac{\alpha^2 - \beta^2}{\alpha^2 - (\beta + t)^2} \right)^{\nu+1/2} = (1 - 2\theta t + \sigma^2 t^2)^{-r/2}.$$

Similarly, taking $f(x) = x^k$ and setting $M_k = \mathbb{E}W^k$ leads to the following recurrence equation for the moments of the Variance-Gamma distributions with $\mu = 0$:

$$(\alpha^2 - \beta^2)M_{k+1} - \beta(2k + 2\nu + 1)M_k - k(2\nu + k)M_{k-1} = 0,$$

which in terms of the first parametrisation is

$$M_{k+1} - \theta(2k + r)M_k - \sigma^2 k(r + k - 1)M_{k-1} = 0.$$

We have that $M_0 = 1$ and $M_1 = (2\nu + 1)\beta/(\alpha^2 - \beta^2) = r\theta$ (see (2.3)), and thus we can solve these recurrence equations by forward substitution to obtain the moments of the Variance-Gamma distributions. As far as the author is aware, these recurrence equations are new, although Scott et al. [40] have already established a formula for the moments of general order of the Variance-Gamma distributions.

3.2 Smoothness estimates for the solution of the Stein equation

We now turn our attention to solving the $\text{VG}_2(\nu, 1, \beta, 0)$ Stein equation (3.7). Handling this particular set of restricted parameters simplifies the calculations and allows

us to write down the solution of $VG_1(r, \theta, \sigma, \mu)$ after a straightforward change of variables.

Since the homogeneous version of the $VG_2(\nu, 1, \beta, 0)$ Stein equation has a simple fundamental system of solutions (see the proof of Lemma 3.3 in Appendix A), we consider variation of parameters to be an appropriate method of solution. We carry out these calculations in Appendix A and present the solution in Lemma 3.3. We could also have solved the Stein equation by using generator theory. Multiplying both sides of (3.7) by $\frac{1}{x}$, we recognise the left-hand side of the equation as the generator of a Bessel process with drift with killing (for an account of the Bessel process with drift see Linetsky [25]). The Stein equation can then be solved using generator theory (see Durrett [14], pp. 249). For a more detailed account of the application of the generator approach to Stein's method for Variance-Gamma distributions see Gaunt [18].

In the following lemma we give the solution to the Stein equation. The proof is given in Appendix A.

Lemma 3.3. *Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a measurable function with $\mathbb{E}|h(X)| < \infty$, where $X \sim VG_2(\nu, 1, \beta, 0)$, and $\nu > -1/2$ and $-1 < \beta < 1$. Then a solution $f : \mathbb{R} \rightarrow \mathbb{R}$ to the Variance-Gamma Stein equation (3.7) is given by*

$$f(x) = -\frac{e^{-\beta x} K_\nu(|x|)}{|x|^\nu} \int_0^x e^{\beta y} |y|^\nu I_\nu(|y|) [h(y) - \widetilde{VG}_{\beta,0}^{\nu,1} h] dy - \frac{e^{-\beta x} I_\nu(|x|)}{|x|^\nu} \int_x^\infty e^{\beta y} |y|^\nu K_\nu(|y|) [h(y) - \widetilde{VG}_{\beta,0}^{\nu,1} h] dy, \quad (3.11)$$

where the modified Bessel functions $I_\nu(x)$ and $K_\nu(x)$ are defined in Appendix B. Suppose further that h is bounded, then $f(x)$ and $f'(x)$ are bounded for all $x \in \mathbb{R}$. Moreover, this is the unique bounded solution when $\nu \geq 0$ and $-1 < \beta < 1$.

Remark 3.4. *The equality*

$$\int_{-\infty}^x e^{\beta y} |y|^\nu K_\nu(|y|) [h(y) - \widetilde{VG}_{\beta,0}^{\nu,1} h] dy = - \int_x^\infty e^{\beta y} |y|^\nu K_\nu(|y|) [h(y) - \widetilde{VG}_{\beta,0}^{\nu,1} h] dy \quad (3.12)$$

is very useful when it comes to obtaining smoothness estimates for the solution to the Stein equation. The equality ensures that we can restrict our attention to bounding the derivatives in the region $x \geq 0$, provided we obtain these bounds for both positive and negative β .

By direct calculations it is possible to bound the derivatives of the solution of the $VG_2(\nu, 1, \beta, 0)$ Variance-Gamma Stein equation (3.7). Gaunt [18] carried out these (rather lengthy) calculations for case $\beta = 0$, to obtain uniform bounds on the solution of the Stein equation (3.7) and its first four derivatives. By a change of variables it is then possible to establish smoothness estimates for the solution of the $VG_1(r, 0, \sigma, \mu)$ Stein equation (1.7).

Bounds on the first four derivatives of the solution of the $VG_1(r, 0, \sigma, \mu)$ Stein equation are sufficient for the limit theorems of Section 4. However, it would be desirable to extend these bounds to the general case of the $VG_1(r, \theta, \sigma, \mu)$ Stein equation. Another open problem is to obtain uniform bounds for the derivatives of all order for the solution of the $VG_1(r, 0, \sigma, \mu)$ Stein equation. This has been achieved for the normal and Gamma Stein equations (see the bounds of Goldstein and Rinott [22] and Luk [27]) using the generator approach. Gaunt [18] made some progress towards the goal of achieving such bounds via the generator approach, but the problem remains unsolved.

The following smoothness estimates for the solution of the $VG_2(\nu, 1, 0, 0)$ Stein equation were established by Gaunt [18].

Lemma 3.5. *Let $\nu > -1/2$ and suppose that $h \in C_b^3(\mathbb{R})$. Then the solution f , as given by (3.11), to the $\text{VG}_2(\nu, 1, 0, 0)$ Stein equation, and its first four derivatives are bounded as follows:*

$$\begin{aligned} \|f\| &\leq \left(\frac{1}{2\nu+1} + \frac{\pi\Gamma(\nu+1/2)}{2\Gamma(\nu+1)} \right) \|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\|, \\ \|f'\| &\leq \frac{2}{2\nu+1} \|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\|, \\ \|f''\| &\leq \left(\frac{\sqrt{\pi}}{2\sqrt{\nu+1/2}} + \frac{1}{2\nu+1} \right) \left[3\|h'\| + 4\|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\| \right], \\ \|f^{(3)}\| &\leq \left(\frac{\sqrt{\pi}}{2\sqrt{\nu+1/2}} + \frac{1}{2\nu+1} \right) \left[5\|h''\| + 18\|h'\| + 18\|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\| \right] + \frac{1}{v(\nu)} \|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\|, \\ \|f^{(4)}\| &\leq \left(\frac{\sqrt{\pi}}{2\sqrt{\nu+1/2}} + \frac{1}{2\nu+1} \right) \left[8\|h^{(3)}\| + 52\|h''\| + 123\|h'\| + 123\|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\| \right] \\ &\quad + \frac{1}{v(\nu)} \left[\|h'\| + \|h - \widetilde{\text{VG}}_{0,0}^{\nu,1}h\| \right], \end{aligned}$$

where $v(\nu)$ is given by

$$v(\nu) = \begin{cases} 2^{2\nu+1}\nu!(\nu+2)!(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}. \end{cases}$$

The bounds given in Lemma 3.5 are of order $\nu^{-1/2}$ as $\nu \rightarrow \infty$, except when 2ν is not equal to an integer, but is sufficiently close to an integer that

$$\sin(\pi\nu)2^{2\nu}\Gamma(\nu+1)\Gamma(\nu+4)\sqrt{\nu} = o(1).$$

Gaunt [18] remarked that the rogue $1/\sin(\pi\nu)$ term appeared to be an artefact of the analysis that was used to obtain the bounds.

It is also worth noting that the bounds of Lemma 3.5 break down as $\nu \rightarrow -1/2$. This is to be expected, because in this limit the $\text{VG}_2(\nu, 1, 0, 0)$ distribution approaches a point mass at the origin.

The bounds simplify in the case that $\nu \in \{0, 1/2, 1, 3/2, \dots\}$, and from these bounds we use a simple change of variables to obtain uniform bounds on the first derivatives of the solution of the $\text{VG}_1(r, 0, \sigma, \mu)$ Stein equation (1.7) for the case that r is a positive integer. These bounds are of order $r^{-1/2}$ as $r \rightarrow \infty$, which is the same order as Pickett's [33] bounds (1.4) for the solution of the $\Gamma(r, \lambda)$ Stein equation (1.3).

Theorem 3.6. *Let r be a positive integer and let $\sigma > 0$. Suppose that $h \in C_b^3(\mathbb{R})$, then the solution of the $\text{VG}_1(r, 0, \sigma, \mu)$ Stein equation (1.7) and its derivatives up to fourth order satisfy*

$$\|f^{(k)}\| \leq M_{r,\sigma}^k(h), \quad k = 0, 1, 2, 3, 4,$$

where

$$\begin{aligned} M_{r,\sigma}^0(h) &\leq \frac{1}{\sigma} \left(\frac{1}{r} + \frac{\pi\Gamma(r/2)}{2\Gamma(r/2+1/2)} \right) \|h - \text{VG}_{\sigma,\mu}^{r,0}h\|, \\ M_{r,\sigma}^1(h) &\leq \frac{2}{\sigma^2 r} \|h - \text{VG}_{\sigma,\mu}^{r,0}h\|, \\ M_{r,\sigma}^2(h) &\leq \frac{1}{\sigma^2} \left(\sqrt{\frac{\pi}{2r}} + \frac{1}{r} \right) \left[3\|h'\| + \frac{4}{\sigma} \|h - \text{VG}_{\sigma,\mu}^{r,0}h\| \right], \\ M_{r,\sigma}^3(h) &\leq \frac{1}{\sigma^2} \left(\sqrt{\frac{\pi}{2r}} + \frac{1}{r} \right) \left[5\|h''\| + \frac{18}{\sigma} \|h'\| + \frac{19}{\sigma^2} \|h - \text{VG}_{\sigma,\mu}^{r,0}h\| \right], \\ M_{r,\sigma}^4(h) &\leq \frac{1}{\sigma^2} \left(\sqrt{\frac{\pi}{2r}} + \frac{1}{r} \right) \left[8\|h^{(3)}\| + \frac{52}{\sigma} \|h''\| + \frac{124}{\sigma^2} \|h'\| + \frac{124}{\sigma^3} \|h - \text{VG}_{\sigma,\mu}^{r,0}h\| \right], \end{aligned}$$

and $f^{(0)} \equiv f$.

Proof. Let $g_{\tilde{h}}(x)$ denote the solution (3.11) to the $\text{VG}_2(\nu, 1, 0, 0)$ Stein equation (3.7)

$$xg''(x) + (2\nu + 1)g'(x) - xg(x) = \tilde{h}(x) - \widetilde{\text{VG}}_{0,0}^{\nu,1}\tilde{h}.$$

Then $f_h(x) = \frac{1}{\sigma}g_{\tilde{h}}(\frac{x-\mu}{\sigma})$ solves the $\text{VG}_1(r, 0, \sigma, \mu)$ Stein equation (1.7)

$$\sigma^2(x - \mu)f''(x) + \sigma^2rf'(x) - (x - \mu)f(x) = h(x) - \text{VG}_{\sigma,\mu}^{r,0}h,$$

where $r = 2\nu + 1$ and $h(x) = \tilde{h}(\frac{x-\mu}{\sigma})$, since $\text{VG}_{\sigma,\mu}^{r,0}h = \widetilde{\text{VG}}_{0,0}^{\nu,1}\tilde{h}$. That $\text{VG}_{\sigma,\mu}^{r,0}h = \widetilde{\text{VG}}_{0,0}^{\nu,1}\tilde{h}$ is verified by the following calculation:

$$\begin{aligned} \text{VG}_{\sigma,\mu}^{r,0}h &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{\pi}\Gamma(\frac{r}{2})} \left(\frac{|x-\mu|}{2\sigma}\right)^{\frac{r-1}{2}} K_{\frac{r-1}{2}}\left(\frac{|x-\mu|}{\sigma}\right) h(x) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}\Gamma(\nu + \frac{1}{2})} \left(\frac{|u|}{2}\right)^{\nu} K_{\nu}(|u|)\tilde{h}(u) du \\ &= \widetilde{\text{VG}}_{0,0}^{\nu,1}\tilde{h}, \end{aligned}$$

where we made the change of variables $u = \frac{x-\mu}{\sigma}$. We have that $\|f_h^{(k)}\| = \sigma^{-k-1}\|g_{\tilde{h}}^{(k)}\|$ for $k \in \mathbb{N}$, and $\|\tilde{h} - \widetilde{\text{VG}}_{0,0}^{\nu,1}\tilde{h}\| = \|h - \text{VG}_{\sigma,\mu}^{r,0}h\|$ and $\|\tilde{h}^{(k)}\| = \sigma^k\|h^{(k)}\|$ for $k \geq 1$, and the result now follows from the bounds of Lemma 3.5. \square

4 Limit theorems for Symmetric-Variance Gamma distributions

We now consider the Symmetric Variance-Gamma ($\theta = 0$) limit theorem that we discussed in the introduction. Let \mathbf{X} be a $m \times r$ matrix of independent and identically random variables X_{ik} with zero mean and unit variance. Similarly, we let \mathbf{Y} be a $n \times r$ matrix of independent and identically random variables Y_{jk} with zero mean and unit variance, where the Y_{jk} are independent of the X_{ik} . Then the statistic

$$W_r = \frac{1}{\sqrt{mn}} \sum_{i,j,k=1}^{m,n,r} X_{ik}Y_{jk}$$

is asymptotically $\text{VG}_1(r, 0, 1, 0)$ distributed, which can be seen by applying the central limit theorem and part (iv) of Proposition 1.2. Pickett [33] showed that the statistic $\frac{1}{m} \sum_{k=1}^d (\sum_{i=1}^m X_{ik})^2$, where the X_{ik} are independent and identically random variables with zero mean, unit variance and bounded eighth moment, converges to a $\chi_{(d)}^2$ random variable at a rate of order m^{-1} for smooth test functions. We now exhibit a proof which gives a bound for the rate of convergence of the statistic W_r to $\text{VG}_1(r, 0, 1, 0)$ random variables, under additional moment assumptions, which is shown to be of order $m^{-1} + n^{-1}$ for smooth test functions, using similar symmetry arguments to obtain this rate of convergence.

