Rates of Convergence of Variance-Gamma Approximations via Stein's Method

Robert E. Gaunt

The Queen's College

University of Oxford

A thesis submitted for the degree of *Doctor of Philosophy*.

Hilary Term 2013

Acknowledgements

I would like to thank my supervisor, Professor Gesine Reinert, for the exceptional guidance she has given me throughout the course of this work. I am very grateful for everything she has done for me, and it has been a pleasure to be her student. I would like to thank Professor Alison Etheridge for supervising me in Trinity Term 2010 and for the advice she has given me. I would also like to thank the Engineering and Physical Sciences Research Council for their financial support. Finally, I would like to thank my family for their love and support.

Abstract

Stein's method is a powerful technique that can be used to obtain bounds for approximation errors in a weak convergence setting. The method has been used to obtain approximation results for a number of distributions, such as the normal, Poisson and Gamma distributions. A major strength of the method is that it is often relatively straightforward to apply it to problems involving dependent random variables.

In this thesis, we consider the adaptation of Stein's method to the class of Variance-Gamma distributions. We obtain a Stein equation for the Variance-Gamma distributions. Uniform bounds for the solution of the Symmetric Variance-Gamma Stein equation and its first four derivatives are given in terms of the supremum norms of derivatives of the test function. New formulas and inequalities for modified Bessel functions are obtained, which allow us to obtain these bounds. We then use local approach couplings to obtain bounds on the error in approximating two asymptotically Variance-Gamma distributed statistics by their limiting distribution. In both cases, we obtain a convergence rate of order n^{-1} for suitably smooth test functions.

The product of two normal random variables has a Variance-Gamma distribution and this leads us to consider the development of Stein's method to the product of r independent mean-zero normal random variables. An elegant Stein equation is obtained, which motivates a generalisation of the zero bias transformation. This new transformation has a number of interesting properties, which we exploit to prove some limit theorems for statistics that are asymptotically distributed as the product of two central normal distributions.

The Variance-Gamma and Product Normal distributions arise as functions of the multivariate normal distribution. We end this thesis by demonstrating how the multivariate normal Stein equation can be used to prove limit theorems for statistics that are asymptotically distributed as a function of the multivariate normal distribution. We establish some sufficient conditions for convergence rates to be of order n^{-1} for smooth test functions, and thus faster than the $O(n^{-1/2})$ rate that would arise from the Berry-Esséen Theorem. We apply the multivariate normal Stein equation approach to prove Variance-Gamma and Product Normal limit theorems, and we also consider an application to Friedman's χ^2 statistic.

Contents

Chapter 1

Introduction

In this thesis, we will consider the problem of bounding the error in approximations involving the class of Variance-Gamma distributions. We will obtain our bounds using a powerful general method for proving limits theorems in probability, which was introduced in 1972 by Charles Stein and is known as Stein's method. When applied the method gives bounds of the form $|\mathbb{E}h(X) - \mathbb{E}h(D)|$, where X is the statistic we are interested in, D is the limit distribution and h is an arbitrarily chosen measurable function, which is known as the test function. In this thesis we will restrict ourselves to the case of smooth h, however it is worth noting that if we take $h = h_x = \chi_{(-\infty,z]}(x)$, where z is a continuity point of the limiting distribution, and take the supremum of the bounds over z then we will obtain traditional convergence in distribution results.

The class of Variance-Gamma distributions form a subclass of the class of Generalised Hyperbolic distributions. The Variance-Gamma distributions are used in financial modelling and were introduced into the financial literature by Madan and Seneta [46] in 1990. An important property of Variance-Gamma random variables is that they have a simple characterisation in terms of normal and Gamma random variables. In particular, the random variable $\sum_{k=1}^{r} Z_k Z_{r+k}$, where the Z_k are i.i.d. mean zero normal random variables, has a Variance–Gamma distribution.

1.1 A motivating example

We now consider a motivating example, which concerns word sequence comparison. Comparison of the similarities between two random sequences arise, for example, in biological sequences comparisons. One way of comparing the sequences uses k-tuples (a sequence of letters of length k). If two sequences with letters chosen from a finite alphabet are closely related in a biological sense, we would expect the k-tuple content of both sequences to be similar. A statistic for sequence comparison based on k-tuple content, known as the D_2^* statistic was suggested in Reinert et al. [62].

More formally, suppose that the two sequences, $\mathbf{A} = A_1A_2\ldots A_m$ and $\mathbf{B} = B_1B_2\ldots B_n$, say, are composed of i.i.d. letters that are drawn from a finite alphabet $\mathcal{A}$ of size d. The null hypothesis is typically that the two sequences are independent. For $a \in \mathcal{A}$ let p_a denote the probability of letter a. For $\mathbf{w} = (w_1, \ldots, w_k) \in \mathcal{A}^k$ let

$$X_{\mathbf{w}} = \sum_{i=1}^{\bar{m}} \mathbf{1}(A_i = w_1, \ldots, A_{i+k-1} = w_k)$$

count the number of occurrences of $\mathbf{w}$ in $\mathbf{A}$. Here $\bar{m} = m - k + 1$. Similarly, we let $Y_{\mathbf{w}}$ count the number of occurrences of $\mathbf{w}$ in $\mathbf{B}$, and let $\bar{n} = n - k + 1$. For $\mathbf{w} = w_1 \cdots w_k$ denote by $p_{\mathbf{w}} = \prod_{i=1}^{k} p_{w_i}$ the probability of occurrence of $\mathbf{w}$. Then D_2^* is defined by

$$D_2^* = \sum_{\mathbf{w}\in\mathcal{A}^k} \frac{(X_{\mathbf{w}} - \bar{m}p_{\mathbf{w}})(Y_{\mathbf{w}} - \bar{n}p_{\mathbf{w}})}{\sqrt{\bar{m}\bar{n}}p_{\mathbf{w}}}.$$

The statistic D_2^* is motivated by estimating the standardised counts

$$X_{\mathbf{w}}^0 = \frac{X_{\mathbf{w}} - \bar{m}p_{\mathbf{w}}}{\sqrt{\mathrm{Var} X_{\mathbf{w}}}}, \qquad \text{and} \qquad Y_{\mathbf{w}}^0 = \frac{Y_{\mathbf{w}} - \bar{n}p_{\mathbf{w}}}{\sqrt{\mathrm{Var} Y_{\mathbf{w}}}}.$$

We have that $\mathrm{Var} X_{\mathbf{w}} = \bar{m}p_{\mathbf{w}}(1-p_{\mathbf{w}})$. For relatively rare words $\mathbf{w}$ (rare words will occur provided that k is reasonably large) we have $1 - p_{\mathbf{w}} \approx 1$ and so we can approximate $\mathrm{Var} X_{\mathbf{w}}$ by $\bar{m}p_{\mathbf{w}}$.

Finally, we note that even if the sequence letters are i.i.d., the different random variable indicators of word occurrences are not independent due to overlaps. For example, if the word $\mathbf{w} = ATAT$ occurs at position i in the sequence then another occurrence of $\mathbf{w}$ is much more likely to occur at position $i+2$ than if $\mathbf{w}$ did not occur at position i, and an occurrence of $\mathbf{w}$ at position $i+1$ is not possible. There is also a dependence structure at the global level, as the word count statistics $X_{\mathbf{w}}$ themselves are not independent. For example, if the word $\mathbf{u} = AAAA$ occurs very frequently, then it may mean that other word sequences, such as $\mathbf{v} = CCCC$ may occur very rarely. For a more detailed account of the dependence structure see Reinert et al. [60].

If we were to perform a statistical test to compare two sequences using the D_2^* statistic it would be important to know the asymptotic distribution of D_2^*, under the null hypothesis that the two sequences are independent, as the length of the two sequences increases. It would also be of interest to have a bound on the error in approximating D_2^* by its asymptotic distribution, especially when the sequence lengths are small. This is important because if we are carrying out a statistical test at the α level, then we calculate a critical value c_α, say, based on the assumption that D_2^* has converged to its limiting distribution. Knowledge of the error bounds for the approximation allows us to determine the accuracy of c_α. This is essentially a quantification of our 'trust' in the outcome of the test.

Stein's method is particularly well suited to the problem of bounding the errors in the approximation of D_2^* by its asymptotic distribution. If successfully applied Stein's method would give a bound in the error in the weak convergence setting, with a bound of the form $|\mathbb{E}h(D_2^*) - \mathbb{E}h(D)|$, where D is the limit distribution and h is a test function. The complex dependence structure of D_2^* also motivates using Stein's method to tackle the problem, as it may be possible through appropriate use of couplings to disentangle the overlap structure of the words.

Also, for D_2^* the interest is in a bound which depends not only on sequence lengths but also on word length and on alphabet size. In practice we can choose the word length depending on sequence length and alphabet size. The beauty of Stein's method is that it can provide such bounds, where in principle more than one quantity can tend to infinity. This therefore motivates using Stein's method to obtain bounds on the error in approximating D_2^* by its limiting distribution.

For large sequence lengths m and n, we can see from the central limit theorem that the standardised counts $X_{\mathbf{w}}^0$ and $Y_{\mathbf{w}}^0$ will be approximately normally distributed with mean zero. Provided that dependence between the counts is not 'too large' (this is certainly the case for i.i.d. sequence letters that are drawn from a very large alphabet) then D_2^* would be approximately distributed as the sum of products of two independent, mean zero, normal random variables, and thus would be approximately Variance-Gamma distributed. However, if the dependence is 'too large' then we would perhaps not expect D_2^* to be asymptotically a Variance-Gamma distributed. The explicit bounds that Stein's method yields would allow us to quantify what 'too large' means in this context.

In this thesis, we will make substantial progress towards this goal. In Chapter 4 we obtain bounds on the distance between a Variance-Gamma distribution and a statistic of a similar form to D_2^*, but with the simpler assumption of i.i.d. random variables. In Chapter 6 we demonstrate how Stein's method can be used to approximate statistics that are asymptotically distributed as a function of the multivariate normal distribution. For large sequence lengths m and n this is the case for the D_2^* statistic (D_2^* is asymptotically of the form $\sum_k S_k T_k$, where the S_k and T_k are asymptotically normally distributed). The approach developed in Chapter 6 would be used to obtain explicit bounds for the error in approximating D_2^* by its limiting distribution, even if the limiting distribution is not a Variance-Gamma distribution. This approach certainly works in principle but, as a result of the complex dependence structure of D_2^*, very detailed calculations will still be required to establish these bounds.

1.2 Outline of thesis

This thesis has as main focus the adaptation of Stein's method to the class of Variance-Gamma distributions. Motivated by our development of Stein's method for Variance-Gamma distributions, we also consider the problem of adapting Stein's method to the Product Normal distributions and functions of the multivariate normal distribution.

In Chapter 2 we give a brief introduction to Stein's method. We begin by considering Stein's original work on the normal distribution and look at the techniques that allows us to produce bounds from the seemingly rather abstract Stein equation. We then see how Stein's method can be extended to produce approximations for other limiting distributions. We also establish new smoothness estimates for the solution of the Gamma Stein equation.

We begin Chapter 3 by reviewing the class of Variance-Gamma distributions and some of their basic properties. A Stein equation for the Variance-Gamma distributions is obtained. Uniform bounds for the solution of the Symmetric Variance-Gamma Stein equation and its first four derivatives are given in terms of the supremum norms of derivatives of the test function.

In Chapter 4 we go on to consider the rate of convergence of two asymptotically Variance-Gamma distributed statistics. Consider two collections of i.i.d. random variables $(X_{ik})_{1\leq i\leq m,\,1\leq k\leq r}$ and $(Y_{jk})_{1\leq j\leq n,\,1\leq k\leq r}$, with the X_{ik} and Y_{jk} independent, each with zero mean and unit variance. Then the statistics $U_r = \frac{1}{m}\sum_{k=1}^{r}(\sum_{i=1}^{m}X_{ik})^2 - \frac{1}{n}\sum_{k=1}^{r}(\sum_{i=1}^{n}Y_{ik})^2$ and $V_r = \frac{1}{\sqrt{mn}}\sum_{i,j,k=1}^{m,n,r}X_{ik}Y_{jk}$ are shown to converge to Variance-Gamma random variables and we obtain a bound on the rate of convergence of order $m^{-1}+n^{-1}$ for smooth test functions.

In Chapter 5 we consider the problem of adapting Stein's method to the distribution of the product of r independent mean zero normal random variables. We obtain a Stein equation for this class of distributions. Motivated by this Stein equation, we establish a generalisation of the zero bias transformation. We obtain a number of useful properties of this new transformation and demonstrate now these properties can be used together with Stein equation to prove Product Normal limit theorems. We use our generalisation of the zero bias transformation and exchangeable pair couplings to prove some Variance-Gamma limit theorems, thereby utilising an alternative method to the local approach coupling that was used in Chapter 4.