4.1 Local approach bounds for Symmetric Variance-Gamma distributions in the case $r = 1$

We first consider the case $r = 1$; the general r case follows easily as W_r is a linear sum of independent W_1 . For ease of reading, in the statement of the following theorem and in its proof we shall set $X_i \equiv X_{i1}$, $Y_j \equiv Y_{j1}$ and $W \equiv W_1$. Then we have the following:

Theorem 4.1. Suppose $X, X_1, \dots, X_m, Y, Y_1, \dots, Y_n$ are independent random variables with zero mean, unit variance and bounded sixth moment, with $X_i \stackrel{D}{=} X$ for all $i = 1, \dots, m$ and $Y_j \stackrel{D}{=} Y$ for all $j = 1, \dots, n$. Let $W = \frac{1}{\sqrt{mn}} \sum_{i,j=1}^{m,n} X_i Y_j$. Then, for $h \in C_b^3(\mathbb{R})$, we have

$$|\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h| \leq \gamma_1(X, Y)M_1^2(h) + \gamma_2(X, Y)M_1^3(h) + \gamma_3(X, Y)M_1^4(h), \quad (4.1)$$

where the $M_1^k(h)$ are defined as in Theorem 3.6, $\text{VG}_{1,0}^{1,0}h$ denotes the expectation of $h(Z)$ for $Z \sim \text{VG}_1(1, 0, 1, 0)$, and

$$\begin{aligned} \gamma_{m,n}^1(X, Y) &= \frac{10}{n}|\mathbb{E}Y^3|\mathbb{E}|Y^3| + \frac{11}{\sqrt{mn}}|\mathbb{E}X^3|\mathbb{E}Y^4, \\ \gamma_{m,n}^2(X, Y) &= \frac{9}{m}\mathbb{E}X^4\mathbb{E}Y^4 + \frac{30}{n}|\mathbb{E}Y^3|\mathbb{E}Y^4 + \frac{85}{\sqrt{mn}}|\mathbb{E}X^3|\mathbb{E}|Y^5| + \frac{46}{\sqrt{mn}}\mathbb{E}|X^3||\mathbb{E}Y^3|\mathbb{E}Y^4, \\ \gamma_{m,n}^3(X, Y) &= \frac{1}{n}\mathbb{E}X^4\mathbb{E}Y^4(1 + 15|\mathbb{E}Y^3|) + \frac{284}{m}|\mathbb{E}X^3|\mathbb{E}|X^3|\mathbb{E}Y^6 + \frac{148}{n}\mathbb{E}X^4|\mathbb{E}Y^3|\mathbb{E}|Y^5| \\ &\quad + \frac{135}{\sqrt{mn}}|\mathbb{E}X^3|\mathbb{E}X^4\mathbb{E}|Y^3| + \frac{248}{\sqrt{mn}}\mathbb{E}X^4|\mathbb{E}Y^3|. \end{aligned}$$

Remark 4.2. Notice that the statistic $W = \frac{1}{\sqrt{mn}} \sum_{i,j=1}^{m,n} X_i Y_j$ is symmetric in m and n and the random variables X_i and Y_j , and yet the bound (4.1) of Theorem 4.1 is not symmetric in m and n and the moments of X and Y . This asymmetry is a consequence of the local couplings that we used to obtain the bound.

In practice, when applying Theorem 4.1, we would compute $\gamma_{m,n}^k(X, Y)$ for $k = 1, 2, 3$, and $\gamma_{n,m}^k(Y, X)$ for $k = 1, 2, 3$, which would yields two bounds for the quantity $\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h$. We would then take the minimum of these two bounds. We proceed in this manner when applying bound (4.1) to prove Theorem 4.12.

Before proving Theorem 4.1, we introduce some notation and preliminary lemmas. We define the standardised sum S and T by

$$S = \frac{1}{\sqrt{m}} \sum_{i=1}^m X_i \quad \text{and} \quad T = \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_j$$

and we have that $W = ST$. In our proof we shall make use of the sums

$$S_i = S - \frac{1}{\sqrt{m}}X_i \quad \text{and} \quad T_j = T - \frac{1}{\sqrt{n}}Y_j$$

which are independent of X_i and Y_j , respectively. We therefore have the following formulas

$$\begin{aligned} W - S_i T &= ST - S_i T = \frac{1}{\sqrt{m}}X_i T \\ W - ST_j &= ST - ST_j = \frac{1}{\sqrt{n}}Y_j S. \end{aligned} \quad (4.2)$$

In the proof of Theorem 4.1 we use the following lemma, which can be found in Pickett [33], Lemma 4.3.

Lemma 4.3. Let X, X_1, \dots, X_m be a collection of independent and identically distributed random variables with mean zero and unit variance. Then, $\mathbb{E}S^p = O(1)$ for all $p \geq 1$. Specifically,

$$\begin{aligned} \mathbb{E}S^2 &= 1, \\ \mathbb{E}S^4 &= \frac{1}{m}[3(m-1) + \mathbb{E}X^4] < 3 + \frac{\mathbb{E}X^4}{m}, \end{aligned}$$

$$\begin{aligned} \mathbb{E}S^6 &= \frac{1}{m^2}[15(m-1)(m-2) + 10(m-1)(\mathbb{E}X^3)^2 + 15(m-1)\mathbb{E}X^4 + \mathbb{E}X^6] \\ &< 15 + \frac{10(\mathbb{E}X^3)^2}{m} + \frac{15\mathbb{E}X^4}{m} + \frac{\mathbb{E}X^6}{m^2}. \end{aligned}$$

and $\mathbb{E}|S| \leq (\mathbb{E}S^2)^{1/2}$, $\mathbb{E}|S^3| \leq (\mathbb{E}S^4)^{3/4}$, $\mathbb{E}|S^5| \leq (\mathbb{E}S^6)^{5/6}$, by Hölder's inequality.

We will also use the following lemma.

Lemma 4.4. *Suppose $p \geq 1$, then $\mathbb{E}|S_i|^p \leq \mathbb{E}|S|^p$.*

Proof. Applying Jensen's inequality gives

$$\mathbb{E}|S|^p = \mathbb{E}(\mathbb{E}(|S_i + n^{-1/2}X_i|^p | S_i)) \geq \mathbb{E}|\mathbb{E}(S_i + n^{-1/2}X_i | S_i)|^p = \mathbb{E}|S_i|^p,$$

as required. □

Using the $\text{VG}_1(1, 0, 1, 0)$ Stein equation (1.9) (with $r = 1$), we require a bound on the expression $\mathbb{E}[Wf''(W) + f'(W) - Wf(W)]$. We split the proof into two parts. In the first part of the proof, we use local couplings and Taylor expansions to bound $\mathbb{E}[Wf''(W) + f'(W) - Wf(W)]$ by the remainder terms that result from our Taylor expansions. Most of these terms are shown to be of the desired order of $O(m^{-1} + n^{-1})$, but the bounding of some of the terms is more involved. The second part of the proof is devoted to bounding these terms to the required order.

4.1.1 Proof Part I: Expansions and Bounding

Due to the independence of the X_i and Y_j variables, we are in the realms of the local approach coupling. We Taylor expand $f(W)$ about S_iT to obtain

$$\begin{aligned} &\mathbb{E}[Wf''(W) + f'(W) - Wf(W)] \\ &= \mathbb{E}STf''(W) + \mathbb{E}f'(W) - \frac{1}{\sqrt{m}} \sum_{i=1}^m \mathbb{E}X_iT \left(f(S_iT) + (ST - S_iT)f'(S_iT) \right. \\ &\quad \left. + \frac{1}{2}(ST - S_iT)^2 f''(S_iT) + \frac{1}{6}(ST - S_iT)^3 f^{(3)}(S_i^{[1]}T) \right), \end{aligned}$$

where $S_i^{[1]} = S_i + \theta_1(S - S_i)$ for some $\theta_1 \in (0, 1)$. Later in the proof we shall write $T_j^{[q]} = T_j + \theta_q(T - T_j)$, where $\theta_q \in (0, 1)$. Using independence and the fact that $\mathbb{E}X_i = 0$, we have

$$\sum_{i=1}^m \mathbb{E}X_iTf(S_iT) = \sum_{i=1}^m \mathbb{E}X_i\mathbb{E}Tf(S_iT) = 0.$$

As $ST - S_iT = \frac{1}{\sqrt{m}}X_iT$, we obtain

$$\mathbb{E}\{Wf''(W) + f'(W) - Wf(W)\} = N_1 + R_1 + R_2,$$

where

$$\begin{aligned} N_1 &= \mathbb{E}STf''(W) + \mathbb{E}f'(W) - \frac{1}{m} \sum_{i=1}^m \mathbb{E}X_i^2T^2 f'(S_iT), \\ R_1 &= -\frac{1}{2m^{3/2}} \sum_{i=1}^m \mathbb{E}X_i^3T^3 f''(S_iT), \\ R_2 &= -\frac{1}{6m^2} \sum_{i=1}^m \mathbb{E}X_i^4T^4 f^{(3)}(S_i^{[1]}T). \end{aligned}$$

We begin by bounding R_1 and R_2 . Taylor expanding $f''(S_iT)$ about W and using (4.2) gives

$$\begin{aligned} |R_1| &= \frac{|\mathbb{E}X^3|}{2m^{3/2}} \left| \sum_{i=1}^m \mathbb{E}T^3 f''(S_iT) \right| \\ &= \frac{|\mathbb{E}X^3|}{2m^{3/2}} \left| \sum_{i=1}^m \mathbb{E}T^3 f''(W) - \frac{1}{\sqrt{m}} \sum_{i=1}^m \mathbb{E}X_i T^4 f^{(3)}(S_i^{[2]}T) \right| \\ &\leq \frac{|\mathbb{E}X^3|}{2\sqrt{m}} |\mathbb{E}T^3 f''(W)| + \frac{\|f^{(3)}\| |\mathbb{E}X^3|}{2m} \left(3 + \frac{\mathbb{E}Y^4}{n} \right), \end{aligned}$$

where we used that the random variables X, X_1, \dots, X_m are identically distributed. In obtaining the last inequality we used that $\mathbb{E}T^4 < 3 + \frac{\mathbb{E}Y^4}{n}$ and that $\mathbb{E}|X_i| \leq \sqrt{\mathbb{E}X_i^2} = 1$. Bounding the term $\frac{1}{\sqrt{m}} |\mathbb{E}T^3 f''(W)|$ to the desired order of $O(m^{-1} + n^{-1})$ is somewhat involved and is deferred until the part II of the proof.

The bound for R_2 is immediate. We have

$$|R_2| \leq \frac{\|f^{(3)}\|}{6m^2} \sum_{i=1}^m \mathbb{E}X^4 \mathbb{E}T^4 \leq \frac{\|f^{(3)}\|}{6m} \mathbb{E}X^4 \left(3 + \frac{\mathbb{E}Y^4}{n} \right).$$

We now consider N_1 . We use independence and that $\mathbb{E}X_i^2 = 1$ and then Taylor expand $f'(S_iT)$ about W to obtain

$$\begin{aligned} \frac{1}{m} \sum_{i=1}^m \mathbb{E}X_i^2 T^2 f'(S_iT) &= \mathbb{E}T^2 f'(W) - \frac{1}{m^{3/2}} \sum_{i=1}^m \mathbb{E}X_i T^3 f''(W) \\ &\quad - \frac{1}{2m^2} \sum_{i=1}^m \mathbb{E}X_i^2 T^4 f^{(3)}(S_i^{[3]}T). \end{aligned}$$

Taylor expanding $f''(W)$ about S_iT gives

$$\begin{aligned} \frac{1}{m^{3/2}} \sum_{i=1}^m \mathbb{E}X_i T^3 f''(W) &= \frac{1}{m^{3/2}} \sum_{i=1}^m \mathbb{E}X_i T^3 f''(S_iT) + \frac{1}{m^2} \sum_{i=1}^m \mathbb{E}X_i^2 T^4 f^{(3)}(S_i^{[4]}T) \\ &= \frac{1}{m^2} \sum_{i=1}^m \mathbb{E}X_i^2 T^4 f^{(3)}(S_i^{[4]}T), \end{aligned}$$

where we used independence and that the X_i have zero mean to obtain the final inequality. Putting this together we have that

$$N_1 = \mathbb{E}ST f''(W) + \mathbb{E}f'(W) - \mathbb{E}T^2 f'(W) + R_3,$$

where

$$|R_3| \leq \frac{1}{2m^2} \left| \sum_{i=1}^m \mathbb{E}X_i^2 T^4 f^{(3)}(S_i^{[3]}T) \right| + \frac{1}{m^2} \left| \sum_{i=1}^m \mathbb{E}X_i^2 T^4 f^{(3)}(S_i^{[4]}T) \right| \leq \frac{3\|f^{(3)}\|}{2m} \left(3 + \frac{\mathbb{E}Y^4}{n} \right).$$

Noting that $T^2 = \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_j T = \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_j (\frac{1}{\sqrt{n}} Y_j + T_j)$, we may write N_1 as

$$N_1 = N_2 + R_3 + R_4,$$

where

$$\begin{aligned} N_2 &= \mathbb{E}ST f''(W) - \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}Y_j T_j f'(W), \\ R_4 &= \mathbb{E}f'(W) - \frac{1}{n} \sum_{j=1}^n \mathbb{E}Y_j^2 f'(W). \end{aligned}$$

We first consider R_4 . Taylor expanding $f'(W)$ about ST_j and using that $ST - ST_j = \frac{1}{\sqrt{n}}Y_jS$ gives

$$\begin{aligned} R_4 &= \frac{1}{n} \sum_{j=1}^n \mathbb{E}(1 - Y_j^2)f'(W) \\ &= \frac{1}{n} \sum_{j=1}^n \mathbb{E}(1 - Y_j^2) \left(f'(ST_j) + \frac{1}{\sqrt{n}}Y_jSf''(ST_j) + \frac{1}{2n}Y_j^2S^2f^{(3)}(ST_j^{[5]}) \right) \\ &= -\frac{\mathbb{E}Y^3}{n^{3/2}} \sum_{j=1}^n \mathbb{E}Sf''(ST_j) + \frac{1}{2n^2} \sum_{j=1}^n \mathbb{E}(Y_j^2 - Y_j^4)S^2f^{(3)}(ST_j^{[5]}), \end{aligned}$$

where we used independence and that $\mathbb{E}Y_j = 0$ and $\mathbb{E}Y_j^2 = 1$ to obtain the final equality. Taylor expanding $f''(ST_j)$ about W gives

$$\frac{\mathbb{E}Y^3}{n^{3/2}} \sum_{j=1}^n \mathbb{E}Sf''(ST_j) = \frac{\mathbb{E}Y^3}{\sqrt{n}} \mathbb{E}Sf''(W) - \frac{\mathbb{E}Y^3}{n^2} \sum_{j=1}^n \mathbb{E}Y_jS^2f^{(3)}(ST_j^{[6]}).$$

Putting this together we have the following bound for R_4 :

$$|R_4| \leq \frac{|\mathbb{E}Y^3|}{\sqrt{n}} |\mathbb{E}Sf''(W)| + \frac{\|f^{(3)}\|}{2n} (1 + 2|\mathbb{E}Y^3| + \mathbb{E}Y^4).$$

As was the case with the term $\frac{1}{\sqrt{n}}|\mathbb{E}T^3f''(W)|$, bounding the quantity $\frac{1}{\sqrt{n}}|\mathbb{E}Sf''(W)|$ to the desired order of $O(m^{-1} + n^{-1})$ is somewhat involved and is deferred until part II of the proof.

We now consider N_2 . Taylor expanding $f'(W)$ about ST_j , then using independence and that $\mathbb{E}Y_j = 0$ and $\mathbb{E}Y_j^2 = 1$ gives

$$\begin{aligned} N_2 &= \mathbb{E}STf''(W) - \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}Y_jT_j \left(f'(ST_j) + \frac{1}{\sqrt{n}}Y_jSf''(ST_j) \right. \\ &\quad \left. + \frac{1}{2n}Y_j^2S^2f^{(3)}(ST_j) + \frac{1}{6n^{3/2}}Y_j^3S^3f^{(4)}(ST_j^{[7]}) \right) \\ &= R_5 + R_6 + R_7, \end{aligned}$$

where

$$\begin{aligned} R_5 &= \frac{\mathbb{E}Y^3}{2n^{3/2}} \sum_{j=1}^n \mathbb{E}S^2T_jf^{(3)}(ST_j), \\ |R_6| &= \frac{1}{6n^2} \left| \sum_{j=1}^n \mathbb{E}Y_j^4S^3T_jf^{(4)}(ST_j^{[7]}) \right| \leq \frac{\|f^{(4)}\|}{6n} \mathbb{E}Y^4 \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4}, \\ R_7 &= \mathbb{E}STf''(W) - \frac{1}{n} \sum_{j=1}^n \mathbb{E}ST_jf''(ST_j). \end{aligned}$$

Using independence and that the Y_j have zero mean and then Taylor expanding $f^{(3)}(ST_j)$

about W gives

$$\begin{aligned} |R_5| &= \frac{|\mathbb{E}Y^3|}{2n^{3/2}} \left| \sum_{j=1}^n \mathbb{E}S^2Tf^{(3)}(ST_j) \right| \\ &= \frac{|\mathbb{E}Y^3|}{2n^{3/2}} \left| \sum_{j=1}^n \mathbb{E} \left[S^2Tf^{(3)}(W) - \frac{1}{\sqrt{n}}Y_jS^3 \left(T_j + \frac{1}{\sqrt{n}}Y_j \right) f^{(4)}(ST_j^{[8]}) \right] \right| \\ &\leq \frac{|\mathbb{E}Y^3|}{2\sqrt{n}} |\mathbb{E}S^2Tf^{(3)}(W)| + \frac{\|f^{(4)}\| |\mathbb{E}Y^3|}{2n} \left(1 + \frac{1}{\sqrt{n}} \right) \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4}. \end{aligned}$$

The term $\frac{1}{\sqrt{n}}|\mathbb{E}S^2Tf^{(3)}(W)|$ is bounded to the required order of $O(m^{-1} + n^{-1})$ in part II of the proof.

To bound R_7 we Taylor expand $f''(W)$ about ST_j and use independence and that the Y_j have zero mean to obtain

$$\begin{aligned} |R_7| &= \frac{1}{n} \left| \sum_{j=1}^n \mathbb{E}ST[f''(ST) - f''(ST_j)] \right| \\ &= \frac{1}{n} \left| \sum_{j=1}^n \mathbb{E}S(T_j + Y_j) \left(\frac{1}{\sqrt{n}}Y_jSf^{(3)}(ST_j) + \frac{1}{2n}Y_j^2S^2f^{(4)}(ST_j^{[9]}) \right) \right| \\ &= \frac{1}{n} \left| \sum_{j=1}^n \mathbb{E} \left[\frac{1}{\sqrt{n}}Y_j^2S^2f^{(3)}(ST_j) + \frac{1}{2n}Y_j^2 \left(T_j + \frac{1}{\sqrt{n}}Y_j \right) S^3f^{(4)}(ST_j^{[9]}) \right] \right| \\ &\leq \frac{\|f^{(3)}\|}{n} + \frac{\|f^{(4)}\|}{2n} \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4} \left(1 + \frac{\mathbb{E}|Y^3|}{\sqrt{n}} \right). \end{aligned}$$

To summarise, at this stage we have shown that $|\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h| \leq \sum_{k=1}^7 |R_k|$. We have also bounded all terms to order $m^{-1} + n^{-1}$, except for the terms $\frac{1}{\sqrt{m}}|\mathbb{E}T^3f''(W)|$, $\frac{1}{\sqrt{n}}|\mathbb{E}Sf''(W)|$ and $\frac{1}{\sqrt{n}}|\mathbb{E}S^2Tf^{(3)}(W)|$. In part II of the proof we shall use symmetry arguments to bound these terms to the required order. But before doing so, we obtain a useful bound for $\mathbb{E}S^2Tf^{(3)}(W)$ that will ensure that our bound for $|\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h|$ will only involve bounds of the first four derivatives of the $\text{VG}_1(1, 0, 1, 0)$ Stein equation (1.7), and hence will only involve the supremum norm of the first three derivatives of the test function h . The bound is given in the following lemma, which is proved in Appendix A.

Lemma 4.5. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be four times differentiable, then*

$$\begin{aligned} |\mathbb{E}S^2Tf^{(3)}(W)| &\leq |\mathbb{E}S^2f''(W)| + |\mathbb{E}Sf''(W)| + \frac{\|f^{(3)}\|}{\sqrt{n}}(1 + \mathbb{E}|Y^3|) \\ &\quad + \frac{\|f^{(4)}\|}{2\sqrt{n}} \left(2 + \frac{2}{\sqrt{n}} + \mathbb{E}|Y^3| + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right) \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4}. \end{aligned}$$

4.1.2 Proof Part II: Symmetry Argument for Optimal Rate

We now obtain bounds for $\frac{1}{\sqrt{n}}\mathbb{E}Sf''(W)$, $\frac{1}{\sqrt{n}}\mathbb{E}S^2Tf''(W)$ and $\frac{1}{\sqrt{m}}\mathbb{E}T^3f''(W)$. To obtain the desired rate of convergence we shall use symmetry arguments, which are similar to those used in Section 4.1.2. of Pickett [33] to achieve the optimal rate of convergence for chi-square limit theorems.

We begin by considering the bivariate standard normal Stein equation (see, for example, Goldstein and Rinott [22]) with test functions $g_1(s, t) = sf''(st)$, $g_2(s, t) =$

$st^2 f''(st)$ and $g_3(s, t) = t^3 f''(st)$. The bivariate standard normal Stein equation with test function $g_k(s, t)$, $k = 1, 2, 3$, and solution ψ_k is given by

$$\frac{\partial^2 \psi_k}{\partial s^2}(s, t) + \frac{\partial^2 \psi_k}{\partial t^2}(s, t) - s \frac{\partial \psi_k}{\partial s}(s, t) - t \frac{\partial \psi_k}{\partial t}(s, t) = g_k(s, t) - \mathbb{E}g_k(Z_1, Z_2), \tag{4.3}$$

where Z_1 and Z_2 are independent standard normal random variables.

For large m and n we have $S \approx N(0, 1)$ and $T \approx N(0, 1)$, so we can apply the $O(m^{-1/2} + n^{-1/2})$ bivariate central limit convergence rate (see, for example, Reinert and Röllin [38]) to bound the quantities $|\mathbb{E}g_k(S, T) - \mathbb{E}g_k(Z_1, Z_2)|$, $k = 1, 2, 3$. However, as the test functions g_k are odd functions ($g_k(s, t) = -g_k(-s, -t)$ for all $s, t \in \mathbb{R}$), the following lemma ensures that $\mathbb{E}g_k(Z_1, Z_2) = 0$, meaning that it should be possible to bound the expectations $\mathbb{E}g_k(S, T)$ to order $m^{-1/2} + n^{-1/2}$, which would yield $O(m^{-1} + n^{-1})$ bounds for $\frac{1}{\sqrt{n}} \mathbb{E}Sf''(W)$, $\frac{1}{\sqrt{n}} \mathbb{E}ST^2 f''(W)$ and $\frac{1}{\sqrt{m}} \mathbb{E}T^3 f''(W)$.