In Chapter 6 we demonstrate how the multivariate normal Stein equation can be used to prove limit theorems for statistics that are asymptotically distributed as a function of the multivariate normal distribution. In Section 6.1 we use this approach to prove some limit theorems for the case of functions that have polynomial growth, and in Section 6.2 we illustrate how we could generalise these theorems. We also obtain some sufficient conditions for convergence rates in these limit theorems to be of order n^{-1}. We apply the theory developed in this chapter to prove χ^2, Variance-Gamma and Product Normal limit theorems, as well as an application to Friedman's χ^2 statistic. In Chapter 7 we summarise the results of the thesis and point out some interesting open questions.

In Appendix A we prove some lemmas of Chapter 3. Appendix B provides a list of some elementary properties of modified Bessel functions that we make use of throughout this thesis. In Appendix C we obtain new formulas and inequalities for some expressions involving modified Bessel functions. In Appendix D we use these formulas and inequalities to bound a number of expressions involving modified Bessel functions. We use these bounds in Chapter 3 to obtain smoothness estimates for the solution of the Variance-Gamma Stein equation.

1.3 Contributions made in this thesis

We now highlight the major results established in this thesis, in order of appearance.

In Theorem 2.19 we establish new smoothness estimates for the solution of the Gamma Stein equation.

We obtain a Stein equation for the Variance-Gamma distributions (see (3.14)). For certain parameter values this Stein equation reduces to the classical normal Stein equation of Stein [72] and the Gamma Stein equation that was given by Luk [44].

We obtain a unique bounded solution of the Variance-Gamma Stein equation (Lemmas 3.13 and 3.14), and obtain uniform bounds for the solution and its first four derivatives for the case of Symmetric Variance-Gamma distributions (Lemma 3.20 and Theorems 3.21 and 3.22).

In Theorems 4.3 and 4.4 we obtain bounds on convergence rates of order n^{-1} for two asymptotically Variance-Gamma distributed statistics, which is faster than the $O(n^{-1/2})$ rate that would arise from the Berry Esséen Theorem. These convergence rates were obtained using symmetry considerations that were introduced by Pickett [55], who obtained $O(n^{-1})$ bounds on convergence rates for χ^2 approximation.

A Stein equation for the distribution of r independent central normal random variables is obtained (see (5.8)). The Stein operator has the interesting property of being of order r. Most

Stein operators in the literature are of order one or two, although Goldstein and Reinert [31] have also obtained Stein operators of order $n \in \mathbb{Z}^+$.

We present a generalisation of the zero bias transformation (see Definition 5.11). This new transformation has a number of interesting properties, which are collected in Lemma 5.13 and Proposition 5.16. These properties generalise the properties of the zero bias transformation that are given in Lemma 2.1 of Goldstein and Reinert [29] and Proposition 2.3 of Chen et al. [18].

In Chapter 6 we develop an approach based on the multivariate normal Stein equation for approximating statistics that are asymptotically distributed as $g(\mathbf{Z})$, where $\mathbf{Z}$ has the multivariate normal distribution and $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is continuous. In Lemma 6.8 we show that if g is an even function then the solution of the multivariate normal Stein equation is an even function. We exploit this property to obtain limit theorems with $O(n^{-1})$ convergence rates later in the chapter.

We use this approach to obtain a limit theorem for the case that $g : \mathbb{R} \rightarrow \mathbb{R}$ has derivatives of polynomial growth and the random variables are i.i.d. (Theorem 6.1). With the additional assumptions that g is an even function (Theorem 6.6) or that $\mathbb{E}X^3 = 0$ (Theorem 6.7) we are able to obtain bounds of order n^{-1}. Theorem 6.6 and Lemmas 6.8 and 6.12 help shed light on the $O(n^{-1})$ bounds for χ^2 approximation of Pickett [55], as well as our $O(m^{-1} + n^{-1})$ bounds for Variance-Gamma approximation.

In Theorem 6.24 we present explicit bounds for the error in approximating Friedman's statistic by its limiting χ^2 distribution.

Chapter 2

Stein's Method

2.1 Weak Convergence

Stein's method is a powerful general technique for obtaining bounds between two probability distributions with respect to a probability metric. This allows us to obtain bounds on the errors in approximations. Before we see how Stein's method can be used to obtain these approximation results we briefly look at the probability metrics that are used in these bounds (see Shorack [68] for a more detailed account of the properties of these metrics).

Many distances between probability measures μ_X and μ_Y can be given by

$$d(\mu_X, \mu_Y) = \sup_{h \in \mathcal{H}} |\mathbb{E}h(X) - \mathbb{E}h(Y)|, \tag{2.1}$$

where $\mathcal{H}$ is some class of functions and X and Y are real valued-random variables with probability measures μ_X and μ_Y respectively. For example, if we take $\mathcal{H}_1 = \{h : \mathbb{R} \to \mathbb{R} : \ \|h\|_\infty = 1\}$, where $\|f\| = \|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$, then we obtain the total variation distance between μ_X and μ_Y,

$$d_{TV}(\mu_X, \mu_Y) = 2 \sup_{A \in \mathcal{B}} |\mu_X(A) - \mu_Y(A)| = \sup_{h \in \mathcal{H}_1} |\mathbb{E}h(X) - \mathbb{E}h(Y)|,$$

where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$. It is worth noting that the total variation distance between μ_X and μ_Y is often defined without the factor 2, which is the case, for example, in Chen et al. [18].

The Wasserstein distance between between μ_X and μ_Y is given by

$$d_W(\mu_X, \mu_Y) = \sup_{h \in \mathcal{H}_2} |\mathbb{E}h(X) - \mathbb{E}h(Y)|,$$

where $\mathcal{H}_2 = \{h : \mathbb{R} \to \mathbb{R} \,:\, h \text{ is Lipschitz with } \|h'\|_\infty = 1\}$.

Taking $\mathcal{H} = \mathcal{H}_3 = \{\chi_{(-\infty,z]}(x) \,:\, z \in \mathbb{R}\}$, where χ_I denotes the indicator function of an interval $I \in \mathbb{R}$, gives the Kolmogorov-Smirnov distance between μ_X and μ_Y,

$$d_K(\mu_X, \mu_Y) = \sup_{z\in\mathbb{R}} |P_X(z) - P_Y(z)| = \sup_{h\in\mathcal{H}_3} |\mathbb{E}h(X) - \mathbb{E}h(Y)|,$$

where P_X and P_Y are the cumulative distribution functions of X and Y respectively. For a sequence of random variables $(X_i)_{i\geq 1}$, if $\lim_{n\to\infty} d_K(\mu_{X_n}, \mu_X) = 0$, then the sequence of random variables $(X_i)_{i\geq 1}$ converges in distribution to X.

When successfully applied Stein's method gives a bound on the quantity $|\mathbb{E}h(X) - \mathbb{E}h(D)|$, where X is the statistic under consideration, D is the limiting distribution of X, and h is chosen from a suitable set $\mathcal{H}$, and is known as the test function. Taking the supremum of this over all test functions from a class of functions $\mathcal{H}$ gives the distance metric defined by (2.1). We have seen that if we take $\mathcal{H} = \{\chi_{(-\infty,z]}(x) \,:\, z \in \mathbb{R}\}$ then we get a bound in terms of the Kolmogorov-Smirnov distance and we can obtain traditional convergence in distribution results. However, in this thesis we restrict ourselves to continuous test functions and so we may wonder why it is useful to have a bound on $|\mathbb{E}h(X) - \mathbb{E}h(D)|$. We now use results from the theory of weak convergence on $\mathbb{R}$ to see that this is in fact an appropriate distance metric.

Definition 2.1 (**Weak Convergence**). *Let μ_n and μ be probability measures on $\mathbb{R}$. Then we say that μ_n* converges weakly *to μ, and write $\mu_n \Rightarrow \mu$, if and only if $\int_{\mathbb{R}} f \,\mathrm{d}\mu_n \to \int_{\mathbb{R}} f \,\mathrm{d}\mu$ for all $f \in C_b(\mathbb{R})$.*

It can be shown (see, for example Billingsley [15]) that convergence in distribution and weak convergence are equivalent statements on $\mathbb{R}$. The condition that $f \in C_b(\mathbb{R})$ is a bit restrictive for our purposes, as bounds arising from Stein's method involve derivatives of the test function. It turns out that we can weaken the condition that $f \in C_b(\mathbb{R})$. We will need the following:

Definition 2.2 (**Convergence-Determining Class**). *A sub-collection $\mathcal{F}$ of $C_b(\mathbb{R})$ is said to be a* convergence determining class *for weak convergence if and only if*

$$\lim_{n\to\infty} \int_{\mathbb{R}} f \,\mathrm{d}\mu_n = \int_{\mathbb{R}} f \,\mathrm{d}\mu, \ \forall f \in \mathcal{F} \quad \Longrightarrow \quad \mu_n \Rightarrow \mu \ \text{ as } n \to \infty.$$

It turns out that $C_b^k(\mathbb{R})$, where k is a positive integer, is a convergence determining class (see Billingsley [15]). Thus, if the bounds on $|\mathbb{E}h(X_n) - \mathbb{E}h(D)|$, which we obtain by Stein's method, are shown to tend to zero as $n \to \infty$, for all $h \in C_b^k(\mathbb{R})$, for some k, then we will have obtained a convergence in distribution result.

2.2 Stein's Method for Normal Approximation

In 1972, Stein [72] introduced a powerful method for studying normal approximations. The method rests on the following characterisation of the normal distribution, which can be found in Stein [73].

Lemma 2.3. *Let Z be a real-valued random variable. Then $\mathscr{L}(Z) = N(0,1)$ if and only if, for all $f : \mathbb{R} \to \mathbb{R}$ such that $f \in C^2(\mathbb{R})$ and $\mathbb{E}|f''(X)|$ is finite for $X \sim N(0,1)$,*

$$\mathbb{E}\{f''(Z) - Zf'(Z)\} = 0. \tag{2.2}$$

Remark 2.4. Stein's original characterisation was given in terms of f rather than f', and this is still the case in much of the literature. However, in this thesis we differ from the convention of Stein [73]. The reason for will become apparent in Section 2.3, when we consider an approach to Stein characterisations based on generators of Markov processes.

The characterisation of Lemma 2.3 gives rise to the following inhomogeneous differential equation, known as the Stein equation:

$$f''(x) - xf'(x) = h(x) - \Phi h, \tag{2.3}$$

where Φh denotes the quantity $\mathbb{E}h(X)$ for $X \sim N(0,1)$, and the test function h is a real-valued bounded function. For any such test function, a solution f' to (2.3) exists. Now, evaluating both sides at any random variable W and taking expectations gives

$$\mathbb{E}\{f''(W) - Wf'(W)\} = \mathbb{E}h(W) - \Phi h. \tag{2.4}$$

If we take $W \sim N(0,1)$ then (2.4) reduces to (2.2). Moreover, we recognise the relationship between the right–hand side of (2.4) and the distance metric (2.1) that we discussed in Section 2.1. Thus the problem of bounding the error in approximating a random variable W by a standard normal random variable in the weak convergence setting, reduces to the bounding the left-hand side of (2.4) for all $h \in \mathcal{H}$. As we shall see, the left-hand side can often be bounded using Taylor expansions about a random variable coupled with W. The bound will involve derivatives of f, and these derivatives can be bounded in terms of derivatives of h by the following lemma, which is given in Stein [73].

Lemma 2.5. *Suppose $h : \mathbb{R} \to \mathbb{R}$ is a bounded absolutely continuous function. Then the unique bounded solution of the Stein equation (2.3) is given by*

$$f'(x) = -\mathrm{e}^{x^2/2} \int_{-\infty}^{x} (h(t) - \Phi h)\mathrm{e}^{-t^2/2}\,\mathrm{d}t, \tag{2.5}$$

and the following bounds on its derivatives hold

$$\|f'\| \leq \sqrt{\frac{\pi}{2}}\|h - \Phi h\|, \qquad \|f''\| \leq 2\|h - \Phi h\|, \qquad \|f^{(3)}\| \leq 2\|h'\|.$$

The bounds on the derivatives are obtained directly from the solution (2.5) and the use of simple inequalities. However, it becomes tedious to bound higher order derivatives this way. Making use of the theory of generators of Markov processes it is possible to achieve a bound on the derivatives of general order of the solution, in terms of derivatives of h. We will see how this is done in Section 2.3, and bounds on the derivatives of general order of the solution are given in Lemma 2.14.

With these bounds we may obtain bounds on the error in asymptotic approximations by bounding the left-hand side of (2.4). Typically this approach involves Taylor expanding about a random variable coupled with W. The basic idea is to construct an auxiliary random variable W^* on the same probability space as W, such that W^* is close, but not identical, to W and has properties of which we can take advantage for our particular situation. Such constructions are known as *couplings*. There are different types of couplings available to us, each with different properties and useful in a different context. Throughout this thesis we will mostly consider local approach couplings, which are generally appropriate when we consider a sum of an independent family $\{X_i\}_{i\in I}$.

To see how Stein's method can be used to obtain approximation results, consider the following simple example, which can be founding in Reinert [58].

Example 2.6. Let $X, X_1, X_2, \ldots, X_n$ be a a collection of i.i.d. random variables, with mean zero, unit variance and $\mathbb{E}|X^3| < \infty$. Let $W = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i$. By the central limit theorem, W converges in distribution to a standard normal random variable. We now apply Stein's method to obtain a bound on the error in approximating W by a standard normal random variable, and by the theory of weak convergence, introduced in Section 2.1, also recover a proof of the central limit theorem.

Define $W_i = W - \frac{1}{\sqrt{n}}X_i$. The coupling is between W and W_i and both are clearly defined on the same probability space. Then, for all $f \in C^3(\mathbb{R})$, we may Taylor expand the left–hand side of (2.4) and use that W_i and X_i are independent, to obtain

$$\begin{aligned}
\mathbb{E}Wf'(W) &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}X_if'(W) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}X_if'(W_i) + \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}X_i(W-W_i)f''(W_i) + R_1 \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}X_i^2\mathbb{E}f''(W_i) + R_1 \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}f''(W) + R_1 + R_2 \\
&= \mathbb{E}f''(W) + R_1 + R_2,
\end{aligned}$$

where R_1 and R_2 denote the errors from the Taylor expansions. These can be bounded as follows

$$\begin{aligned}
|R_1| &\leq \frac{\|f^{(3)}\|}{2\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}|X_i(W-W_i)^2| = \frac{\|f^{(3)}\|}{2n\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}|X_i^3| = \frac{\|f^{(3)}\|\mathbb{E}|X^3|}{2\sqrt{n}}; \\
|R_2| &\leq \frac{\|f^{(3)}\|}{n\sqrt{n}}\sum_{i=1}^{n}\mathbb{E}|X_i| \leq \frac{\|f^{(3)}\|}{\sqrt{n}},
\end{aligned}$$

where the final inequality follows since $\mathbb{E}|X_i| \leq \sqrt{\mathbb{E}X_i^2} = 1$, by the Cauchy-Schwarz inequality. We can use Lemma 2.5 to bound these remainders in terms of derivatives of h. Therefore, for $h \in C^1(\mathbb{R})$ the right-hand side of (2.4) may be bounded as follows

$$|\mathbb{E}h(W) - \Phi h| = |\mathbb{E}f''(W) - \mathbb{E}Wf'(W)| \leq |R_1| + |R_2| \leq \frac{1}{\sqrt{n}}\|h'\|(\mathbb{E}|X^3| + 2). \tag{2.6}$$

We have obtained an explicit bound in the error in approximation, in the Wasserstein distance, of the distribution of W by the standard normal random distribution. Note that this $O(n^{-1/2})$ rate of convergence is the same as the that of the Berry-Esséen Theorem (Berry [13] and Esséen [25]) which gives a bound on the Kolmogorov-Smirnov distance between W and a standard normal random variable:

$$d_K(W,Z) = \sup_{z\in\mathbb{R}}|P(W \leq z) - \Phi(z)| \leq \frac{K\mathbb{E}|X^3|}{\sqrt{n}},$$

where K is a universal constant and $\Phi(z)$ is the standard normal distribution function.

Goldstein and Reinert [29] showed that for the case $\mathbb{E}X^3 = 0$ the convergence rate is of order n^{-1}. When the first p moments of the distribution of X agree with the first p moments of the standard normal distribution, we can use Stein's method to obtain a bound of order $n^{(p-1)/2}$ on the convergence rate, as the following theorem demonstrates. This result appears to be new, but there are other instances in the literature of faster convergence rates for normal approximation in the case of vanishing moments; see, for example, Hall [35].

Theorem 2.7. *Let $X, X_1, X_2, \ldots, X_n$ be a collection of i.i.d. random variables with $\mathbb{E}X^k = \mathbb{E}Z^k$, where $Z \sim N(0,1)$, for all positive integers $k \leq p$, and suppose that $\mathbb{E}|X|^{p+1} < \infty$. Then, for all $h \in C_b^{p+1}(\mathbb{R})$, we have*

$$|\mathbb{E}h(W) - \Phi h| \leq \frac{M_p}{(p-1)!\, n^{(p-1)/2}}\bigg(\frac{2^{(p-1)/2}\Gamma(\frac{p}{2})}{\sqrt{\pi}} + \frac{\mathbb{E}|X|^{p+1}}{p}\bigg), \tag{2.7}$$

where

$$M_p = \min\bigg\{\frac{\sqrt{\pi}\Gamma(\frac{p+1}{2})}{2\Gamma(\frac{p}{2}+1)}\|h^{(p)}\|, \ \frac{\|h^{(p+1)}\|}{p+1}\bigg\}.$$

Proof. We Taylor expand $f''(W)$ and $f'(W)$ about $W_i = W - \frac{1}{\sqrt{n}}X_i$ to obtain

$$\begin{aligned}\mathbb{E}f''(W) - \mathbb{E}Wf'(W) &= \frac{1}{n}\sum_{i=1}^n \mathbb{E}f''(W) - \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbb{E}X_i f'(W) \\ &= \sum_{i=1}^n\sum_{j=0}^{p-2}\frac{1}{j!\, n^{j/2+1}}\mathbb{E}X_i^j\mathbb{E}f^{(j+2)}(W_i) \\ &\quad - \sum_{i=1}^n\sum_{k=0}^{p-1}\frac{1}{k!\, n^{k/2+1/2}}\mathbb{E}X_i^{k+1}\mathbb{E}f^{(k+1)}(W_i) + R_1 + R_2,\end{aligned}$$

where

$$\begin{aligned}|R_1| &\leq \frac{\mathbb{E}|X|^{p+1}\|f^{(p+1)}\|}{p!\, n^{(p-1)/2}}, \\ |R_2| &\leq \frac{\mathbb{E}|X|^{p-1}\|f^{(p+1)}\|}{(p-1)!\, n^{(p-1)/2}} = \frac{\mathbb{E}|Z|^{p-1}\|f^{(p+1)}\|}{(p-1)!\, n^{(p-1)/2}} = \frac{\|f^{(p+1)}\|}{(p-1)!\, n^{(p-1)/2}}\cdot\frac{2^{(p-1)/2}\Gamma(\frac{p}{2})}{\sqrt{\pi}},\end{aligned}$$

where we used that $\mathbb{E}|Z|^k = \frac{2^{k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}}$ (see formula 17 of Winklebauer [75]). Since the X_i have mean zero and are identically distributed, we have

$$\begin{aligned}\mathbb{E}f''(W) - \mathbb{E}Wf'(W) &= \sum_{j=0}^{p-2}\frac{1}{j!\, n^{j/2+1}}\mathbb{E}X_i^j\mathbb{E}f^{(j+2)}(W_i) \\ &\quad - \sum_{k=1}^{p-1}\frac{1}{k!\, n^{k/2+1/2}}\mathbb{E}X_i^{k+1}\mathbb{E}f^{(k+1)}(W_i) + R_1 + R_2\end{aligned}$$

$$= \sum_{k=1}^{p-1} \frac{1}{k!\, n^{k/2+1/2}} [k\mathbb{E}X^{k-1} - \mathbb{E}X^{k+1}] f^{(k+1)}(W_i) + R_1 + R_2.$$

For even k we have $k\mathbb{E}X^{k-1} - \mathbb{E}X^{k+1} = 0$, and for odd k we have

$$k\mathbb{E}X^{k-1} - \mathbb{E}X^{k+1} = k \cdot \frac{2^{(k-1)/2}\Gamma(\frac{k}{2})}{\sqrt{\pi}} - \frac{2^{(k+1)/2}\Gamma(\frac{k}{2}+1)}{\sqrt{\pi}} = 0,$$

where we used that $\Gamma(x+1) = x\Gamma(x)$. Hence,

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

Therefore $|\mathbb{E}h(W) - \Phi h| \leq |R_1| + |R_2|$, and (2.7) now follows from using inequalities (2.17) and (2.18) (see Lemma 2.14, below) to bound $\|f^{(p+1)}\|$. □

Note that the bound of Theorem 2.7 involves derivatives of the test function h that have order greater than one. Hence, unlike in Example 2.6, we have not obtained a bound in Wasserstein distance. This will be the case for most of the limit theorems obtained in this thesis.

A major advantage of Stein's method over traditional Fourier methods is that the extension to the case of locally dependent random variables is often straightforward. We demonstrate this with the next example, which can be found in Barbour and Chen [7].

Example 2.8. Suppose that the random variables $X_1, X_2, \ldots, X_n$ have zero mean and are normalized so that $\mathbb{E}W^2 = 1$, where $W = \sum_{i=1}^n X_i$. We suppose that the X_i have the following local dependence structure. Suppose for each $i = 1, \ldots, n$, there exist index sets $\{i\} \subseteq A_i^{(1)} \subseteq A_i^{(2)} \subseteq \{1, \ldots, n\}$ such that $X_i \perp\!\!\!\perp \sigma\{X_j \,:\, j \in A_i^{(1)}\}$, and if $X_p \in \sigma\{X_j \,:\, j \in A_i^{(1)}\}$ then $X_p \perp\!\!\!\perp \sigma\{X_j \,:\, j \in A_i^{(2)}\}$, where $\sigma\{X_j\}$ denotes the σ–algebra generated by X_j.

For $k = 1, 2$ we define $X_i^{(k)} = \sum_{j \in A_i^{(k)}} X_j$ and $W_i^{(k)} = W - X_i^{(k)}$. Then, for all $f \in C^3(\mathbb{R})$,

$$\begin{aligned}
\mathbb{E}Wf'(W) &= \sum_{i=1}^n \mathbb{E}X_i f'(W) \\
&= \sum_{i=1}^n \mathbb{E}X_i f'(W_i^{(1)}) + \sum_{i=1}^n \mathbb{E}X_i X_i^{(1)} f''(W_i^{(1)}) + R_1 \\
&= \sum_{i=1}^n \mathbb{E}X_i X_i^{(1)} f''(W_i^{(2)}) + R_1 + R_2 \\
&= \sum_{i=1}^n \mathbb{E}X_i X_i^{(1)} \mathbb{E}f''(W_i^{(2)}) + R_1 + R_2
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}f''(W) + R_1 + R_2 + R_3 \\
&= \mathbb{E}f''(W) + R_1 + R_2 + R_3,
\end{aligned}
$$

where we used that $\mathbb{E}W^2 = \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} = 1$ to obtain the last equality. The remainders R_1, R_2 and R_3 are the errors from the Taylor expansions and can be bounded as follows

$$
\begin{aligned}
|R_1| &\leq \frac{\|f^{(3)}\|}{2} \sum_{i=1}^{n} \mathbb{E}|X_i(W - W_i^{(1)})^2| = \frac{\|f^{(3)}\|}{2} \sum_{i=1}^{n} \mathbb{E}|X_i(X_i^{(1)})^2|; \\
|R_2| &\leq \|f^{(3)}\| \sum_{i=1}^{n} \mathbb{E}|X_i X_i^{(1)}(W_i^{(1)} - W_i^{(2)})| = \|f^{(3)}\| \sum_{i=1}^{n} \mathbb{E}|X_i X_i^{(1)}(X_i^{(2)} - X_i^{(1)})|; \\
|R_3| &\leq \|f^{(3)}\| \sum_{i=1}^{n} |\mathbb{E}X_i X_i^{(1)}| \mathbb{E}|W - W_i^{(2)}| = \|f^{(3)}\| \sum_{i=1}^{n} |\mathbb{E}X_i X_i^{(1)}| \mathbb{E}|X_i^{(2)}|.
\end{aligned}
$$

Using Lemma 2.5 we can obtain the following bound on the left-hand of (2.4),

$$
|\mathbb{E}h(W) - \Phi h| \leq \|h'\| \sum_{i=1}^{n} \Big\{ \mathbb{E}|X_i(X_i^{(1)})^2| + 2\mathbb{E}|X_i X_i^{(1)}(X_i^{(2)} - X_i^{(1)})| + 2|\mathbb{E}X_i X_i^{(1)}| \mathbb{E}|X_i^{(2)}| \Big\}. \tag{2.8}
$$

If we were interested in a case with independent random variables then we could use (2.8), with $A_i^{(1)} = A_i^{(2)} = \{i\}$, thus $X_i = X_i^{(1)} = X_i^{(2)}$. Rescaling $X_i = \frac{1}{\sqrt{n}} Y_i$ and noting that $\mathbb{E}|Y_i| \leq \sqrt{\mathbb{E}Y_i^2} = 1$, by the Cauchy-Schwarz inequality, gives

$$
|\mathbb{E}h(W) - \Phi h| \leq \frac{\|h'\|}{n^{3/2}} \sum_{i=1}^{n} \{\mathbb{E}|Y_i|^3 + 2\mathbb{E}|Y_i| \mathbb{E}Y_i^2\} \leq \frac{1}{\sqrt{n}} \|h'\| (\mathbb{E}|Y^3| + 2),
$$

where Y is a random variable with the same distribution as Y_i. This bound is the same as bound (2.6) for the i.i.d. case, which we considered in Example 2.6.