Lemma 4.6. *Suppose $g(x, y) = -g(-x, -y)$, then $\mathbb{E}g(Z_1, Z_2) = 0$ where Z_1, Z_2 are independent standard normal random variables. In particular, if Z_1 and Z_2 are independent standard normal random variables, then $\mathbb{E}g_k(Z_1, Z_2) = 0$, for $k = 1, 2, 3$.*

Proof. Let $Z'_1 = -Z_1$ and $Z'_2 = -Z_2$. Then $Z'_1 \stackrel{\mathcal{D}}{=} Z_1$ and $Z'_2 \stackrel{\mathcal{D}}{=} Z_2$, so $\mathbb{E}g(Z_1, Z_2) = -\mathbb{E}g(Z'_1, Z'_2) = -\mathbb{E}g(Z_1, Z_2)$, and therefore $\mathbb{E}g(Z_1, Z_2) = 0$. \square

We now apply Lemma 4.6 and then perform Taylor expansions to bound the expectations $\mathbb{E}g_k(S, T)$. Providing that a solution ψ_k exists for the test function g_k , we have

$$\begin{aligned} \mathbb{E}g_k(S, T) &= \mathbb{E} \left\{ \frac{\partial^2 \psi_k}{\partial s^2}(S, T) + \frac{\partial^2 \psi_k}{\partial t^2}(S, T) - S \frac{\partial \psi_k}{\partial s}(S, T) - T \frac{\partial \psi_k}{\partial t}(S, T) \right\} \\ &= R_8^k + R_9^k + R_{10}^k + R_{11}^k, \end{aligned}$$

where

$$\begin{aligned} R_8^k &= \frac{1}{2m^{3/2}} \sum_{i=1}^m \mathbb{E}X_i^3 \frac{\partial^3 \psi_k}{\partial s^3} \left(S_i + \phi_1 \frac{X_i}{\sqrt{m}}, T \right), \\ R_9^k &= \frac{1}{2n^{3/2}} \sum_{j=1}^n \mathbb{E}Y_j^3 \frac{\partial^3 \psi_k}{\partial t^3} \left(S, T_j + \phi_2 \frac{Y_j}{\sqrt{n}} \right), \\ R_{10}^k &= \frac{1}{m^{3/2}} \sum_{i=1}^m \mathbb{E}X_i \frac{\partial^3 \psi_k}{\partial s^3} \left(S_i + \phi_3 \frac{X_i}{\sqrt{m}}, T \right), \\ R_{11}^k &= \frac{1}{n^{3/2}} \sum_{j=1}^n \mathbb{E}Y_j \frac{\partial^3 \psi_k}{\partial t^3} \left(S, T_j + \phi_4 \frac{Y_j}{\sqrt{n}} \right), \end{aligned}$$

with $\phi_1, \phi_2, \phi_3, \phi_4 \in (0, 1)$.

Before we bound the remainder terms, we need bounds for the third order partial derivatives of the solution ψ_k in terms of the derivatives of f . We achieve this task by using the following lemma, the proof of which is given in Appendix A. Before stating the lemma, we define the double factorial function. The double factorial of a positive integer n is given by

$$n!! = \begin{cases} 1 \cdot 3 \cdot 5 \cdots (n-2) \cdot n, & n > 0 \text{ odd,} \\ 2 \cdot 4 \cdot 6 \cdots (n-2) \cdot n, & n > 0 \text{ even,} \end{cases} \tag{4.4}$$

and we define $(-1)!! = 0!! = 1$ (Arfken [1], p.547).

Lemma 4.7. *Suppose that $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is four times differentiable and let $g(s, t) = s^{a+b} f''(st)$, where $a, b \in \mathbb{N}$. Then, the third order partial derivatives of the solution ψ to the standard bivariate normal Stein equation (4.3) with test function g are bounded as follows*

$$\begin{aligned} \left| \frac{\partial^3 \psi}{\partial s^3} \right| &\leq \frac{\pi}{4} \{ 2^{a+b} \|f^{(4)}\| (|s|^a + a!!)(|t|^{b+2} + (b+1)!!) \\ &\quad + a 2^{a+b-1} \|f^{(3)}\| (|s|^{a-1} + (a-1)!!)(|t|^{b+1} + b!!) \\ &\quad + a(a-1) 2^{a+b-4} \|f''\| (|s|^{a-2} + (a-2)!!)(|t|^b + (b-1)!!) \}, \end{aligned} \tag{4.5}$$

$$\begin{aligned} \left| \frac{\partial^3 \psi}{\partial t^3} \right| &\leq \frac{\pi}{4} \{ 2^{a+b} \|f^{(4)}\| (|s|^{a+2} + (a+1)!!)(|t|^b + b!!) \\ &\quad + b 2^{a+b-1} \|f^{(3)}\| (|s|^{a+1} + a!!)(|t|^{b-1} + (b-1)!!) \\ &\quad + b(b-1) 2^{a+b-4} \|f''\| (|s|^{a-2} + (a-1)!!)(|t|^{b-2} + (b-2)!!) \}. \end{aligned} \tag{4.6}$$

With these bounds it is straightforward to bound the remainder terms. The following lemma allows us to easily deduce bounds for the remainder terms R_8^k, R_9^k, R_{10}^k and R_{11}^k , $k = 1, 2, 3$.

Lemma 4.8. *Suppose that $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is four times differentiable, and let $g(s, t) = s^{a+b} f''(st)$, where $a, b \in \mathbb{N}$, then*

$$\begin{aligned} |R_8^k| &\leq \frac{\pi}{8\sqrt{m}} \left\{ 2^{a+b} \|f^{(4)}\| \left[2^{a-1} \left[\mathbb{E}|X|^3 \mathbb{E}|S|^a + \frac{\mathbb{E}|X|^{a+3}}{m^{a/2}} \right] + a!! \mathbb{E}|X|^3 \right] (\mathbb{E}|T|^{b+2} + (b+1)!!) \right. \\ &\quad + a 2^{a+b-1} \|f^{(3)}\| \left[2^{a-2} \left[\mathbb{E}|X|^3 \mathbb{E}|S|^{a-1} + \frac{\mathbb{E}|X|^{a+2}}{m^{(a-1)/2}} \right] + (a-1)!! \mathbb{E}|X|^3 \right] (\mathbb{E}|T|^{b+1} + b!!) \\ &\quad + a(a-1) 2^{a+b-4} \|f''\| \left[2^{a-3} \left[\mathbb{E}|X|^3 \mathbb{E}|S|^{a-2} + \frac{\mathbb{E}|X|^{a+1}}{m^{(a-2)/2}} \right] + (a-2)!! \mathbb{E}|X|^3 \right] \\ &\quad \left. \times (\mathbb{E}|T|^b + (b-1)!!) \right\}, \\ |R_9^k| &\leq \frac{\pi}{8\sqrt{n}} \left\{ 2^{a+b} \|f^{(4)}\| \left[2^{b-1} \left[\mathbb{E}|Y|^3 \mathbb{E}|T|^b + \frac{\mathbb{E}|Y|^{b+3}}{n^{b/2}} \right] + b!! \mathbb{E}|Y|^3 \right] (\mathbb{E}|S|^{a+2} + (a+1)!!) \right. \\ &\quad + b 2^{a+b-1} \|f^{(3)}\| \left[2^{b-2} \left[\mathbb{E}|Y|^3 \mathbb{E}|T|^{b-1} + \frac{\mathbb{E}|Y|^{b+2}}{n^{(b-1)/2}} \right] + (b-1)!! \mathbb{E}|Y|^3 \right] (\mathbb{E}|S|^{a+1} + a!!) \\ &\quad + b(b-1) 2^{a+b-4} \|f''\| \left[2^{b-3} \left[\mathbb{E}|Y|^3 \mathbb{E}|T|^{b-2} + \frac{\mathbb{E}|Y|^{b+1}}{n^{(b-2)/2}} \right] + (b-2)!! \mathbb{E}|Y|^3 \right] \\ &\quad \left. \times (\mathbb{E}|S|^a + (a-1)!!) \right\}. \end{aligned}$$

The bound for R_{10}^k is similar to the bound for R_8^k but with $\mathbb{E}X^p$ and $\mathbb{E}|X|^p$ replaced with $\mathbb{E}X^{p-2}$ and $\mathbb{E}|X|^{p-2}$ respectively. The bound for R_{11}^k is similar to the bound for R_9^k but with $\mathbb{E}Y^p$ and $\mathbb{E}|Y|^p$ replaced with $\mathbb{E}Y^{p-2}$ and $\mathbb{E}|Y|^{p-2}$ respectively.

Proof. We prove that the bound for R_8^k holds; the bound for R_9^k then follows by symmetry. We begin by defining $S_i^* = S_i + \frac{\phi_1}{\sqrt{m}} X_i$. We note the following simple bound for $|S_i^*|^p$, for $p \geq 1$:

$$|S_i^*|^p = \left| S_i + \frac{\phi_1}{\sqrt{m}} X_i \right|^p \leq 2^{p-1} \left(|S_i|^p + \frac{\phi_1^p}{m^{p/2}} |X_i|^p \right) \leq 2^{p-1} \left(|S_i|^p + \frac{|X_i|^p}{m^{p/2}} \right). \tag{4.7}$$

Using our bound (4.5) for the third order partial derivative of ψ with respect to s , we

have

$$\begin{aligned}
 |R_8^k| &= \frac{1}{2m^{3/2}} \left| \sum_{i=1}^m \mathbb{E} X_i^3 \frac{\partial^3 \psi}{\partial s^3}(S_i^*, T) \right| \\
 &\leq \frac{\pi}{8m^{3/2}} \sum_{i=1}^m \mathbb{E} \left| X_i^3 \left\{ 2^{a+b} \|f^{(4)}\| (|S_i^*|^a + a!!)(|T|^{b+2} + (b+1)!!) \right. \right. \\
 &\quad \left. \left. + a 2^{a+b-1} \|f^{(3)}\| (|S_i^*|^{a-1} + (a-1)!!)(|T|^{b+1} + b!!) \right. \right. \\
 &\quad \left. \left. + a(a-1) 2^{a+b-4} \|f''\| (|S_i^*|^{a-2} + (a-2)!!)(|T|^b + (b-1)!!) \right\} \right| \\
 &\leq \frac{\pi}{8m^{3/2}} \sum_{i=1}^m \mathbb{E} \left| X_i^3 \left\{ 2^{a+b} \|f^{(4)}\| \left[2^{a-1} \left[|S_i|^a + \frac{|X_i|^a}{m^{a/2}} \right] + a!! \right] (|T|^{b+2} + (b+1)!!) \right. \right. \\
 &\quad \left. \left. + a 2^{a+b-1} \|f^{(3)}\| \left[2^{a-2} \left[|S_i|^{a-1} + \frac{|X_i|^{a-1}}{m^{(a-1)/2}} \right] + a!! \right] (|T|^{b+1} + b!!) \right. \right. \\
 &\quad \left. \left. + a(a-1) 2^{a+b-4} \|f''\| \left[2^{a-3} \left[|S_i|^{a-2} + \frac{|X_i|^{a-2}}{m^{(a-2)/2}} \right] + a!! \right] (|T|^b + (b-1)!!) \right\} \right|,
 \end{aligned}$$

where we used (4.7) to obtain the final inequality. Applying the triangle inequality, that X_i and S_i are independent and that, by Lemma 4.4, $\mathbb{E}|S_i|^p \leq \mathbb{E}|S|^p$ gives the desired bound. The final statement of the lemma is clear. \square

We can bound R_8^k , R_9^k , R_{10}^k and R_{11}^k by using the bounds in Lemma 4.8. We illustrate the argument by bounding R_{11}^k . In this case we have $g_1(s, t) = sf''(st)$, that is $a = 1$ and $b = 0$. We have

$$\begin{aligned}
 |R_8^1| &\leq \frac{\pi}{8\sqrt{m}} \left\{ 2 \|f^{(4)}\| \left(2\mathbb{E}|X^3| + \frac{\mathbb{E}X^4}{\sqrt{m}} \right) (\mathbb{E}T^2 + 1!!) + \|f^{(3)}\| (2\mathbb{E}|X^3|)(\mathbb{E}|T| + 0!!) \right\} \\
 &= \frac{\pi}{2\sqrt{m}} \left\{ \|f^{(4)}\| \left(2\mathbb{E}|X^3| + \frac{\mathbb{E}X^4}{\sqrt{m}} \right) + \|f^{(3)}\| \mathbb{E}|X^3| \right\},
 \end{aligned}$$

where we used that $0!! = 1!! = 1$, and $\mathbb{E}|T| \leq \sqrt{\mathbb{E}T^2} = 1$ to obtain the second equality. Continuing in this manner gives the following bounds:

$$\begin{aligned}
 |R_9^1| &\leq \frac{\pi \|f^{(4)}\|}{2\sqrt{n}} \mathbb{E}|Y^3| \left[2 + \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4} \right], \\
 |R_8^2| &\leq \frac{\pi}{\sqrt{m}} \left\{ \|f^{(4)}\| \left(2\mathbb{E}|X^3| + \frac{\mathbb{E}X^4}{\sqrt{m}} \right) \left(6 + \frac{\mathbb{E}Y^4}{n} \right) + \|f^{(3)}\| \mathbb{E}|X^3| \left[2 + \left(3 + \frac{\mathbb{E}Y^4}{n} \right)^{3/4} \right] \right\}, \\
 |R_9^2| &\leq \frac{2\pi}{\sqrt{n}} \left\{ \|f^{(4)}\| \left(2\mathbb{E}|Y^3| + \frac{\mathbb{E}|Y^5|}{n} \right) \left(2 + \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4} \right) \right. \\
 &\quad \left. + \|f^{(3)}\| \left[2\mathbb{E}|Y^3| + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right] + \|f''\| \mathbb{E}|Y^3| \right\}, \\
 |R_8^3| &\leq \frac{2\pi \|f^{(4)}\|}{\sqrt{m}} \mathbb{E}|X^3| \left[8 + \left(15 + \frac{10(\mathbb{E}Y^3)^2}{n} + \frac{15\mathbb{E}Y^4}{n^2} + \frac{\mathbb{E}Y^6}{n^3} \right)^{5/6} \right], \\
 |R_9^3| &\leq \frac{\pi}{4\sqrt{n}} \left\{ 8 \|f^{(4)}\| \left[\mathbb{E}|Y^3| \left[3 + 4 \left(3 + \frac{\mathbb{E}Y^4}{m} \right)^{3/4} \right] + \frac{\mathbb{E}Y^6}{m^{3/2}} \right] \right. \\
 &\quad \left. + 24 \|f^{(3)}\| \left[2\mathbb{E}|Y^3| + \frac{\mathbb{E}|Y^5|}{n} \right] + 3 \|f''\| \left[2\mathbb{E}|Y^3| + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right] \right\}.
 \end{aligned}$$

The bound for R_{10}^k is similar to the bound for R_8^k but with $\mathbb{E}X^p$ and $\mathbb{E}|X^p|$ replaced with $\mathbb{E}X^{p-2}$ and $\mathbb{E}|X^{p-2}|$ respectively. The bound for R_{11}^k is similar to the bound for R_9^k but with $\mathbb{E}Y^p$ and $\mathbb{E}|Y^p|$ replaced with $\mathbb{E}Y^{p-2}$ and $\mathbb{E}|Y^{p-2}|$, respectively.