Remark 2.9. There are a number of different local dependence conditions used in the Stein literature; see Chen et al. [18] for a list of conditions (the one used in Example 2.8 is (LD4*) of p. 245 of [18]).

2.3 Extension to Other Distributions

We now see how Stein's method can be applied to a variety of distributions. Provided we can obtain a characterisation like Lemma 2.3 for our chosen distribution, then we can obtain a Stein equation and bound the derivatives. Then we use couplings to obtain approximation results.

There is not a unique Stein equation for a given distribution, and therefore the idea is to find a Stein equation with 'good' properties, such as having a bounded solution and having a simple form that allows us to easily use Taylor expansions and couplings to produce bounds. The standard normal Stein equation (2.3) has these 'good' properties.

There are a number of techniques for finding Stein equations (see, for example, Reinert [59]). However, we will only consider two: the *generator method* which was developed by Barbour [5], [6] and also by Götze [33], and the *density approach* which was introduced by Stein et al. [74].

2.3.1 The generator approach

We seek a characterisation of our chosen distribution μ, say, similar to that of Lemma 2.3. This is achieved if we can find an operator $\mathcal{A}$ such that $X \sim \mu$ if and only if for all sufficiently smooth function f, $\mathbb{E}\mathcal{A}f(X) = 0$ holds. The idea is to choose $\mathcal{A}$ to be the generator of a Markov process $(X_t)_{t\geq 0}$ with stationary distribution μ. The generator $\mathcal{A}$ is a partial differential operator defined by $\mathcal{A}f(x) = \lim_{t\downarrow 0} \frac{1}{t}(T_t f(x) - f(x))$, where T_t is the *transition semigroup* operator, defined by $T_t f(x) = \mathbb{E}\{f(X_t)\,|\,X_0 = x\}$. A standard result for generators (see, for example, Ethier and Kurtz [22]) is: if μ is the stationary distribution of a Markov process then

$$X \sim \mu \Longleftrightarrow \mathbb{E}\mathcal{A}f(X) = 0,$$

for all real valued functions f for which $\mathcal{A}f(x)$ is defined. This parallels our characterisation of the standard normal distribution in Lemma 2.3, and gives rise to a Stein equation

$$\mathcal{A}f(x) = h(x) - h_\mu, \tag{2.9}$$

where $h_\mu = \int h\,\mathrm{d}\mu$. A Stein equation of the form (2.9) often simplifies the problem of obtaining smoothness estimates. Indeed, another standard result form Ethier and Kurtz [22] (Proposition 1.5) is that the solution of (2.9) is given by

$$f(x) = -\int_0^\infty \{T_t h(x) - \mu_h\}\,\mathrm{d}t, \tag{2.10}$$

if the integral exists. We now consider some examples.

Example 2.10 (Standard Normal distribution). The operator $\mathcal{A}f(x) = f''(x) - xf'(x)$ is the generator of the Ornstein-Uhlenbeck process with unit drift and noise $\sqrt{2}$, for which the stationary distribution is the standard normal distribution.

Example 2.11 (**Multivariate Normal distribution**). By considering a system of d independent Ornstein-Uhlenbeck processes each with unit drift and noise $\sqrt{2}$ we may deduce from the previous example that the Stein equation for the $\mathrm{MVN}(0, \mathrm{I}_d)$ distribution is given by

$$\sum_{i=1}^{d}\left(\frac{\partial^2 f(x)}{\partial x_i^2} - x_i\frac{\partial f(x)}{\partial x_i}\right) = h(x) - \mathbb{E}h(\mathbf{Z}), \tag{2.11}$$

where $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and $\mathbf{Z} \sim \mathrm{MVN}(0, \mathrm{I}_d)$.

Let Σ be a positive-definite $d \times d$ matrix, then the $\mathrm{MVN}(0, \Sigma)$ distribution is the stationary distribution of a Markov process given by the the unique strong solution of the stochastic differential equation

$$\mathrm{d}X_t = -X_t\,\mathrm{d}t + \sqrt{2}\Sigma^{1/2}\,\mathrm{d}B_t,$$

where dB_t denotes standard Brownian motion in $\mathbb{R}^d$. It can be shown, using a standard theorem concerning generators of Itô diffusions (see, for example Øksendal [51], p. 123) that the generator of this process is

$$\mathcal{A}f(x) = \frac{1}{2}\sum_{i,j=1}^{d}(\sqrt{2}\Sigma^{1/2}(\sqrt{2}\Sigma^{1/2})^T)_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} x_i\frac{\partial f}{\partial x_i}(x).$$

Therefore, a Stein equation for the $\mathrm{MVN}(0, \Sigma)$ distribution is

$$\nabla^T\Sigma\nabla f(x) - x^T\nabla f(x) = h(x) - \mathbb{E}h(\Sigma^{1/2}\mathbf{Z}). \tag{2.12}$$

Note that in the case $d = 1$ and $\Sigma = 1$ the multivariate normal Stein equation (2.12) reduces to the standard normal Stein equation (2.3).

Example 2.12. (**Gamma distribution**) Luk [44] obtained a Stein equation for the Gamma distribution using the generator method. Since the Gamma distribution is not consistently parametrized, we define it below to avoid confusion.

Definition 2.13 (**Gamma Distribution**). *The random variable X is said to have a* Gamma *distribution with parameters $r > 0, \lambda > 0$ if and only if it has p.d.f. given by*

$$\gamma_X(x;\ r, \lambda) = \frac{\lambda^r}{\Gamma(r)}\mathrm{e}^{-\lambda x}x^{r-1}, \qquad x > 0. \tag{2.13}$$

If (2.13) *holds then we write $X \sim \Gamma(r, \lambda)$.*

The $\Gamma(r, \lambda)$ distribution is the stationary distribution of a Markov process given by the solution of the stochastic differential equation

$$X_t = x + \int_0^t (r - \lambda X_s)\,\mathrm{d}s + \int_0^t \sqrt{2X_s}\,\mathrm{d}B_s,$$

where B_s denotes standard Brownian motion and $x \geq 0$. Again, it can be shown by a standard theorem concerning generators of Itô diffusions (see, for example Øksendal [51], p. 123) that the generator of this process is

$$\mathcal{A}f(x) = \frac{1}{2}(\sqrt{2x})^2 f''(x) + (r - \lambda x)f'(x).$$

Therefore, a Stein equation for the $\Gamma(r, \lambda)$ distribution is

$$xf''(x) + (r - \lambda x)f'(x) = h(x) - \Gamma_{r,\lambda}h, \tag{2.14}$$

where $\Gamma_{r,\lambda}h$ denotes the quantity $\mathbb{E}h(X)$ for $X \sim \Gamma(r, \lambda)$. Hence, a Stein equation for the χ^2 distribution with p degrees of freedom is

$$xf''(x) + \frac{1}{2}(p - x)f'(x) = h(x) - \chi^2_{(p)}h, \tag{2.15}$$

where $\chi^2_{(p)}h$ denotes the quantity $\mathbb{E}h(X)$ for $X \sim \chi^2_{(p)}$ (see Pickett [55] for a detailed account of Stein's method for χ^2 distributions).

Another useful feature of the generator method is that, provided that we can obtain a formula for the semigroup $T_t f(x)$, we can often exploit properties of the Markov process $(X_t)_{t\geq 0}$ to obtain a general formula for the k-th derivative of the solution to the Stein equation (2.9). We can then apply standard theorems for integrals, such as the dominated convergence theorem, to obtain a bound on the k–th derivative of the solution in terms of supremum norms of derivatives of the test function.

We now demonstrate this approach by considering the multivariate normal Stein equation (2.12). Suppose Σ is a positive-definite $d \times d$ matrix. The multivariate normal Stein equation is of the form $\mathcal{A}f(x) = h(x) - \mathbb{E}h(\Sigma^{1/2}h)$, where $\mathcal{A}f(x)$ is the generator of a Markov process that has the $\mathrm{MVN}(0, \Sigma)$ distribution as its stationary distribution. The Mehler formula (see for example Nualart [50], formula 1.54) for the semigroup of this process is:

$$T_t h(x) = \mathbb{E}h(x\mathrm{e}^{-s} + \sqrt{1 - \mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{Z}),$$

and applying this formula along with formula (2.10) shows that

$$f(x)=-\int_0^\infty[\mathbb{E}h(x\mathrm{e}^{-s}+\sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{Z})-\mathbb{E}h(\Sigma^{1/2}\mathbf{Z})]\,\mathrm{d}s \tag{2.16}$$

solves the Stein equation (2.12). Meckes [48] showed that, for each x, f is well-defined, and verified that it solves the Stein equation (2.12) by calculating its first and second order partial derivatives.

In the following lemma we obtain bounds for the k-th derivative of the solution (2.16) of the multivariate normal Stein equation (2.12).

Lemma 2.14. *Suppose Σ is a $d\times d$ positive-definite matrix and $h\in C_b^k(\mathbb{R}^d)$, where $k\geq 1$. Then, we have the following bounds for the partial derivatives of the solution (2.16) of the Stein equation (2.12):*

$$\left\|\frac{\partial^k f(x)}{\prod_{j=1}^k\partial x_{i_j}}\right\| \leq \frac{1}{k}\left\|\frac{\partial^k h(x)}{\prod_{j=1}^k\partial x_{i_j}}\right\|, \tag{2.17}$$

$$\left\|\frac{\partial^k f(x)}{\prod_{j=1}^k\partial x_{i_j}}\right\| \leq \frac{\Gamma(\frac{k}{2})}{\sqrt{2}\Gamma(\frac{k+1}{2})}\|\Sigma^{-1/2}\|_\infty\min_{1\leq l\leq k}\left\|\frac{\partial^{k-1} h(x)}{\prod_{j\neq l}^k\partial x_{i_j}}\right\|, \tag{2.18}$$

where the operator norm of a matrix A over $\mathbb{R}$ is defined by $\|A\|_\infty=\max_{1\leq i<\infty}\sum_{j=1}^d|a_{ij}|$, and $\partial^0 h\equiv h$. For bound (2.18) to hold we require that $\Sigma^{-1/2}$ exists.

Proof. We begin by proving inequality (2.17). If $\frac{\partial^k h(x)}{\prod_{j=1}^k\partial x_{i_j}}$ is bounded, then, by dominated convergence,

$$\frac{\partial^k f(x)}{\prod_{j=1}^k\partial x_{i_j}}=-\int_0^\infty \mathrm{e}^{-ks}\mathbb{E}\bigg[\frac{\partial^k h}{\prod_{j=1}^k\partial x_{i_j}}(x\mathrm{e}^{-s}+\sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{Z})\bigg]\,\mathrm{d}s. \tag{2.19}$$

Taking absolute values and using that $\int_0^\infty \mathrm{e}^{-ks}\,\mathrm{d}s=\frac{1}{k}$ yields (2.17).

We now prove inequality (2.18). Again, by dominated convergence,

$$\begin{aligned}\frac{\partial f(x)}{\partial x_{i_l}}&=-\int_0^\infty \mathrm{e}^{-s}\mathbb{E}\bigg[\frac{\partial h}{\partial x_{i_l}}(x\mathrm{e}^{-s}+\sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{Z})\bigg]\,\mathrm{d}s\\&=-\int_0^\infty\int_{\mathbb{R}^d}\mathrm{e}^{-s}\frac{\partial h}{\partial x_{i_l}}(x\mathrm{e}^{-s}+\sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{y})p(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s,\end{aligned}$$

where

$$p(\mathbf{y})=\frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}}\exp\bigg(-\frac{1}{2}\mathbf{y}^T\Sigma^{-1}\mathbf{y}\bigg).$$

Let $(\Sigma^{-1})_{ij} = \tilde{\sigma}_{ij}^2$. Then,

$$\frac{\partial}{\partial y_{i_l}}(\mathbf{y}^T\Sigma^{-1}\mathbf{y}) = \frac{\partial}{\partial y_{i_l}}\bigg(\sum_{j,k=1}^d \tilde{\sigma}_{jk}^2 y_{i_j}y_{i_k}\bigg) = 2\sum_{k=1}^d \tilde{\sigma}_{lk}^2 y_{i_k} = 2(\Sigma^{-1}\mathbf{y})_l,$$

and integration by parts therefore gives

$$\begin{aligned}\frac{\partial f(x)}{\partial x_{i_l}} &= -\int_0^\infty \int_{\mathbb{R}^d} \frac{\mathrm{e}^{-s}}{\sqrt{1-\mathrm{e}^{-2s}}}(\Sigma^{-1/2}\mathbf{y})_l h(x\mathrm{e}^{-s} + \sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{y})p(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s \\ &= -\int_0^\infty \frac{\mathrm{e}^{-s}}{\sqrt{1-\mathrm{e}^{-2s}}}\mathbb{E}\Big[(\Sigma^{-1/2}\mathbf{Z})_l h(x\mathrm{e}^{-s} + \sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{Z})\Big]\,\mathrm{d}s,\end{aligned}$$

where to obtain the first equality we used that h is bounded and that $\lim_{y_l\to-\infty}(\Sigma^{-1/2}\mathbf{y})_l p(\mathbf{y}) = 0$ and $\lim_{y_l\to\infty}(\Sigma^{-1/2}\mathbf{y})_l p(\mathbf{y}) = 0$. Again, by dominated convergence, we have

$$\frac{\partial^k f(x)}{\prod_{j=1}^k \partial x_{i_j}} = -\int_0^\infty \frac{\mathrm{e}^{-ks}}{\sqrt{1-\mathrm{e}^{-2s}}}\mathbb{E}\bigg[(\Sigma^{-1/2}\mathbf{Z})_l \frac{\partial^{k-1}h}{\prod_{j\neq l}^k \partial x_{i_j}}(x\mathrm{e}^{-s} + \sqrt{1-\mathrm{e}^{-2s}}\Sigma^{1/2}\mathbf{Z})\bigg]\,\mathrm{d}s, \quad (2.20)$$

thus

$$\bigg\|\frac{\partial^k f(x)}{\prod_{j=1}^k \partial x_{i_j}}\bigg\| \leq \mathbb{E}|(\Sigma^{-1/2}\mathbf{Z})_l| \min_{1\leq l\leq k}\bigg\|\frac{\partial^{k-1}h(x)}{\prod_{j\neq l}^k \partial x_{i_j}}\bigg\| \int_0^\infty \frac{\mathrm{e}^{-ks}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s.$$

Now, letting Z_1 denote a univariate standard normal random variable, we have $\mathbb{E}|(\Sigma^{-1/2}\mathbf{Z})_l| \leq \|\Sigma^{-1/2}\|_\infty \mathbb{E}|Z_1| = \sqrt{\frac{2}{\pi}}\|\Sigma^{-1/2}\|_\infty$, and the definite integral formula

$$\int_0^\infty \frac{\mathrm{e}^{-ks}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s = \frac{1}{2}\int_0^1 (1-t)^{-1/2}t^{k/2-1}\,\mathrm{d}t = \frac{1}{2}B\bigg(\frac{1}{2},\frac{k}{2}\bigg) = \frac{\Gamma(\frac{1}{2})\Gamma(\frac{k}{2})}{2\Gamma(\frac{k+1}{2})} = \frac{\sqrt{\pi}\Gamma(\frac{k}{2})}{2\Gamma(\frac{k+1}{2})}, \quad (2.21)$$

where $B(a,b)$ is the beta function. The desired inequality now follows. □

Remark 2.15. Lemma 2.14 gives us the choice of two bounds on the partial derivatives of the solution to the multivariate normal Stein equation. Bound (2.18) has a smoothing property, whereby the partial derivatives of h are one degree lower than f. This allows us to impose weaker differentiability conditions on the test functions in multivariate approximation limit theorems obtained by use of Stein's method. In some cases bound (2.17) may be preferable, since to use bound (2.18) we must compute $\|\Sigma^{-1/2}\|_\infty$.

Remark 2.16. Bound (2.17) of Lemma 2.14 is very well known in the literature (see, for example, Barbour [5], Goldstein and Rinott [32] and Reinert and Röllin [61]). However, bounds like (2.18) that have a smoothing property are much less common. Meckes [48] obtained bounds similar to (2.18) for derivatives of f up to third order, and these bounds were obtained using a very similar to proof to the one we used to prove inequality (2.18).

Luk [44] used a similar approach to that used in the proof of Lemma 2.14 to obtain bounds on the derivatives of the solution to the $\Gamma(r,\lambda)$ Stein equation (2.14). These bounds are as follows:

$$\|f^{(k)}\| \leq \frac{\|h^{(k)}\|}{k\lambda}, \qquad k \geq 1. \tag{2.22}$$

Pickett [55] was able to improve the bounds to involve the parameter r and require one less derivative of the test function h. This improvement resulted from an application of integration by parts similar to that used in our proof of bound (2.18), and Pickett was able to show that derivatives of the solution to the $\Gamma(r,\lambda)$ Stein equation satisfy

$$\|f^{(k)}\| \leq \left\{\sqrt{\frac{2\pi}{r}} + \frac{2}{r}\right\}\|h^{(k-1)}\|, \qquad k \geq 1, \tag{2.23}$$

where $h^{(0)} \equiv h$. Unfortunately, Pickett's proof contained an error. In Section 2.4 we present a corrected version of Pickett's proof that actually leads to an improvement of Pickett's bound to involve the order, k, of the derivative as well as the shape parameter r.

2.3.2 The density approach

We now review the *density approach* for obtaining Stein equations. The density approach was introduced by Stein et al. [74] and an account of the method is also given in Reinert [59]. Under several regularity, differentiability and integrability conditions (see Proposition 1.4 of Stein et al. [74] for these conditions) on the target density p, a random variable X has density p if and only if for all differentiable functions f for which

$$\int_{-\infty}^{\infty} |f'(x)|p(x)\,\mathrm{d}x < \infty$$

we have

$$\mathbb{E}\{f'(X) + \psi(X)f(X)\} = 0,$$

where

$$\psi(x) = \frac{p'(x)}{p(x)}.$$

For example, for the standard normal distribution $\psi(x) = -x$, and we recover the classical standard normal Stein equation. For a $\Gamma(r,\lambda)$ distribution $\psi(x) = \frac{r-1-\lambda x}{x}$, which leads to the characterisation

$$\mathbb{E}\left[f'(X) - \frac{r-1-\lambda X}{X}f(X)\right] = 0.$$

Setting $g'(x) = xf(x)$ then leads to the characterisation of Luk [44].

For a target density p which satisfies the various conditions of Proposition 1.4 of Stein et al. [74] it is straightforward to obtain a Stein characterisation for the target distribution. However, if the function $\psi(x)$ is too 'complicated' then the corresponding Stein equation may not be amenable to standard coupling techniques. When this is the case, another approach will be required to obtain a more tractable Stein equation.

For instance, Peköz et al. [54] considered the problem of obtaining a Stein operator for the family of densities

$$\kappa_s(x) = \Gamma(s)\sqrt{\frac{2}{s\pi}}\exp\bigg(-\frac{x^2}{2s}\bigg)U\bigg(s-1,\frac{1}{2},\frac{x^2}{2s}\bigg), \qquad x>0,\ s\geq\frac{1}{2},$$

where $U(a,b,x)$ denotes the confluent hypergeometric function of the second kind (also known as the Kummer U function); see Olver et al. [52], Chapter 13. The density approach would have lead to a complicated first order Stein operator involving the confluent hypergeometric function of the second kind. However, via a different approach, Peköz et al. obtained the relatively simple second order Stein operator:

$$\mathcal{A}f(x) = sf''(x) - xf'(x) - 2(s-1)f(x).$$

Recently, Ley and Swan [41] have extended the scope of the density approach. Their method allows the writing of many of the known univariate Stein characterisations, as well as some new characterisations. With their approach it may be possible to obtain second order Stein operators, which would not be possible using the direct density approach of Stein et al. [74].

In Chapter 3 and 5 we obtain Stein equations for distributions whose densities involve modified Bessel and Meijer G-functions. Like Peköz et al. [54], we arrive at Stein operators that are of order greater than one.

2.4 Smoothness Estimates for the solution of the Gamma Stein equation

We now consider the problem of obtaining smoothness estimates for the solution of the Gamma Stein equation (2.14). We draw on the work of Luk [44] and Pickett [55] to establish new bounds for the derivatives of the solution of the Gamma Stein equation. Our bounds involve the shape parameter r as well as the order, k, of the derivative.

Like Luk [44] and Pickett [55], we use the generator approach to arrive at our smoothness estimates for the solution of the Gamma Stein equation (2.14). Before writing down an expression for the solution to (2.14), we need a generalisation of the Gamma distribution (see Patnaik [53]).

Definition 2.17. *(***Non-central Gamma Distribution***) The random variable X is said to have a non-central Gamma distribution with parameters $r > 0$, $\lambda > 0$ and non-centrality parameter θ if and only if it has p.d.f. given by*

$$\hat{\gamma}_X(x; r, \lambda, \theta) = \sum_{i=0}^{\infty} \mathbb{P}(R = i)\gamma_X(x; i + r, \lambda), \qquad x > 0, \tag{2.24}$$

where $R \sim Poisson(\theta/2)$. If (2.24) holds then we write $X \sim \Gamma(r, \lambda, \theta)$.

Recall from Example 2.12 that the $\Gamma(r, \lambda)$ distribution is the stationary distribution of a Markov process given by the solution of the stochastic differential equation

$$X_t = x + \int_0^t (\lambda - rX_s)\,\mathrm{d}s + \int_0^t \sqrt{2X_s}\,\mathrm{d}B_s.$$

Luk [44] obtained the following expression for the transition semigroup $T_t h(x)$ of this process:

$$T_t h(x) = \int_0^{\infty} h((1 - \mathrm{e}^{-\lambda t})y)\hat{\gamma}(y; r, \lambda, 1 - \mathrm{e}^{-\lambda t})\,\mathrm{d}y.$$

With this formula for the semigroup, we can use the theory of Section 2.3.1 to write down the solution of the Gamma Stein equation (2.14),

$$f(x) = -\int_0^{\infty} [T_t h(x) - \Gamma_{r,\lambda} h]\,\mathrm{d}t, \tag{2.25}$$

which can easily be shown to exist by expanding out the integrand and applying the dominated convergence theorem. As is the case for the Ornstein-Uhlenbeck process, which we considered when establishing smoothness estimates for the solution of the normal Stein equation, the k-th derivative of the semigroup $T_t h(x)$ has a simple expression. This expression is given by the following lemma (see Luk [44], pp. 21–23).

Lemma 2.18. *Let $\mathcal{C}_\lambda = \{h : \mathbb{R}_+ \to \mathbb{R} \;:\; |h(x)| \le c\mathrm{e}^{ax}$ for $c \in \mathbb{R}_+$ and $a < \lambda\}$, and suppose $h \in \mathcal{C}_\lambda \cap C_b^k(\mathbb{R})$. Then*

$$(T_t h)^{(k)}(x) = \mathrm{e}^{-k\lambda t}\mathbb{E}h^{(k)}((1 - \mathrm{e}^{-\lambda t})Y_k),$$

where $Y_k \sim \Gamma(r + k, \lambda, 2\lambda\mathrm{e}^{-\lambda t}(1 - \mathrm{e}^{-\lambda t})^{-1}x)$.

We now make use of this lemma to derive the following smoothness estimates for the solution of the Gamma Stein equation (2.14).