We have therefore been able to bound the terms $\frac{1}{\sqrt{n}}\mathbb{E}Sf''(W)$, $\frac{1}{\sqrt{n}}\mathbb{E}ST^2f''(W)$ and $\frac{1}{\sqrt{m}}\mathbb{E}T^3f''(W)$ to order $m^{-1/2} + n^{-1/2}$. It therefore follows that the remainder terms R_1, \dots, R_7 are of order $m^{-1} + n^{-1}$. We showed in part I of the proof that $|\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h| \leq \sum_{k=1}^7 |R_k|$, and so we have achieved the desired $O(m^{-1} + n^{-1})$ bound. We can now sum up the remainder terms to obtain the following bound:

$$|\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h| \leq \tilde{\gamma}_{m,n}^1(X, Y)M_1^2(h) + \tilde{\gamma}_{m,n}^2(X, Y)M_1^3(h) + \tilde{\gamma}_{m,n}^3(X, Y)M_1^4(h),$$

where

$$\begin{aligned} \tilde{\gamma}_{m,n}^1(X, Y) &= \frac{\pi|\mathbb{E}Y^3|}{n}(\mathbb{E}|Y^3| + 2) + \frac{3\pi|\mathbb{E}X^3|}{8\sqrt{mn}} \left[4 + 2\mathbb{E}|Y^3| + \frac{2}{\sqrt{n}} + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right], \\ \tilde{\gamma}_{m,n}^2(X, Y) &= \frac{1}{6m}(9 + 3|\mathbb{E}X^3| + \mathbb{E}X^4) \left(3 + \frac{\mathbb{E}Y^4}{n} \right) + \frac{\pi|\mathbb{E}Y^3|}{n} \left[4 + 2\mathbb{E}|Y^3| + \frac{2}{\sqrt{n}} + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right] \\ &\quad + \frac{|\mathbb{E}Y^3|}{2n}(1 + \mathbb{E}|Y^3|) + \frac{3\pi|\mathbb{E}X^3|}{\sqrt{mn}} \left[4 + 2\mathbb{E}|Y^3| + \frac{2|\mathbb{E}Y^3|}{n} + \frac{\mathbb{E}|Y^5|}{n} \right] \\ &\quad + \frac{\pi|\mathbb{E}Y^3|}{\sqrt{mn}}(2 + \mathbb{E}|X^3|) \left(2 + \left(3 + \frac{\mathbb{E}Y^4}{n} \right)^{3/4} \right), \\ \tilde{\gamma}_{m,n}^3(X, Y) &= \frac{1}{24n} \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4} \left(4\mathbb{E}Y^4 + 6|\mathbb{E}Y^3| \left[4 + \mathbb{E}|Y^3| + \frac{4}{\sqrt{n}} + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right] \right) \\ &\quad + \frac{\pi|\mathbb{E}X^3|}{m}(\mathbb{E}|X^3| + 2) \left(8 + \left(15 + \frac{10(\mathbb{E}Y^3)^2}{n} + \frac{15\mathbb{E}Y^4}{n^2} + \frac{\mathbb{E}Y^6}{n^3} \right)^{5/6} \right) \\ &\quad + \frac{\pi|\mathbb{E}Y^3|}{4n} \left(2 + \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4} \right) \left(18 + \left(9 + \frac{8}{n} \right) \mathbb{E}Y^3 + \frac{4\mathbb{E}|Y^5|}{n} \right) \\ &\quad + \frac{\pi|\mathbb{E}X^3|}{\sqrt{mn}}(\mathbb{E}|Y^3| + 2) \left(3 + 4 \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4} \right) \\ &\quad + \frac{5\pi|\mathbb{E}Y^3|}{4\sqrt{mn}} \left(4 + 2\mathbb{E}|X^3| + \frac{2}{\sqrt{m}} + \frac{\mathbb{E}X^4}{\sqrt{m}} \right) \left(6 + \frac{\mathbb{E}Y^4}{n} \right). \end{aligned}$$

To complete the proof of Theorem 4.1 we simplify this bound by using that $m, n \geq 1$, as well as that for $a \geq b \geq 2$ we have $\mathbb{E}X^a \geq \mathbb{E}X^b \geq 1$, and then round all numbers up to the nearest integer. Doing so leads to bound (4.1), as required. \square

4.2 Extension to the case $r > 1$

For the case of $r > 1$, we have the following generalisation of Theorem 4.1:

Theorem 4.9. *Suppose the X_{ik} and Y_{jk} are defined as before, each with bounded sixth moment. Let $W_r = \frac{1}{\sqrt{mn}} \sum_{i,j,k=1}^{m,n,r} X_{ik}Y_{jk}$. Then, for any positive integer r and $h \in C_b^3(\mathbb{R})$, we have*

$$|\mathbb{E}h(W_r) - \text{VG}_{1,0}^{r,0}h| \leq r(\gamma_1(X, Y)M_{r,1}^2(h) + \gamma_2(X, Y)M_{r,1}^3(h) + \gamma_3(X, Y)M_{r,1}^4(h)), \quad (4.8)$$

where the $M_{r,1}^i(h)$ are defined as in Theorem 3.6, $\text{VG}_{1,0}^{r,0}h$ denotes the expectation of $h(Z)$ for $Z \sim \text{VG}(r, 0, 1, 0)$, and the γ_i are as in Theorem 4.1.

Proof. Define $W_{(k)} = \frac{1}{\sqrt{mn}} \sum_{i,j=1}^{m,n} X_{ik}Y_{jk}$, so that $W_r = \sum_{k=1}^r W_{(k)}$. Using the $\text{VG}_1(r, 0, 1, 0)$ Stein equation (1.9) we have

$$\begin{aligned} \mathbb{E}h(W_r) - \text{VG}_{1,0}^{r,0}h &= \mathbb{E}\{W_r f''(W_r) + r f'(W_r) - W_r f(W_r)\} \\ &= \sum_{k=1}^r \mathbb{E}[W_{(k)} f''(W_r) + f'(W_r) - W_{(k)} f(W_r)] \end{aligned}$$

$$= \sum_{k=1}^r \mathbb{E} \left[\mathbb{E} [W_{(k)} f''(W_r) + f'(W_r) - W_{(k)} f(W_r) \mid W_{(1)}, \dots, W_{(k-1)}, W_{(k+1)}, \dots, W_{(r)}] \right].$$

Since $\|g^{(n)}(x+c)\| = \|g^{(n)}(x)\|$ for any constant c , we may use bound (4.1) from Theorem 4.1 and the bounds of Theorem 3.6 for the derivatives of the solution of the $VG_1(r, 0, 1, 0)$ Stein equation to bound the above expression, which yields (4.8). \square

Remark 4.10. The terms $M_{r,1}^k(h)$, for $k = 2, 3, 4$, are of order $r^{-1/2}$ as $r \rightarrow \infty$ (recall Theorem 3.6), and therefore the bound of Theorem 4.9 is of order $r^{1/2}(m^{-1} + n^{-1})$. This is in agreement with bound of Theorem 4.7 of Pickett [33] for chi-square approximation, which is of order $d^{1/2}m^{-1}$.

Remark 4.11. The premise that the test function must be smooth is vital, as a non smooth test function will enforce a square-root convergence rate (cf. Berry-Essén theorem). Consider the following example in the case of a $VG_1(1, 0, 1, 0)$ random variable with test function $h \equiv \chi_{\{0\}}$. Let X_i , $i = 1, \dots, m = 2k$ and Y_j , $j = 1, \dots, n = 2l$, be random variables taking values in the set $\{-1, 1\}$ with equal probability. Then $\mathbb{E}X_i = \mathbb{E}Y_j = 0$, $\text{Var}X_i = \text{Var}Y_j = 1$ and

$$\begin{aligned} \mathbb{E}h(W) &= \mathbb{P} \left(\sum_{i,j} X_i Y_j = 0 \right) \\ &= \mathbb{P} \left(\sum_i X_i = 0 \right) + \mathbb{P} \left(\sum_j Y_j = 0 \right) - \mathbb{P} \left(\sum_i X_i = 0 \right) \mathbb{P} \left(\sum_j Y_j = 0 \right) \\ &= \binom{2k}{k} \left(\frac{1}{2} \right)^{2k} + \binom{2l}{l} \left(\frac{1}{2} \right)^{2l} - \binom{2k}{k} \left(\frac{1}{2} \right)^{2k} \binom{2l}{l} \left(\frac{1}{2} \right)^{2l} \\ &\approx \frac{1}{\sqrt{\pi k}} + \frac{1}{\sqrt{\pi l}} - \frac{1}{\pi kl} = \sqrt{\frac{2}{\pi m}} + \sqrt{\frac{2}{\pi n}} - \frac{4}{\pi mn}, \end{aligned}$$

by Stirling's approximation. Furthermore, $VG_{1,0}^{1,0}h = \mathbb{P}(VG(1, 0, 1, 0) = 0) = 0$, and hence the univariate bound (4.1) fails.

4.3 Application: Binary Sequence Comparison

We now consider a straightforward application of Theorem 4.1 to binary sequence comparison. This example is a simple special case of a more general problem of word sequence comparison, which is of particular importance to biological sequence comparisons. One way of comparing the sequences uses k -tuples (a sequence of letters of length k). If two sequences are closely related, we would expect the k -tuple content of both sequences to be very similar. A statistic for sequence comparison based on k -tuple content, known as the D_2 statistic was suggested by Blaisdell [8] (for other statistics based on k -tuple content see Reinert et al. [39]). Letting \mathcal{A} denote an alphabet of size d , and X_w and Y_w the number of occurrences of the word $w \in \mathcal{A}^k$ in the first and second sequences, respectively, then the D_2 statistic is defined by

$$D_2 = \sum_{w \in \mathcal{A}^k} X_w Y_w.$$

Due to the complicated dependence structure at both the local and global level (for a detailed account of the dependence structure see Reinert et al. [37]) 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; see Lippert et al. [26] for a detailed account of the limiting distributions of D_2 for different parameter values.

We now consider the case of an alphabet of size 2 with comparison based on the content of 1-tuples. We suppose that the sequences are of length m and n . We assume that 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 , respectively, then

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

Clearly, X and Y are independent binomial variables with expectation $\frac{m}{2}$ and $\frac{n}{2}$ respectively. Since $\mathbb{E}X^2 = \frac{m(m+1)}{4}$, it is easy to compute the mean and variance of D_2 , which are given by

$$\mathbb{E}D_2 = \frac{mn}{2} \quad \text{and} \quad \text{Var}D_2 = \frac{mn}{4}.$$

We now consider the standardised D_2 statistic,

$$\begin{aligned} W &= \frac{D_2 - \mathbb{E}D_2}{\sqrt{\text{Var}D_2}} && (4.9) \\ &= \frac{2}{\sqrt{mn}} \left(XY + (m - X)(n - Y) - \frac{mn}{2} \right) \\ &= \frac{2}{\sqrt{mn}} \left(2XY - mX - nY + \frac{mn}{2} \right) \\ &= \left(\frac{X - \frac{m}{2}}{\sqrt{\frac{m}{4}}} \right) \left(\frac{Y - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right). \end{aligned}$$

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. Therefore W has an approximate $\text{VG}_1(1, 0, 1, 0)$ distribution. We now apply Theorem 4.1 to obtain a bound on the error, in a weak convergence setting, in approximating the standardised D_2 statistic by its limiting $\text{VG}_1(1, 0, 1, 0)$ distribution.

Theorem 4.12. *For uniform i.i.d. binary sequences of lengths m and n , the standardised D_2 statistic W , defined as in equation (4.9), based on 1-tuple content is approximately $\text{VG}_1(1, 0, 1, 0)$ distributed. Moreover, for $h \in C_b^3(\mathbb{R})$ the following bound on the error in approximating W by its limiting distribution holds,*

$$|\mathbb{E}h(W) - \text{VG}_{1,0}^{1,0}h| \leq \min\{A, B\}, \tag{4.10}$$

where

$$A = \frac{9}{m}M_{1,1}^3(h) + \frac{1}{n}M_{1,1}^4(h) \quad \text{and} \quad B = \frac{9}{n}M_{1,1}^3(h) + \frac{1}{m}M_{1,1}^4(h),$$

where the $M_{1,1}^k(h)$ is defined as in Theorem 3.6, and $\text{VG}_{1,0}^{1,0}h$ denotes the expectation of $h(Z)$, for $Z \sim \text{VG}_1(1, 0, 1, 0)$.

Proof. We can write the number of occurrences that letter 0 occurs in the first sequence as $X = \sum_{i=1}^m \mathbb{I}_i$, where \mathbb{I}_i is the indicator random variable that letter 0 appears at position i in the first sequence. Similarly, the number of occurrences of letter 0 in the second sequence is given by $Y = \sum_{j=1}^n \mathbb{J}_j$, where \mathbb{J}_j is the indicator random variable that letter 0 appears at position j in the second sequence. The standardised D_2 statistic W may therefore be write as

$$W = \frac{D_2 - \mathbb{E}D_2}{\sqrt{\text{Var}D_2}} = \left(\frac{X - \frac{m}{2}}{\sqrt{\frac{m}{4}}} \right) \left(\frac{Y - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right) = \frac{1}{\sqrt{mn}} \sum_{i,j=1}^{m,n} X_i Y_j,$$

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 and have zero mean and unit variance. We may therefore apply Theorem 4.1 with $\mathbb{E}X_i^3 = \mathbb{E}Y_j^3 = 0$ and $\mathbb{E}X_i^4 = \mathbb{E}Y_j^4 = 1$ to obtain bound (4.10). \square

References

- [1] ARFKEN, G. *Mathematical Methods for Physicists*, 3rd ed. Orlando, FL: Academic Press, 1985. MR-0205512
- [2] BARBOUR, A. D. Stein's method and Poisson process convergence. *J. Appl. Probab.* **25A** (1988), pp. 175–184. MR-0974580
- [3] BARBOUR, A. D. Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** (1990), pp. 297–322. MR-1035659
- [4] BARNDORFF-NIELSEN, O. E. and HALGREEN, C. Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **38** (1977), pp. 309–312. MR-0436260
- [5] BARNDORFF-NIELSEN, O. E., KENT, J. and SØRENSEN, M. Normal Variance-Mean Mixtures and z Distributions. *International Statistical Review* **50** (1982), pp. 145–159. MR-0678296
- [6] BERRY, A. C. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** (1941), pp. 126–136. MR-0003498
- [7] BIBBY, B. M., and SØRENSEN, M. Hyperbolic Processes in Finance. In S. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance* (2003), pp. 211–248. Amsterdam: Elsevier Science.
- [8] BLAISDELL, B. A measure of the similarity of sets of sequences not requiring sequence alignment. *Proc. Natl. Acad. Sci. USA* **83** (1986), pp. 5155–5159.
- [9] CHATTERJEE, S., FULMAN, J. and RÖLLIN, A. Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** (2011), pp. 197–223. MR-2802856
- [10] CHEN, L. H. Y. Poisson approximation for dependent trials. *Ann. Probab.* **3** (1975), pp. 534–545. MR-0428387
- [11] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q-M. *Normal Approximation By Stein's Method*. Springer, 2011. MR-2732624
- [12] COLLINS, P. J. *Differential and Integral Equations*. Oxford University Press, 2006.
- [13] DÖBLER, C. Distributional transformations without orthogonality relations. arXiv:1312.6093, 2013.
- [14] DURRETT, R. *Stochastic Calculus: a Practical Introduction*. CRC Press, 1996. MR-1398879
- [15] EBERLEIN, E. and HAMMERSTEIN E. Generalised Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes. in: R.C. Dalang, M. Dozzi, F. Russo (Eds.), Seminar on Stochastic Analysis, Random Fields and Applications IV, in: *Progress in Probability* **58** Birkhäuser Verlag, (2004), pp. 105–153.
- [16] ESSÉEN, C-G. Fourier analysis of distribution functions, a mathematical study of the Laplace-Gaussian law. *Acta Math.* **77** (1945), pp. 1–125. MR-0014626
- [17] FINLAY, R. and SENETA, E. Option pricing with VG-like models, *International Journal of Theoretical and Applied Finance.* **11**, (2008), pp. 943–955. MR-2492088
- [18] GAUNT, R. E. *Rates of Convergence of Variance-Gamma Approximations via Stein's Method*. DPhil thesis, University of Oxford, 2013.
- [19] GAUNT, R. E. Uniform bounds for expressions involving modified Bessel functions. arXiv:1309.4309, 2013.
- [20] GAUNT, R. E. On Stein's Method for products of normal, gamma and beta random variables and a generalisation of the zero bias coupling. Preprint 2014.
- [21] GOLDSTEIN, L. and REINERT, G. Stein's Method and the zero bias transformation with application to simple random sampling *Ann. Appl. Probab.* **7** (1997), pp. 935–952. MR-1484792
- [22] GOLDSTEIN, L. and RINOTT, Y. Multivariate normal approximations by Stein's method and size bias couplings *J. Appl. Probab.* **33** (1996), pp. 1–17.
- [23] HOLM, H. and ALOUINI, M-S. Sum and Difference of two squared correlated Nakagami variates with the McKay distribution. *IEEE Transactions on Communications.* **52** (2004), pp. 1367-1376.

- [24] LEY, C. and SWAN, Y. A unified approach to Stein characterizations. arXiv:1105.4925, 2011.
- [25] LINETSKY, V. The spectral representation of Bessel processes with drift: Applications in queueing and finance. *J. Appl. Probab.* **41** (2004) pp. 327–344. MR-2052575
- [26] LIPPERT, R. A., HUANG, H. and WATERMAN, M. S. Distributional regimes for the number of k-word matches between two random sequences. *Proc. Natl Acad. Sci. USA* **99** (2002), pp. 13980–13989. MR-1944413
- [27] LUK, H. *Stein's Method for the Gamma Distribution and Related Statistical Applications*. PhD thesis, University of Southern California, 1994. MR-2693204
- [28] MADAN, D. B. and SENETA, E. The Variance Gamma (V.G.) Model for Share Market Returns. *Journal of Business* **63** (1990), pp. 511–524.
- [29] NOURDIN, I. and PECCATI, G. Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** (2011) pp. 75–118. MR-2520122
- [30] OLVER, F. W. J., LOZIER, D. W., BOISVERT, R. F. and CLARK, C. W. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010. MR-2723248
- [31] PEKÖZ, E. and RÖLLIN, A. New rates for exponential approximation and the theorems of Rnyi and Yaglom. *Ann. Probab.* **39** (2011) pp. 587–608. MR-2789507
- [32] PEKÖZ, E., RÖLLIN, A. and ROSS, N. Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Prob.* **23**, pp. 1188–1218. MR-3076682
- [33] PICKETT, A. *Rates of Convergence of χ^2 Approximations via Stein's Method*. DPhil thesis, University of Oxford, 2004.
- [34] PIKE, J. and REN, H. Stein's method and the Laplace distribution. arXiv:1210.5775, 2012.
- [35] RAIČ, M. Normal approximation by Stein's method. In: *Proceedings of the 7th Young Statisticians Meeting* (2003), pp. 71–97.
- [36] REINERT, G. Three general approaches to Stein's method. In *An Introduction to Stein's Method. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **4** (2005), pp. 183–221. Singapore Univ. Press, Singapore.
- [37] REINERT, G., SCHBATH, S. and WATERMAN, M. S. in LOTHAIRE, M. *Applied Combinatorics on Words*. Cambridge University Press, 2005. MR-2165687
- [38] REINERT, G. and RÖLLIN, A. Multivariate Normal Approximations with Stein's Method of Exchangeable Pairs Under a General Linearity Condition. *Ann. Probab.* **37** (2009), pp. 2150–2173. MR-2573554
- [39] REINERT, G., CHEW, D., SUN, F. and WATERMAN, M. S. Alignment free sequence comparison (I): statistics and power. *J. Comput. Biol.* **16** (2009), pp. 1615–1634. MR-2578699
- [40] SCOTT, D. J., WÜRTZ, D. DONG, C. and TRAN T. T. Moments of the generalized hyperbolic distribution. *Computational Statistics* **26** (2011), pp. 459–476. MR-2833142
- [41] STEIN, C. A bound for the error in the normal approximation to the the distribution of a sum of dependent random variables. In *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* (1972), vol. 2, Univ. California Press, Berkeley, pp.583–602. MR-0402873
- [42] STEIN, C. *Approximate Computation of Expectations*. IMS, Hayward, California, 1986. MR-0882007
- [43] STEIN, C., DIACONIS, S, HOLMES, S. and REINERT, G. Use of exchangeable pairs in the analysis of simulations. In *Stein's method: expository lectures and applications*, volume 46 of IMS Lecture Notes Monogr. Ser., Inst. Math. Statist., Beachwood, OH (2004), pp. 1–26.
- [44] WINKELBAUER, A. Moments and absolute moments of the normal distribution, arXiv:1209.4340, 2012.

A Proofs from the text

Here we prove the lemmas that we stated in the main text without proof.

A.1 Proof of Proposition 1.2

For clarity, we restate the proposition.

Proposition A.1. (i) Let $\sigma > 0$ and $\mu \in \mathbb{R}$ and suppose that Z_r has the $\text{VG}_1(r, 0, \sigma/\sqrt{r}, \mu)$ distribution. Then Z_r converges in distribution to a $N(\mu, \sigma^2)$ random variable in the limit $r \rightarrow \infty$.

(ii) Let $\sigma > 0$ and $\mu \in \mathbb{R}$, then a $\text{VG}_1(2, 0, \sigma, \mu)$ random variable has the Laplace(μ, σ) distribution with probability density function

$$p_{\text{VG}_1}(x; 2, 0, \sigma, \mu) = \frac{1}{2\sigma} \exp\left(-\frac{|x - \mu|}{\sigma}\right), \quad x \in \mathbb{R}. \tag{A.1}$$

(iii) Suppose that (X, Y) has the bivariate normal distribution with correlation ρ and marginals $X \sim N(0, \sigma_X^2)$ and $Y \sim N(0, \sigma_Y^2)$. Then the product XY follows the $\text{VG}_1(1, \rho\sigma_X\sigma_Y, \sigma_X\sigma_Y\sqrt{1 - \rho^2}, 0)$ distribution.

(iv) Let X_1, \dots, X_r and Y_1, \dots, Y_r be independent standard normal random variables. Then $\mu + \sigma \sum_{k=1}^r X_k Y_k$ has the $\text{VG}_1(r, 0, \sigma, \mu)$ distribution. As a special case we have that a Laplace random variable with density (A.1) has the representation $\mu + \sigma(X_1 Y_1 + X_2 Y_2)$.

(v) The Gamma distribution is a limiting case of the Variance-Gamma distributions: for $r > 0$ and $\lambda > 0$, the random variable $X_\sigma \sim \text{VG}_1(2r, (2\lambda)^{-1}, \sigma, 0)$ converges in distribution to a $\Gamma(r, \lambda)$ random variable in the limit $\sigma \downarrow 0$.

(vi) Suppose that (X, Y) follows a bivariate gamma distribution with correlation ρ and marginals $X \sim \Gamma(r, \lambda_1)$ and $Y \sim \Gamma(r, \lambda_2)$. Then the random variable $X - Y$ has the $\text{VG}_1(2r, (2\lambda_1)^{-1} - (2\lambda_2)^{-1}, (\lambda_1 \lambda_2)^{-1/2}(1 - \rho)^{1/2}, 0)$ distribution.