Theorem 2.19. *Suppose $h:\mathbb{R}\to\mathbb{R}$ is an absolutely bounded functions such that $h\in\mathcal{C}_\lambda\cap C_b^k(\mathbb{R})$. Then the solution f of the $\Gamma(r,\lambda)$ Stein equation (2.14), as given by (2.25), is k times differentiable with*

$$\|f^{(j)}\|\leq\left\{\sqrt{\frac{2\pi}{r+j-1}}+\frac{\log(r+j)}{\mathrm{e}(r+j-1)}+\frac{2}{r+j-1}+\frac{1}{\mathrm{e}(r+j)}\right\}\|h^{(j-1)}\| \tag{2.26}$$

$$\leq\left\{\frac{\sqrt{2\pi}+\mathrm{e}^{-1}}{\sqrt{r+j-1}}+\frac{2}{r+j-1}\right\}\|h^{(j-1)}\|,\qquad j=1,\dots,k, \tag{2.27}$$

where $h^{(0)}\equiv h$. In the case $j=1$ the above bound can improved to

$$\|f'\|\leq\left\{\sqrt{\frac{2\pi}{r}}+\frac{2}{r}\right\}\|h\|. \tag{2.28}$$

Proof. By Lemma 2.18, for $j=1,2\dots,k$, we have $|(T_th)^{(j)}(x)|\leq\|h^{(j)}\|\mathrm{e}^{-j\lambda t}$. As $\mathrm{e}^{-j\lambda t}$ is integrable on $[0,\infty)$, by the dominated convergence theorem,

$$f^{(j)}(x)=-\int_0^\infty(T_th)^{(j)}(x)\,\mathrm{d}t. \tag{2.29}$$

We now concentrate on the case $j=1$ to prove inequality (2.28). We will then consider the case $j\geq 2$ to prove inequalities (2.26) and (2.27). Using Definition 2.17, Lemma 2.18 and equation (2.29), we have

$$\begin{aligned}f'(x)&=-\int_0^\infty\int_0^\infty\sum_{i=0}^\infty\frac{1}{i!}\mathrm{e}^{-\lambda t}h'((1-\mathrm{e}^{-\lambda t})y)\exp\left\{-\frac{\lambda\mathrm{e}^{-\lambda t}}{1-\mathrm{e}^{-\lambda t}}x\right\}\\&\quad\times\left(\frac{\lambda\mathrm{e}^{-\lambda t}}{1-\mathrm{e}^{-\lambda t}}x\right)^i\gamma(y;r+i+1,\lambda)\,\mathrm{d}y\,\mathrm{d}t\\&=-\sum_{i=0}^\infty\frac{1}{i!}\int_0^\infty\int_0^\infty\mathrm{e}^{-\lambda t}h'((1-\mathrm{e}^{-\lambda t})y)\exp\left\{-\frac{\lambda\mathrm{e}^{-\lambda t}}{1-\mathrm{e}^{-\lambda t}}x\right\}\\&\quad\times\left(\frac{\lambda\mathrm{e}^{-\lambda t}}{1-\mathrm{e}^{-\lambda t}}x\right)^i\mathrm{d}t\,\gamma(y;r+i+1,\lambda)\,\mathrm{d}y,\end{aligned}$$

where the interchange of summation and integration is justified in Pickett [55], Lemma 7.1. Substituting $s=\mathrm{e}^{-\lambda t}$ gives

$$f'(x)=-\frac{1}{\lambda}\sum_{i=0}^\infty\frac{1}{i!}\int_0^\infty\int_0^1h'((1-s)y)\exp\left\{-\frac{s}{1-s}\lambda x\right\}\left(\frac{s}{1-s}\lambda x\right)^i\mathrm{d}s\,\gamma(y;r+i+1,\lambda)\,\mathrm{d}t.$$

Integration by parts, and using that h is bounded and that $\lim_{s\uparrow1}\exp\{-\frac{s}{1-s}\lambda x\}(\frac{s}{1-s}\lambda x)^i=0$

for $\lambda > 0$, $x > 0$ and $i \geq 0$, gives

$$\begin{aligned} f'(x) &= -\frac{1}{\lambda}\sum_{i=0}^{\infty}\frac{1}{i!}\int_0^{\infty}\bigg\{\bigg[-h((1-s)y)\exp\bigg\{-\frac{s}{1-s}\lambda x\bigg\}\bigg(\frac{s}{1-s}\lambda x\bigg)^i\bigg]_0^1 \\ &\quad + \int_0^1 h((1-s)y)\frac{\mathrm{d}}{\mathrm{d}s}\bigg[\exp\bigg\{-\frac{s}{1-s}\lambda x\bigg\}\bigg(\frac{s}{1-s}\lambda x\bigg)^i\bigg]\,\mathrm{d}s\bigg\}\frac{\gamma(y;r+i+1,\lambda)}{y}\,\mathrm{d}y \\ &= \frac{1}{\lambda}\int_0^{\infty}h(y)\frac{\gamma(y;r+1,\lambda)}{y}\,\mathrm{d}y - \frac{1}{\lambda}\sum_{i=0}^{\infty}\frac{1}{i!}\int_0^{\infty}\int_0^1 h((1-s)y) \\ &\quad \times \frac{\mathrm{d}}{\mathrm{d}s}\bigg[\exp\bigg\{-\frac{s}{1-s}\lambda x\bigg\}\bigg(\frac{s}{1-s}\lambda x\bigg)^i\bigg]\,\mathrm{d}s\,\frac{\gamma(y;r+i+1,\lambda)}{y}\,\mathrm{d}y. \end{aligned}$$

Substituting $u = \frac{s}{1-s}\lambda x$ gives

$$\begin{aligned} f'(x) &= \frac{1}{\lambda}\int_0^{\infty}h(y)\frac{\gamma(y;r+1,\lambda)}{y}\,\mathrm{d}y \\ &\quad - \frac{1}{\lambda}\sum_{i=0}^{\infty}\frac{1}{i!}\int_0^{\infty}\int_0^{\infty}h\bigg(\frac{\lambda xy}{u+\lambda x}\bigg)\frac{\mathrm{d}}{\mathrm{d}u}\{\mathrm{e}^{-u}u^i\}\,\mathrm{d}u\,\frac{\gamma(y;r+i+1,\lambda)}{y}\,\mathrm{d}y, \end{aligned} \tag{2.30}$$

by use of the chain rule. The first term of (2.30) is easily bounded:

$$\frac{1}{\lambda}\bigg|\int_0^{\infty}h(y)\frac{\gamma(y;r+1,\lambda)}{y}\,\mathrm{d}y\bigg| \leq \frac{\|h\|}{\lambda}\int_0^{\infty}\frac{\gamma(y;r+1,\lambda)}{y}\,\mathrm{d}y = \frac{\|h\|}{r},$$

since $\mathbb{E}(X^{-1}) = \frac{\beta}{\alpha-1}$ if $X \sim \Gamma(\alpha,\beta)$, for $\alpha > 1$ and $\beta > 0$. We now concentrate on bounding the second term of equation (2.30). Now, for $i \geq 1$,

$$\frac{\mathrm{d}}{\mathrm{d}u}\{\mathrm{e}^{-u}u^i\} = (i-u)\mathrm{e}^{-u}u^{i-1},$$

and hence the derivative is strictly positive on $(0,i)$ and strictly negative on (i,∞). This allows us to bound the inner integral for $i \geq 1$ as follows.

$$\begin{aligned} &\bigg|\int_0^{\infty}h\bigg(\frac{\lambda xy}{u+\lambda x}\bigg)\{\mathrm{e}^{-u}u^i\}'\,\mathrm{d}u\bigg| \\ &= \bigg|\int_0^{i}h\bigg(\frac{\lambda xy}{u+\lambda x}\bigg)\{\mathrm{e}^{-u}u^i\}'\,\mathrm{d}u + \int_i^{\infty}h\bigg(\frac{\lambda xy}{u+\lambda x}\bigg)\{\mathrm{e}^{-u}u^i\}'\,\mathrm{d}u\bigg| \\ &\leq \|h\|\bigg\{\int_0^{i}(\mathrm{e}^{-u}u^i)'\,\mathrm{d}u - \int_i^{\infty}(\mathrm{e}^{-u}u^i)'\,\mathrm{d}u\bigg\} \\ &= 2\|h\|\mathrm{e}^{-i}i^i. \end{aligned} \tag{2.31}$$

For $i = 0$, the inner integral is bounded by $\|h\|$. Hence,

$$\begin{aligned}|f'(x)| &\leq \frac{\|h\|}{r} + \frac{\|h\|}{\lambda}\bigg\{\int_0^\infty \frac{\gamma(y;r+1,\lambda)}{y}\,\mathrm{d}y + 2\sum_{i=1}^\infty \frac{\mathrm{e}^{-i}i^i}{i!}\int_0^\infty \frac{\gamma(y;r+i+1,\lambda)}{y}\,\mathrm{d}y\bigg\}\\ &= 2\|h\|\bigg\{\frac{1}{r} + \sum_{i=1}^\infty \frac{\mathrm{e}^{-i}i^i}{i!\,(r+i)}\bigg\},\end{aligned}$$

where we, again, used that $\mathbb{E}(X^{-1}) = \frac{\beta}{\alpha-1}$ if $X \sim \Gamma(\alpha,\beta)$, for $\alpha > 1$ and $\beta > 0$.

The sum above is not expressible in closed form, and so, we apply Stirling's approximation $n! \geq n^n\mathrm{e}^{-n}\sqrt{2\pi n}$ (see, for example, Robbins [64]) to obtain

$$2\sum_{i=1}^\infty \frac{\mathrm{e}^{-i}i^i}{i!\,(r+i)} \leq \sqrt{\frac{2}{\pi}}\sum_{i=1}^\infty \frac{1}{\sqrt{i}(r+i)} \tag{2.32}$$

$$\leq \sqrt{\frac{2}{\pi}}\int_0^\infty \frac{1}{\sqrt{t}(r+t)}\,\mathrm{d}t \tag{2.33}$$

$$= \sqrt{\frac{2\pi}{r}}, \tag{2.34}$$

where we used the integral test to obtain the final inequality. This completes the proof for the case $j = 1$.

We now consider the case $j \geq 2$. We have

$$\begin{aligned}f^{(j)}(x) = -\sum_{i=0}^\infty \frac{1}{i!}\int_0^\infty\int_0^\infty \mathrm{e}^{-j\lambda t}h^{(j)}((1-\mathrm{e}^{-\lambda t})y)\exp\bigg\{-\frac{\lambda\mathrm{e}^{-\lambda t}}{1-\mathrm{e}^{-\lambda t}}x\bigg\}\\ \times\bigg(\frac{\lambda\mathrm{e}^{-\lambda t}}{1-\mathrm{e}^{-\lambda t}}x\bigg)^i\,\mathrm{d}t\,\gamma(y;r+i+j,\lambda)\,\mathrm{d}y,\end{aligned}$$

which, substituting $s = \mathrm{e}^{-\lambda t}$, becomes

$$f^{(j)}(x) = -\frac{1}{\lambda}\sum_{i=0}^\infty \frac{1}{i!}\int_0^\infty\int_0^1 h^{(j)}((1-s)y)\exp\bigg\{-\frac{s}{1-s}\lambda x\bigg\}\bigg(\frac{s}{1-s}\lambda x\bigg)^i s^{j-1}\,\mathrm{d}s\,\gamma(y;r+i+j,\lambda)\,\mathrm{d}y.$$

Integrating by parts as before gives

$$\begin{aligned}f^{(j)}(x) = -\frac{1}{\lambda}\sum_{i=0}^\infty \frac{1}{i!}\int_0^\infty\int_0^1 h^{(j-1)}((1-s)y)\frac{\mathrm{d}}{\mathrm{d}s}\bigg[\exp\bigg\{-\frac{s}{1-s}\lambda x\bigg\}\bigg(\frac{s}{1-s}\lambda x\bigg)^i s^{j-1}\bigg]\,\mathrm{d}s\\ \times\gamma(y;r+i+j,\lambda)\,\mathrm{d}y,\end{aligned}$$

which, substituting $u = \frac{s}{1-s}\lambda x$, becomes

$$f^{(j)}(x) = -\frac{1}{\lambda}\sum_{i=0}^{\infty}\frac{1}{i!}\int_0^\infty\int_0^\infty h^{(j-1)}\left(\frac{\lambda xy}{u+\lambda x}\right)\frac{\mathrm{d}}{\mathrm{d}u}\left[\mathrm{e}^{-u}u^i\left(\frac{u}{u+\lambda x}\right)^{j-1}\right]\mathrm{d}u\,\frac{\gamma(y;r+i+j,\lambda)}{y}\,\mathrm{d}y.$$

We now concentrate on bounding the inner integral, which is given by

$$\begin{aligned} &\int_0^\infty h^{(j-1)}\left(\frac{\lambda xy}{u+\lambda x}\right)\left(\frac{u}{u+\lambda x}\right)^{j-1}\frac{\mathrm{d}}{\mathrm{d}u}\{\mathrm{e}^{-u}u^i\}\,\mathrm{d}u \qquad (A)\\ +&\int_0^\infty h^{(j-1)}\left(\frac{\lambda xy}{u+\lambda x}\right)\mathrm{e}^{-u}u^i\frac{\mathrm{d}}{\mathrm{d}u}\left[\left(\frac{u}{u+\lambda x}\right)^{j-1}\right]\mathrm{d}u. \qquad (B)\end{aligned}$$

To bound (A) we use that $\frac{u}{u+\lambda x} < 1$. and a similar argument to the one used to bound the inner integral in the case $j = 1$. This gives $|(A)| \leq \|h^{(j-1)}\|$ for $i = 0$, and $|(A)| \leq 2\|h^{(j-1)}\|e^i i^i$ for $i \geq 1$.

We now bound (B), and suppose $x > 0$, because $(B) = 0$ for $x = 0$. For $i = 0$, we have

$$\begin{aligned}|(B)| &= (j-1)\int_0^\infty h^{(j-1)}\left(\frac{\lambda xy}{u+\lambda x}\right)\mathrm{e}^{-u}\frac{u^{j-2}\lambda x}{(u+\lambda x)^j}\,\mathrm{d}u\bigg|\\ &\leq (j-1)\|h^{(j-1)}\|\int_0^\infty\frac{u^{j-2}\lambda x}{(u+\lambda x)^j}\,\mathrm{d}u\\ &= \|h^{(j-1)}\|,\end{aligned}$$

where we used that $\int_0^\infty \frac{t^a}{(t+b)^{a+2}}\,\mathrm{d}t = \frac{1}{b}\int_0^1(1-v)^a\,\mathrm{d}v = \frac{1}{(a+1)b}$ for $a > -1$ and $b > 0$ (and we used the substitution $v = \frac{b}{t+b}$).

For $i \geq 1$, we have

$$|(B)| \leq (j-1)\|h^{(j-1)}\|\left|\int_0^\infty \mathrm{e}^{-u}\frac{u^{i+j-2}\lambda x}{(u+\lambda x)^j}\,\mathrm{d}u\right|.$$

From calculus we have that for $u > 0$ and $x > 0$ the function $\frac{\lambda x}{(u+\lambda x)^j}$ has a maximum at $\lambda x = \frac{u}{j-1}$. Hence,

$$\frac{\lambda x}{(u+\lambda x)^j} \leq \frac{u}{(j-1)\left(u+\frac{u}{j-1}\right)^j} = \frac{u^{1-j}}{(j-1)\left(1+\frac{1}{j-1}\right)^j} < \frac{u^{1-j}}{\mathrm{e}(j-1)}.$$

Therefore, for $i \geq 1$, we have

$$|(B)| \leq \frac{\|h^{(j-1)}\|}{e}\int_0^\infty \mathrm{e}^{-u}u^{i-1}\,\mathrm{d}u = \frac{\|h^{(j-1)}\|}{\mathrm{e}}(i-1)!\,.$$

Using our bounds on the inner integrals gives

$$\begin{aligned}\|f^{(j)}\| &\leq \frac{\|h^{(j-1)}\|}{\lambda}\bigg\{2\int_0^\infty \frac{\gamma(y;r+j,\lambda)}{y}\,\mathrm{d}y + \bigg\{2\sum_{i=1}^\infty \frac{\mathrm{e}^{-i}i^i}{i!} + \frac{1}{\mathrm{e}}\sum_{i=1}^\infty \frac{1}{i}\bigg\}\int_0^\infty \frac{\gamma(y;r+i+j,\lambda)}{y}\,\mathrm{d}y\bigg\} \\ &= \|h^{(j-1)}\|\bigg\{\frac{2}{r+j-1} + 2\sum_{i=1}^\infty \frac{\mathrm{e}^{-i}i^i}{i!\,(r+i+j-1)} + \frac{1}{\mathrm{e}}\sum_{i=1}^\infty \frac{1}{i(r+i+j-1)}\bigg\}.\end{aligned}$$

From inequality (2.34) we have

$$2\sum_{i=1}^\infty \frac{\mathrm{e}^{-i}i^i}{i!\,(r+i+j-1)} \leq \sqrt{\frac{2\pi}{r+j-1}}.$$

From inequality (2.34) we also have

$$\sum_{i=1}^\infty \frac{1}{i(r+i+j-1)} \leq \sum_{i=1}^\infty \frac{1}{\sqrt{i}(r+i+j-1)} \leq \frac{\pi}{\sqrt{r+j-1}}.$$

This completes the proof of inequality (2.26). The above bound can be improved;

$$\begin{aligned}\sum_{i=1}^\infty \frac{1}{i(r+i+j-1)} &= \frac{1}{r+j} + \sum_{i=2}^\infty \frac{1}{i(r+i+j-1)} \\ &\leq \frac{1}{r+j} + \int_1^\infty \frac{1}{t(t+r+j-1)}\,\mathrm{d}t \\ &= \frac{1}{r+j} + \frac{\log(r+j)}{r+j-1},\end{aligned}$$

and this yields inequality (2.27). □

Hence, the smoothness estimates of the solution of the $\chi^2_{(p)}$ Stein equation (2.15) are given by

$$\begin{aligned}\|f^{(k)}\| &\leq \bigg\{\frac{\sqrt{2\pi}+\mathrm{e}^{-1}}{\sqrt{2p+k-1}} + \frac{2}{2p+k-1}\bigg\}\|h^{(k-1)}\| \\ &\leq \bigg\{\frac{\sqrt{2\pi}+\mathrm{e}^{-1}}{\sqrt{2p}} + \frac{1}{p}\bigg\}\|h^{(k-1)}\|, \qquad k \geq 1. \end{aligned} \tag{2.35}$$

We end this section with some comments on the bounds of Theorem 2.19.

Remark 2.20. The bounds given in Theorem 2.19 are of order $r^{-1/2}$ as $r \to \infty$, but could we achieve a faster convergence rate? It is worth noting that we cannot do better than order r^{-1} for the bound on the first derivative of the solution of the Stein equation. This can be seen by evaluating both sides of the Gamma Stein equation (2.14) at $x = 0$, which gives

$$f'(0) = \frac{1}{r}[h(0) - \Gamma_{r,\lambda}h].$$

In the proof of (2.28) we made use of three inequalities, and in the proofs of (2.26) and (2.27) we used five inequalities. We now examine each of the inequalities used in the proof of (2.28) to determine whether the use of that inequality may have resulted in a slower convergence rate. (We consider this case for simplicity, but similar comments will apply to proofs of (2.26) and (2.27).) To obtain inequality (2.32) we used Stirling's approximation. This is a very good approximation and would not have resulted in a slower convergence rate. To arrive at inequality (2.33) we used the integral test. From the integral test, we have

$$\frac{2\tan^{-1}(r)}{\sqrt{r}} = \int_0^\infty \frac{1}{\sqrt{x+1}(x+r+1)}\,\mathrm{d}x \leq \sum_{i=1}^\infty \frac{1}{\sqrt{i}(r+i)} \leq \int_0^\infty \frac{1}{\sqrt{x}(x+r)}\,\mathrm{d}x = \frac{\pi}{\sqrt{r}}.$$

Hence,

$$\lim_{r\to\infty}\sum_{i=1}^\infty \frac{\sqrt{r}}{\sqrt{i}(r+i)} = \pi,$$

and so this inequality could also not have resulted in a slower convergence rate. Inequality (2.31) certainly seems rather crude, and therefore it would perhaps not be surprising if the the true of order of convergence is faster than order $r^{-1/2}$. Therefore in order to achieve a faster convergence rate than order $r^{-1/2}$ (provided that $O(r^{-1/2})$ is not the optimal rate) we would need to argue more carefully at this point of the proof.

Remark 2.21. The bounds of Theorem 2.19 involve the shape parameter r, but not the scale parameter λ, whereas bound (2.22) of Luk [44] involves only the scale parameter λ. It is an unsolved problem to obtain a bound on the smoothness estimates that involves both the shape and scale parameters, although in practice we can, of course, use the minimum of the bounds of Theorem 2.19 and the bounds of Luk [44]. We now give some insight into this problem and show that it is not possible to extend the bounds of Theorem 2.19 to involve the scale parameter; that is, for a bound of the form $\|f^{(k)}\| \leq C\|h^{(k-1)}\|$, (i.e. a bound only involving the $(k-1)th$ derivative of h) the constant C may depend on r and k, but will not involve λ.

Recall that the Gamma Stein equation is given by

$$xf''(x) + (r - \lambda x)f'(x) = h(x) - \Gamma_{r,\lambda}h.$$

Substituting $y = \lambda x$ and $f(x) = g(\lambda x) = g(y)$ gives

$$yg''(y) + (r - y)g'(y) = \tilde{h}(y), \tag{2.36}$$

where

$$\tilde{h}(y) = \frac{1}{\lambda}\left[h\left(\frac{y}{\lambda}\right) - \Gamma_{r,\lambda}h\right].$$

Notice that the left-hand side of (2.36) is independent of the scale parameter λ. Therefore bounds on the k-th derivative of the solution g of (2.36) that involve only the sup norm of $(k-1)$-th derivative of the function $\tilde{h}$ will be of the form $\|g^{(k)}\| \leq C(r,k)\|\tilde{h}^{(k-1)}\|$. Therefore,

$$\|f^{(k)}\| = \lambda^k\|g^{(k)}\| \leq \lambda^k C(r,k)\|\tilde{h}^{(k-1)}\| = C(r,k)\|h^{(k-1)}\|,$$

and so the bound is independent of λ.

Chapter 3

Stein's Method for Variance-Gamma distributions

In this chapter we shall consider Stein's method for the class of Variance-Gamma distributions. We begin by deriving a Stein equation for the Variance-Gamma distributions. We obtain the unique bounded solution of the Stein equation and establish bounds for the solution of the Symmetric-Variance Gamma distribution and its first four derivatives in terms of supremum norms of the test function h. Our limit theorems in Chapters 4 and 5 require uniform bounds for the first four derivatives of the solution to the Symmetric Variance-Gamma Stein equation, and so our bounds for the first four derivatives are sufficient for this purpose. We end this chapter by discussing how the generator approach may enable us to obtain bounds for derivatives of general order.

3.1 The class of Variance-Gamma distributions

Here we present the *Variance-Gamma* distributions and some of their basic properties which will be of use to us throughout this thesis. Variance-Gamma distributions are used in applications in financial modelling and were introduced into the financial literature by Madan and Seneta [46]. Variance-Gamma random variables have a simple characterisation in terms of normal and Gamma random variables, and perhaps a more natural name for the full class is the *Normal-Gamma* distributions. Throughout this thesis we will make use of two different parametrisations of the Variance-Gamma distribution; we present both of these below. The first parametrisation is similar to that of Finlay and Seneta [26], and the second parameterisation can be found in Eberlein and Hammerstein [21].

Definition 3.1 (**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 p.d.f. given by*

$$p_{\mathrm{VG}}(x; r, \theta, \sigma, \mu) = \frac{1}{\sigma\sqrt{\pi}\Gamma(\frac{r}{2})} \mathrm{e}^{\frac{\theta}{\sigma^2}(x-\mu)} \left(\frac{|x-\mu|}{2\sqrt{\theta^2+\sigma^2}}\right)^{\frac{r-1}{2}} K_{\frac{r-1}{2}}\left(\frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x-\mu|\right), \tag{3.1}$$

where $x \in \mathbb{R}$. If (3.1) holds then we write $X \sim \mathrm{VG}(r, \theta, \sigma, \mu)$.

Definition 3.2 (**Variance-Gamma Distribution**, **second parametrisation**). *The random variable X is said to have a* Variance-Gamma *distribution with parameters ν, α, β, μ, where $\nu > -1/2$, $\mu \in \mathbb{R}$, $\alpha > |\beta| \geq 0$, if and only if it has p.d.f. given by*

$$p_{\mathrm{VG}_2}(x; \nu, \alpha, \beta, \mu) = \frac{(\alpha^2 - \beta^2)^{\nu+1/2}}{\sqrt{\pi}\Gamma(\nu + \frac{1}{2})} \left(\frac{|x-\mu|}{2\alpha}\right)^{\nu} \mathrm{e}^{\beta(x-\mu)} K_{\nu}(\alpha|x-\mu|), \quad x \in \mathbb{R}. \tag{3.2}$$

If (3.2) holds then we write $X \sim \mathrm{VG}_2(\nu, \alpha, \beta, \mu)$.

Definition 3.3. *If $X \sim \mathrm{VG}(r, 0, \sigma, \mu)$, for r, σ, and μ defined as in Definition 3.1 (or equivalently $X \sim \mathrm{VG}_2(\nu, \alpha, 0, \mu)$), then X is said to have a* Symmetric Variance-Gamma *distribution.*

Remark 3.4. The density (3.1) may at first appear to be undefined in the limit $\sigma \to 0$, 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.11) from Appendix B). As we shall see in Proposition 3.8 (below), taking the limit $\sigma \to 0$ and putting $\mu = 0$ gives the family of Gamma distributions.