Proof. (i) Let X_1, X_2, \dots and Y_1, Y_2, \dots be independent standard normal random variables, and define $Z_r = \mu + \frac{\sigma}{\sqrt{r}} \sum_{i=1}^r X_i Y_i$. Then by Corollary 2.5 it follows that $Z_r \sim \text{VG}_1(r, 0, \sigma/\sqrt{r}, \mu)$. Moreover, the products $X_i Y_i, i = 1, 2, \dots$, are independent and identically random variables with mean zero and unit variance, and by the central limit theorem $\frac{1}{\sqrt{r}} \sum_{i=1}^r X_i Y_i \xrightarrow{D} N(0, 1)$. Hence, $Z_r \xrightarrow{D} N(\mu, \sigma^2)$ as $r \rightarrow \infty$.

(ii) This follows by applying the formula $K_{\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}} e^{-x}$ to the density (1.5).

(iii) Let $\tilde{X} = \frac{X}{\sigma_X} \sim N(0, 1)$ and $\tilde{Y} = \frac{Y}{\sigma_Y} \sim N(0, 1)$, and define the random variable W by

$$W = \frac{1}{\sqrt{1 - \rho^2}} (\tilde{Y} - \rho \tilde{X}).$$

It is straightforward to show that $W \sim N(0, 1)$, and that W and \tilde{X} are jointly normally distributed with correlation 0. We can therefore express the product $Z = XY$ in terms of independent standard normal random variables \tilde{X} and W as follows

$$Z = XY = \sigma_X \sigma_Y \tilde{X} \tilde{Y} = \sigma_X \sigma_Y \tilde{X} (\sqrt{1 - \rho^2} W + \rho \tilde{X}) = \sigma_X \sigma_Y \sqrt{1 - \rho^2} \tilde{X} W + \rho \sigma_X \sigma_Y \tilde{X}^2.$$

Hence, by Corollary 2.5, we have that $Z \sim \text{VG}_1(1, \rho\sigma_X\sigma_Y, \sigma_X\sigma_Y\sqrt{1 - \rho^2}, 0)$.

(iv) Taking $\theta = 0$ in Corollary (2.5) leads to the general representation. The representation for the Laplace distribution now follows from part (ii).

(v) This follows on letting $\sigma \rightarrow 0$ in Proposition 2.4 and then using the fact that if $Y \sim \Gamma(\alpha, \beta)$ then $kY \sim \Gamma(\alpha, \beta/k)$.

(vi) Theorem 6 of Holm and Alouini [23] gives the following formula for the probability density function of $Z = U - V$:

$$p_Z(x) = \frac{|x|^{r-1/2}}{\Gamma(r)\sqrt{\pi}\sqrt{\beta_1\beta_2(1-\rho)}} \left(\frac{1}{(\beta_1 + \beta_2)^2 - 4\beta_1\beta_2\rho} \right)^{\frac{2r-1}{4}} \exp\left(\frac{x}{2(1-\rho)} \left(\frac{1}{\beta_1} - \frac{1}{\beta_2} \right) \right) \\ \times K_{r-\frac{1}{2}} \left(|x| \frac{\sqrt{(\beta_1 + \beta_2)^2 - 4\beta_1\beta_2\rho}}{2\beta_1\beta_2(1-\rho)} \right), \quad x \in \mathbb{R},$$

where

$$\beta_1 = \frac{1}{\lambda_1} \quad \text{and} \quad \beta_2 = \frac{1}{\lambda_2}.$$

We can write the density of Z as follows

$$p_Z(x) = \frac{|x|^{r-1/2}}{\Gamma(r)\sqrt{\pi}\sqrt{\beta_1\beta_2(1-\rho)}} \left(\frac{1}{(\beta_1 + \beta_2)^2 - 4\beta_1\beta_2\rho} \right)^{\frac{2r-1}{4}} \exp\left(x \cdot \frac{\frac{1}{2}(\beta_1 - \beta_2)}{\beta_1\beta_2(1-\rho)}\right) \times K_{r-\frac{1}{2}}\left(|x| \frac{\sqrt{[(\beta_1 - \beta_2)/2]^2 - \beta_1\beta_2(1-\rho)}}{\beta_1\beta_2(1-\rho)}\right), \quad x \in \mathbb{R}. \tag{A.2}$$

Comparing (A.2) with the Variance-Gamma density function (1.5), we see that Z has a $VG_1(2r, \theta, \sigma, 0)$ distribution, where θ and σ are given by

$$\begin{aligned} \theta &= \frac{\beta_1 - \beta_2}{2} = \frac{1}{2\lambda_1} - \frac{1}{2\lambda_2}, \\ \sigma &= \sqrt{\beta_1\beta_2(1-\rho)} = \sqrt{\frac{1-\rho}{\lambda_1\lambda_2}}, \end{aligned}$$

as required. □

A.2 Proof of Lemma 3.3

We begin by proving that there is at most one bounded solution to the Variance-Gamma Stein equation (3.7) when $\nu \geq 0$. Suppose u and v are solutions to the Stein equation that satisfy $\|u^{(k)}\|, \|v^{(k)}\| < \infty$. Define $w = u - v$. Then w satisfies $\|w^{(k)}\| = \|u^{(k)} - v^{(k)}\| \leq \|u^{(k)}\| + \|v^{(k)}\| < \infty$, and is a solution to the following differential equation

$$xw''(x) + (2\nu + 1 + 2\beta x)w'(x) + ((2\nu + 1)\beta - (1 - \beta^2)x)w(x) = 0.$$

This homogeneous differential equation has general solution

$$w(x) = Ae^{-\beta x}x^{-\nu}K_\nu(x) + Be^{-\beta x}x^{-\nu}I_\nu(x).$$

From the asymptotic formula (B.3) for $I_\nu(x)$, it follows that to have a bounded solution we must take $B = 0$. From the asymptotic formula (B.2) for $K_\nu(x)$, we see that $w(x)$ has a singularity at the origin if $\nu \geq 0$. Therefore if $\nu \geq 0$, then for $w(x)$ to be bounded we must take $A = 0$, and therefore $w = 0$ and so $u = v$.

We now use variation of parameters (see Collins [12]) to solve the Stein equation (3.7). The method allows us to solve differential equations of the form

$$v''(x) + p(x)v'(x) + q(x)v(x) = g(x).$$

Suppose $v_1(x)$ and $v_2(x)$ are linearly independent solutions of the homogeneous equation

$$v''(x) + p(x)v'(x) + q(x)v(x) = 0.$$

Then the general solution to the inhomogeneous equation is given by

$$v(x) = -v_1(x) \int_a^x \frac{v_2(t)g(t)}{W(t)} dt + v_2(x) \int_b^x \frac{v_1(t)g(t)}{W(t)} dt,$$

where a and b are arbitrary constants and $W(t) = W(v_1, v_2) = v_1v_2' - v_2v_1'$ is the Wronskian.

It is easy to verify that a pair of linearly independent solutions to the homogeneous equation

$$f''(x) + \left(\frac{(2\nu + 1)}{x} + 2\beta\right)f'(x) + \left(\frac{(2\nu + 1)\beta}{x} - (1 - \beta^2)\right)f(x) = 0$$

are $e^{-\beta x} x^{-\nu} K_\nu(x)$ and $e^{-\beta x} x^{-\nu} I_\nu(x)$. However, we take $f_1(x) = e^{-\beta x} |x|^{-\nu} K_\nu(|x|)$ and $f_2(x) = e^{-\beta x} |x|^{-\nu} I_\nu(|x|)$ as our linearly independent solutions to the homogeneous equation. It will become clear later why this is a more suitable basis of solutions to the homogeneous equation. We now show that f_1 and f_2 are indeed linearly independent solutions to the homogeneous equation. From (B.1) we have

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

Formula (B.6) states that $K_\nu(-x) = (-1)^\nu K_\nu(x) - \pi i I_\nu(x)$ and therefore

$$\frac{K_\nu(-x)}{(-x)^\nu} = \frac{K_\nu(x)}{x^\nu} - \frac{\pi i}{(-1)^\nu} \frac{I_\nu(x)}{x^\nu},$$

and so

$$\frac{e^{-\beta x} K_\nu(|x|)}{|x|^\nu} = \frac{e^{-\beta x} K_\nu(x)}{x^\nu} - \frac{\pi i}{(-1)^\nu} \frac{e^{-\beta x} I_\nu(x)}{x^\nu} \chi_{(-\infty, 0]}(x).$$

Since $e^{-\beta x} x^\nu I_\nu(x)$ is a solution to the homogeneous equation that is linearly independent of $e^{-\beta x} x^\nu K_\nu(x)$, it follows that $e^{-\beta x} |x|^\nu K_\nu(|x|)$ is a solution to the homogeneous equation.

From (B.9) and (B.8) we have

$$\frac{d}{dx} \left(\frac{K_\nu(|x|)}{|x|^\nu} \right) = -\frac{K_{\nu+1}(|x|)}{x|x|^{\nu-1}}, \quad \frac{d}{dx} \left(\frac{I_\nu(|x|)}{|x|^\nu} \right) = \frac{I_{\nu+1}(|x|)}{x|x|^{\nu-1}},$$

and therefore

$$W(x) = \frac{e^{-2\beta x} (I_\nu(|x|)K_{\nu+1}(|x|) + K_\nu(|x|)I_{\nu+1}(|x|))}{x|x|^{2\nu-1}} = \frac{e^{-2\beta x}}{x|x|^{2\nu}},$$

where we used (B.5) to obtain the equality in the above display. Therefore the general solution to the inhomogeneous equation is given by

$$f(x) = -\frac{e^{-\beta x} K_\nu(|x|)}{|x|^\nu} \int_a^x e^{\beta y} |y|^\nu I_\nu(|y|) [h(y) - \widetilde{VG}_{\beta,0}^{\nu,1} h] dy + \frac{e^{-\beta x} I_\nu(|x|)}{|x|^\nu} \int_b^x e^{\beta y} |y|^\nu K_\nu(|y|) [h(y) - \widetilde{VG}_{\beta,0}^{\nu,1} h] dy.$$

This solution is clearly bounded everywhere except possibly for $x = 0$ or in the limits $x \rightarrow \pm\infty$. We therefore choose a and b so that our solution is bounded at these points and thus for all real x . To ensure the solution is bounded at the origin we must take $a = 0$. We choose b so that the solution is bounded in the limits $x \rightarrow \pm\infty$. If we take $b = \infty$ then we obtain solution (3.11). Taking $b = -\infty$ would lead to the same solution (see Remark 3.4).

Solution (3.11) is a candidate bounded solution, and we now verify that this solution and its first derivative are indeed bounded for all $x \in \mathbb{R}$. Straightforward calculations show that, for $x \geq 0$,

$$\begin{aligned} \|f\| &\leq \|\tilde{h}\| \left| \frac{e^{-\beta x} K_\nu(x)}{x^\nu} \int_0^x e^{\beta y} y^\nu I_\nu(y) dy \right| + \|\tilde{h}\| \left| \frac{e^{-\beta x} I_\nu(x)}{x^\nu} \int_x^\infty e^{\beta y} y^\nu K_\nu(y) dy \right|, \\ \|f'\| &\leq \|\tilde{h}\| \left| \frac{d}{dx} \left(\frac{e^{-\beta x} K_\nu(x)}{x^\nu} \right) \int_0^x e^{\beta y} y^\nu I_\nu(y) dy \right| \\ &\quad + \|\tilde{h}\| \left| \frac{d}{dx} \left(\frac{e^{-\beta x} I_\nu(x)}{x^\nu} \right) \int_x^\infty e^{\beta y} y^\nu K_\nu(y) dy \right|, \end{aligned}$$

where $\tilde{h} = h(x) - \widetilde{VG}_{\beta,0}^{\nu,1}h$. From inequalities (B.11), (B.12) and (B.13) it follows that the expressions involving modified Bessel functions are bounded for all $x \geq 0$. Recalling Remark 3.4, it is sufficient to bound these expressions in the region $x \geq 0$ and then consider the case of both positive and negative β , and so we have shown the f and its first derivative are bounded for all $x \in \mathbb{R}$.

A.3 Proof of Lemma 4.5

Taylor expanding $f''(W)$ about ST_j gives

$$\begin{aligned} \mathbb{E}ST^2 f''(W) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}Y_j \left(T_j + \frac{1}{\sqrt{n}} Y_j \right) S f''(W) \\ &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \mathbb{E}Y_j \left(T_j + \frac{1}{\sqrt{n}} Y_j \right) S \left(f''(ST_j) \right. \\ &\quad \left. + \frac{1}{\sqrt{n}} Y_j S f^{(3)}(ST_j) + \frac{1}{2n} Y_j^2 S^2 f^{(4)}(ST_j^{[1]}) \right) \\ &= R_1 + N_1 + R_2, \end{aligned}$$

where

$$\begin{aligned} R_1 &= \frac{1}{n} \sum_{j=1}^n \mathbb{E}S f''(ST_j), \\ N_1 &= \frac{1}{n} \sum_{j=1}^n \mathbb{E}S^2 \left(Y_j^2 T_j + \frac{1}{\sqrt{n}} Y_j^3 \right) f^{(3)}(ST_j), \\ |R_2| &\leq \frac{\|f^{(4)}\|}{2n^{3/2}} \sum_{j=1}^n \mathbb{E}S^3 \left(Y_j^3 T_j + \frac{1}{\sqrt{n}} Y_j^4 \right) \leq \frac{\|f^{(4)}\|}{2\sqrt{n}} \left(\mathbb{E}|Y^3| + \frac{\mathbb{E}Y^4}{\sqrt{n}} \right) \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4}. \end{aligned}$$

Here we used that $\mathbb{E}Y_j = 0$ and $\mathbb{E}Y_j^2 = 1$ to simplify R_1 . To bound R_1 we Taylor expand $f''(ST_j)$ about W to obtain

$$|R_1| \leq |\mathbb{E}S f''(W)| + \frac{\|f^{(3)}\|}{\sqrt{n}}.$$

We now consider bound N_1 . Using independence and that $\mathbb{E}Y_j = 0$ and $\mathbb{E}Y_j^2 = 1$ gives

$$N_1 = \frac{1}{n} \sum_{j=1}^n \mathbb{E}S^2 T f^{(3)}(W_j) + \frac{\mathbb{E}Y^3}{n^{3/2}} \sum_{j=1}^n \mathbb{E}S^2 f^{(3)}(ST_j).$$

Taylor expanding the $f^{(3)}(ST_j)$ about W allows us to write N_1 as

$$N_1 = \mathbb{E}S^2 T f^{(3)}(W) + R_3,$$

where

$$\begin{aligned} |R_3| &= \left| \frac{\mathbb{E}Y^3}{n^{3/2}} \sum_{j=1}^n \mathbb{E}S^2 f^{(3)}(ST_j) - \frac{1}{n^{3/2}} \sum_{j=1}^n \mathbb{E}Y_j S^3 \left(T_j + \frac{1}{\sqrt{n}} Y_j \right) f^{(4)}(ST_j^{[1]}) \right| \\ &\leq \frac{\|f^{(3)}\| \mathbb{E}|Y^3|}{\sqrt{n}} + \frac{\|f^{(4)}\|}{\sqrt{n}} \left(1 + \frac{1}{\sqrt{n}} \right) \left(3 + \frac{\mathbb{E}X^4}{m} \right)^{3/4}. \end{aligned}$$

Putting this together, we have shown that

$$\mathbb{E}ST^2 f''(W) = \mathbb{E}S^2 T f^{(3)}(W) + R_1 + R_2 + R_3.$$

Rearranging and apply the triangle inequality now gives

$$|\mathbb{E}S^2Tf^{(3)}(W)| \leq |\mathbb{E}ST^2f''(W)| + |R_1| + |R_2| + |R_3|,$$

and summing up the remainder terms completes the proof.

A.4 Proof of Lemma 4.7

We prove that inequality (4.5) holds; inequality (4.6) then follows by symmetry. We begin by obtaining a formula for the third order partial derivative of ψ with respect to s . Using a straightforward generalisation of the proof of Lemma 3.2 of Raič [35] it can be shown that

$$\frac{\partial^3\psi}{\partial s^3} = \int_0^\infty \int_{\mathbb{R}^2} \frac{e^{-u}}{\sqrt{1-e^{-2u}}} \frac{\partial^2}{\partial s^2} g(z_s, z_t) \phi'(x) \phi(y) \, dx \, dy \, du, \tag{A.3}$$

where $z_s = e^{-u}s + \sqrt{1-e^{-2u}}x$, $z_t = e^{-u}t + \sqrt{1-e^{-2u}}y$, and $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$ so that $\phi'(x) = -x\phi(x)$. We now calculate the second order partial derivative of g with respect to s . Since

$$\frac{\partial z_s}{\partial s} = \frac{\partial z_t}{\partial t} = e^{-u} \quad \text{and} \quad \frac{\partial z_t}{\partial s} = \frac{\partial z_s}{\partial t} = 0,$$

we have that

$$\frac{\partial^2 g}{\partial s^2} = e^{-2s} \{ z_s^a z_t^{b+2} f^{(4)}(z_s z_t) + 2a z_s^{a-1} z_t^{b+1} f^{(3)}(z_s z_t) + a(a-1) z_s^{a-2} z_t^b f''(z_s z_t) \}.$$

We now use the simple inequality that $|p+q|^n \leq 2^{n-1}(|p|^n + |q|^n)$ to obtain the following bound on z_s

$$|z_s^n| = |e^{-u}s + \sqrt{1-e^{-2u}}x|^n \leq 2^{n-1}(e^{-nu}|s|^n + (1-e^{-2u})^{n/2}|x|^n) \leq 2^{n-1}(|s|^n + |x|^n)$$

and a similar inequality holds for z_t . With these inequalities we have the following bound

$$\begin{aligned} \left| \frac{\partial^2 g}{\partial s^2} \right| &\leq e^{-2u} \{ 2^{a+b} \|f^{(4)}\| (|s|^a + |x|^a)(|t|^{b+2} + |y|^{b+2}) \\ &\quad + 2a 2^{a+b-2} \|f^{(3)}\| (|s|^{a-1} + |x|^{a-1})(|t|^{b+1} + |y|^{b+1}) \\ &\quad + a(a-1) 2^{a+b-4} \|f''\| (|s|^{a-2} + |x|^{a-2})(|t|^b + |y|^b) \}. \end{aligned}$$

Applying this bound to equation (A.3) gives the following bound on the third order partial derivative of ψ with respect to s :

$$\begin{aligned} \left| \frac{\partial^3\psi}{\partial s^3} \right| &\leq \int_0^\infty \int_{\mathbb{R}^2} \frac{e^{-u}}{\sqrt{1-e^{-2u}}} \left| \frac{\partial^2 g}{\partial s^2} \right| |x| \phi(x) \phi(y) \, dx \, dy \, du \\ &\leq \int_0^\infty \int_{\mathbb{R}^2} \frac{e^{-3u}}{\sqrt{1-e^{-2u}}} \{ 2^{a+b} \|f^{(4)}\| (|s|^a + |x|^a)(|t|^{b+2} + |y|^{b+2}) \\ &\quad + a 2^{a+b-1} \|f^{(3)}\| (|s|^{a-1} + |x|^{a-1})(|t|^{b+1} + |y|^{b+1}) \\ &\quad + a(a-1) 2^{a+b-4} \|f''\| (|s|^{a-2} + |x|^{a-2})(|t|^b + |y|^b) \} |x| \phi(x) \phi(y) \, dx \, dy \, du \\ &= \frac{\pi}{4} \int_{\mathbb{R}^2} \{ 2^{a+b} \|f^{(4)}\| (|s|^a + |x|^a)(|t|^{b+2} + |y|^{b+2}) \\ &\quad + a 2^{a+b-1} \|f^{(3)}\| (|s|^{a-1} + |x|^{a-1})(|t|^{b+1} + |y|^{b+1}) \\ &\quad + a(a-1) 2^{a+b-4} \|f''\| (|s|^{a-2} + |x|^{a-2})(|t|^b + |y|^b) \} |x| \phi(x) \phi(y) \, dx \, dy, \tag{A.4} \end{aligned}$$

where the final equality follows from the formula $\int_0^\infty \frac{e^{-3u}}{\sqrt{1-e^{-2u}}} du = \frac{\pi}{4}$ (see Gaunt [18], p. 19). We can now obtain the desired bound by using the following formula (see formula 17 of Winkelbauer [44]) to evaluate (A.4):

$$\int_{-\infty}^\infty |x|^k \phi(x) dx = \frac{2^{k/2} \Gamma(\frac{k+1}{2})}{\sqrt{\pi}} = (k-1)!! \begin{cases} \sqrt{\frac{2}{\pi}}, & k \text{ odd,} \\ 1, & k \text{ even,} \end{cases}$$

$$\leq (k-1)!!,$$

which completes the proof.

B Elementary properties of modified Bessel functions

Here we list standard properties of modified Bessel functions that are used throughout this paper. All these formulas can be found in Olver et al. [30], except for the inequalities, which are given in Gaunt [19].

B.1 Definitions

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

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

The *modified Bessel function of the second kind* of order $\nu \in \mathbb{R}$ can be defined in terms of the modified Bessel function of the first kind as follows

$$K_\nu(x) = \frac{\pi}{2 \sin(\nu\pi)} (I_{-\nu}(x) - I_\nu(x)), \quad \nu \notin \mathbb{Z}, x \in \mathbb{R},$$

$$K_\nu(x) = \lim_{\mu \rightarrow \nu} K_\mu(x) = \lim_{\mu \rightarrow \nu} \frac{\pi}{2 \sin(\mu\pi)} (I_{-\mu}(x) - I_\mu(x)), \quad \nu \in \mathbb{Z}, x \in \mathbb{R}.$$

B.2 Basic properties

For $\nu \in \mathbb{R}$, the modified Bessel function of the first kind $I_\nu(x)$ and the modified Bessel function of the second kind $K_\nu(x)$ are regular functions of x . For $\nu > -1$ and $x > 0$ we have $I_\nu(x) > 0$ and $K_\nu(x) > 0$. For all $\nu \in \mathbb{R}$ the modified Bessel function $K_\nu(x)$ is complex-valued in the region $x < 0$.

B.3 Asymptotic expansions

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

$$I_\nu(x) \sim \frac{e^x}{\sqrt{2\pi x}}, \quad x \rightarrow \infty, \tag{B.3}$$

$$K_\nu(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x}, \quad x \rightarrow \infty. \tag{B.4}$$

B.4 Identities

$$I_\nu(x)K_{\nu+1}(x) - I_{\nu+1}(x)K_\nu(x) = \frac{1}{x}, \tag{B.5}$$

$$(-1)^\nu K_\nu(x) - \pi i I_\nu(x) = K_\nu(-x). \tag{B.6}$$

B.5 Differentiation

$$K'_\nu(x) = -\frac{1}{2}(K_{\nu+1}(x) - K_{\nu-1}(x)), \tag{B.7}$$

$$\frac{d}{dx} \left(\frac{I_\nu(x)}{x^\nu} \right) = \frac{I_{\nu+1}(x)}{x^\nu}, \tag{B.8}$$

$$\frac{d}{dx} \left(\frac{K_\nu(x)}{x^\nu} \right) = -\frac{K_{\nu+1}(x)}{x^\nu}, \tag{B.9}$$

B.6 Modified Bessel differential equation

The modified Bessel differential equation is

$$x^2 f''(x) + x f'(x) - (x^2 + \nu^2) f(x) = 0. \tag{B.10}$$

The general solution is $f(x) = AI_\nu(x) + BK_\nu(x)$.

B.7 Inequalities

Let $-1 < \beta < 1$ and $n = 0, 1, 2, \dots$, then for $x \geq 0$ we have

$$\left| \frac{d}{dx} \left(\frac{e^{-\beta x} K_\nu(x)}{x^\nu} \right) \int_0^x e^{\beta t} t^\nu I_\nu(x) dt \right| < \frac{2(\beta + 1)}{2\nu + 1} x K_{\nu+1}(x) I_\nu(x) < \infty, \quad \nu > -\frac{1}{2}, \tag{B.11}$$

$$\left| \frac{d^n}{dx^n} \left(\frac{e^{-\beta x} I_\nu(x)}{x^\nu} \right) \int_x^\infty e^{\beta t} t^\nu K_\nu(t) dt \right| < \frac{\sqrt{\pi} \Gamma(\nu + 1/2)}{(1 - \beta^2)^{\nu+1/2} \Gamma(\nu + 1)}, \quad \nu \geq \frac{1}{2}, \tag{B.12}$$

$$\left| \frac{d^n}{dx^n} \left(\frac{e^{-\beta x} I_\nu(x)}{x^\nu} \right) \int_x^\infty e^{\beta t} t^\nu K_\nu(t) dt \right| < \frac{(e + 1) 2^{2\nu} \Gamma(\nu + 1/2)}{1 - |\beta|}, \quad |\nu| < \frac{1}{2}. \tag{B.13}$$

Acknowledgments. During the course of this research the author was supported by an EPSRC DPhil Studentship and an EPSRC Doctoral Prize. The author would like to thank Gesine Reinert for the valuable guidance she provided on this project. The author would also like to thank two anonymous referees for their helpful comments which have lead to a substantial improvement in the presentation of this paper.

Electronic Journal of Probability

Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS¹)
- Non profit, sponsored by IMS², BS³, PKP⁴
- Purely electronic and secure (LOCKSS⁵)

Help keep the journal free and vigorous

- Donate to the IMS open access fund⁶ (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

¹OJS: Open Journal Systems <http://pkp.sfu.ca/ojs/>

²IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

³BS: Bernoulli Society <http://www.bernoulli-society.org/>

⁴PK: Public Knowledge Project <http://pkp.sfu.ca/>

⁵LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

⁶IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>