Also, from the asymptotic formula (B.8) for $K_\nu(x)$, we can deduce the following asymptotics for the density (3.1) in the limit $x \to \mu$. For $\sigma > 0$ we have

$$p_{\mathrm{VG}}(x; r, \theta, \sigma, \mu) \sim \begin{cases} \dfrac{\sigma^{r-2}\Gamma(\frac{r-1}{2})}{2\sqrt{\pi}\Gamma(\frac{r}{2})(\theta^2+\sigma^2)^{\frac{r-1}{2}}}, & x \to \mu,\ r > 1, \\ -\dfrac{1}{\pi\sigma}\log|x-\mu|, & x \to \mu,\ r = 1, \\ \dfrac{1}{(2\sigma)^r\sqrt{\pi}} \dfrac{\Gamma(\frac{1-r}{2})}{\Gamma(\frac{r}{2})}|x-\mu|^{r-1}, & x \to \mu,\ 0 < r < 1. \end{cases}$$

In Chapter 4 we shall use a Stein equation for the Variance-Gamma distributions in terms of the first parametrisation to prove some Variance-Gamma limit theorems. However, since the second parametrisation takes a slightly simpler form we will make use of it throughout this chapter. We can then obtain the results for the first parametrisation, that we had obtained using the second parametrisation, by using the fact that the parameters are related by

$$\nu = \frac{r-1}{2}, \qquad \alpha = \frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}, \qquad \beta = \frac{\theta}{\sigma^2}. \tag{3.3}$$

The Variance-Gamma distribution has moments of arbitrary order, in particular the mean and variance (for both parametrisations) of a random variable X with a Variance-Gamma distribution are given by (see Eberlein and Hammerstein [21], formula 3.22):

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

The following proposition, which can be found in Bibby and Sørensen [14], shows that the class of Variance-Gamma distributions is closed under convolution.

Proposition 3.5. *The class of Variance-Gamma distributions is closed under convolution. If X_1 and X_2 are independent random variables such that $X_i \sim \mathrm{VG}(r_i,\theta,\sigma,\mu_i)$, $i=1,2$, then we have that*

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

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. [11].

Proposition 3.6. *Let $r>0$, $\theta\in\mathbb{R}$, $\sigma>0$ and $\mu\in\mathbb{R}$. Suppose that U and V are independent random variables and that $U\sim N(0,1)$, $V\sim\Gamma(r,\lambda)$ and $Z\sim\mathrm{VG}(2r,\theta(2\lambda)^{-1},\sigma\sqrt{2\lambda},\mu)$, then*

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

Corollary 3.7. *Let $\theta\in\mathbb{R}$, $\sigma>0$, $\mu\in\mathbb{R}$, and r be a positive integer. Let $X_1,X_2,\ldots,X_r$ and $Y_1,Y_2,\ldots,Y_r$ be independent standard normal random variables and let Z be a $\mathrm{VG}(r,\theta,\sigma,\mu)$ random variable, then*

$$Z \stackrel{\mathcal{D}}{=} \mu + \theta\sum_{i=1}^r X_i^2 + \sigma\sum_{i=1}^r X_iY_i. \tag{3.4}$$

Proof. Let $X_1,X_2,\ldots,X_r$ and $Y_1,Y_2,\ldots,Y_r$ be independent standard normal random variables.

Then X_i^2, $i=1,2,\ldots,m$ has a $\chi^2_{(1)}$ distribution, that is a $\Gamma(\frac{1}{2},\frac{1}{2})$ distribution. Define

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

Note that $X_iY_i \stackrel{\mathcal{D}}{=} |X_i|Y_i$. Hence, by Proposition 3.6, we have that Z_1 is a $\mathrm{VG}(1,\theta,\sigma,\mu)$ random variable and Z_i, $i=1,2,\ldots,m$, are $\mathrm{VG}(1,\theta,\sigma,0)$ random variables. It therefore follows from Proposition 3.5 that the random variable $Z=\sum_{i=1}^r Z_i$ has a $\mathrm{VG}(r,\theta,\sigma,\mu)$ distribution. $\square$

The representation (3.4) for Variance-Gamma distributions is particularly useful in helping us decide which statistics are suitable candidates to be approximated by Variance-Gamma distributions via Stein's method. From Corollary 3.7 we see that if $X_1,X_2,\ldots,X_r$ and $Y_1,Y_2,\ldots,Y_r$ are a sequence of independent standard normal random variables then the statistic $Z=\sum_{i=1}^r X_iY_i$ has a $\mathrm{VG}(r,0,1,0)$ distribution. This suggests that the D_2^* statistic, defined in Section 1.1, may have an approximate $\mathrm{VG}(r,0,1,0)$ distribution for certain parameter values.

In the following proposition we provide a list of characterisations of the Variance-Gamma distributions in terms of well-known distributions, which would also motivate the study of the asymptotic behaviour of various statistics via Stein's method for Variance-Gamma distributions. In particular, we note that Laplace distribution is a member of the class of Variance-Gamma distributions and that the normal and Gamma distributions occur as limiting cases. The difference of two dependent Gamma distributions and the product of two dependent normal distributions are also seen to be in the class of Variance-Gamma distributions.

Proposition 3.8. *(i) Suppose $\sigma>0$ and $\mu\in\mathbb{R}$, then* $\lim_{r\to\infty}\mathrm{VG}(r,0,\sigma/\sqrt{r},\mu)=N(\mu,\sigma^2)$.

(ii) Suppose $\sigma>0$ and $\mu\in\mathbb{R}$, then a $\mathrm{VG}(2,0,\sigma,\mu)$ random variable has the Laplace distribution with p.d.f.

$$p_{\mathrm{VG}}(x;2,0,\sigma,\mu)=\frac{1}{2\sigma}\exp\bigg(-\frac{|x-\mu|}{\sigma}\bigg), \quad x\in\mathbb{R}.$$

(iii) Suppose that $X\sim N(0,\sigma_X^2)$ and $Y\sim N(0,\sigma_Y^2)$ have correlation ρ. Let $Z=XY$, then $Z\sim\mathrm{VG}(1,\rho\sigma_X\sigma_Y,\sigma_X\sigma_Y\sqrt{1-\rho^2},0)$.

(iv) The Gamma distribution is a limiting case of the Variance-Gamma distribution: for $r>0$ and $\lambda>0$ we have

$$\lim_{\sigma\to 0} p_{\mathrm{VG}}(x;2r,(2\lambda)^{-1},\sigma,0)=\begin{cases}\dfrac{\lambda^r}{\Gamma(r)}x^{r-1}\mathrm{e}^{-\lambda x}, & x>0,\\ 0, & x<0.\end{cases} \tag{3.5}$$

(v) Suppose that $U\sim\Gamma(r,\lambda_1)$ and $V\sim\Gamma(r,\lambda_2)$ have correlation ρ. Let $Z=U-V$, then

$Z \sim \mathrm{VG}(2r, (2\lambda_1)^{-1} - (2\lambda_2)^{-1}, (\lambda_1\lambda_2)^{-1/2}(1-\rho)^{1/2}, 0)$.

Proof. (i) Let $X_1, X_2, \ldots$ and $Y_1, Y_2 \ldots$ be i.i.d. standard normal random variables. Let $Z = \mu + \frac{\sigma}{\sqrt{r}}\sum_{i=1}^r X_iY_i$. Then by Corollary 3.7 it follows that $Z \sim \mathrm{VG}(r, 0, \sigma/\sqrt{r}, \mu)$. Moreover, $Z_i = X_iY_i$, $i = 1, 2, \ldots$ are i.i.d. random variables with mean zero and unit variance, and by the central limit theorem $\frac{1}{\sqrt{r}}\sum_{i=1}^r X_iY_i \xrightarrow{\mathcal{D}} N(0,1)$. Hence, $Z \xrightarrow{\mathcal{D}} N(\mu, \sigma)$ as $r \to \infty$.

(ii) This follows by applying the formula $K_{1/2}(x) = \sqrt{\frac{\pi}{2x}}\mathrm{e}^{-x}$ to the density (3.1).

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

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

We can express Z in terms of $\tilde{X}$ and W as follows

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

Now W is independent of X (see Mardia et al. [47] p. 63), and it is straightforward to show that W has a standard normal distribution. Hence, by Corollary 3.7, it follows that $Z \sim \mathrm{VG}(1, \rho\sigma_X\sigma_Y, \sigma_X\sigma_Y\sqrt{1-\rho^2}, 0)$.

(iv) Suppose $\theta > 0$. By (3.1), we have

$$p_{\mathrm{VG}}(x; 2r, \theta, \sigma, 0) = \frac{1}{\sigma\sqrt{\pi}\Gamma(r)}\mathrm{e}^{\frac{\theta}{\sigma^2}x}\left(\frac{|x|}{2\sqrt{\theta^2+\sigma^2}}\right)^{r-\frac{1}{2}}K_{r-\frac{1}{2}}\left(\frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x|\right).$$

By the asymptotic formula (B.11), we have

$$K_{r-\frac{1}{2}}\left(\frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x|\right) \sim \frac{\sigma}{\sqrt{\theta^2+\sigma^2}}\sqrt{\frac{\pi}{2|x|}}\exp\left(-\frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x|\right), \qquad \sigma \to 0,$$

and consequently for σ small enough

$$p_{\mathrm{VG}}(x; 2r, \theta, \sigma, 0) \sim \frac{1}{2^r\Gamma(r)(\theta^2+\sigma^2)^{r/2}}|x|^{r-1}\exp\left(\frac{\theta}{\sigma^2}x - \frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x|\right).$$

If $x > 0$, then by Taylor series expansion,

$$\frac{\theta}{\sigma^2}x - \frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x| = \frac{\theta x}{\sigma^2}\left(1 - \sqrt{1+\frac{\sigma^2}{\theta^2}}\right) = -\frac{x}{2\theta} + O(\sigma^2), \qquad \sigma \to 0,$$

and if $x<0$, since $\theta>0$, we have

$$\frac{\theta}{\sigma^2}x-\frac{\sqrt{\theta^2+\sigma^2}}{\sigma^2}|x|=\frac{\theta x}{\sigma^2}\left(1+\sqrt{\theta^2+\sigma^2}\right)\rightarrow-\infty,\qquad\sigma\rightarrow 0.$$

Hence,

$$\lim_{\sigma\rightarrow 0}p_{\mathrm{VG}}(x;2r,\theta,\sigma,0)=\begin{cases}\dfrac{1}{(2\theta)^r\Gamma(r)}x^{r-1}\exp\left(-\dfrac{x}{2\theta}\right), & x>0,\\ 0, & x<0.\end{cases}\tag{3.6}$$

Substituting $\theta=(2\lambda)^{-1}$ into (3.6) gives (3.5), as required.

(v) Theorem 6 of Holm and Alouini [36] gives the following formula for the p.d.f. of $Z=U-V$:

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

where

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

We can write the probability density function of Z as follows

$$\begin{aligned}p_Z(x)&=\frac{|x|^{r-1/2}}{\Gamma(r)\sqrt{\pi}\sqrt{\beta_1\beta_2(1-\rho)}}\left(\frac{1}{(\beta_1+\beta_2)^2-4\beta_1\beta_2\rho}\right)^{\frac{2r-1}{4}}\exp\left(x\cdot\frac{\frac{1}{2}(\beta_1-\beta_2)}{\beta_1\beta_2(1-\rho)}\right)\\&\quad\times K_{r-\frac{1}{2}}\left(|x|\frac{\sqrt{[(\beta_1-\beta_2)/2]^2-\beta_1\beta_2(1-\rho)}}{\beta_1\beta_2(1-\rho)}\right),\qquad x\in\mathbb{R}.\end{aligned}\tag{3.7}$$

Comparing (3.7) with the Variance-Gamma p.d.f. (3.1), we see that Z has a $\mathrm{VG}(2r,\theta,\sigma,0)$ 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. $\square$

Remark 3.9. An alternative proof of part (i) of Proposition 3.8 involves the use of the asymptotic properties of modified Bessel functions of large orders, which are given in Stadje [71].

3.2 A Stein Equation for the Variance-Gamma distributions

The following lemma which characterizes the Variance-Gamma distribution, for the two different parametrisations, will lead to a Stein equation for the Variance-Gamma distribution.

Lemma 3.10. *Let Z be a real-valued random variable. Then $\mathscr{L}(Z) = \mathrm{VG}_2(\nu, \alpha, \beta, \mu)$ if and only if, for all $f : \mathbb{R} \to \mathbb{R}$ such that f is piecewise twice continuously differentiable, with f and f' bounded, and $\mathbb{E}|Wf''(W)|$, $\mathbb{E}|f'(W)|$, $\mathbb{E}|Wf'(W)|$, $\mathbb{E}|f(W)|$ and $\mathbb{E}|Wf(W)|$ are finite for $W \sim \mathrm{VG}_2(\nu, \alpha, \beta, \mu)$,*

$$\mathbb{E}\{(Z-\mu)f''(Z) + (2\nu+1+2\beta(Z-\mu))f'(Z) + ((2\nu+1)\beta - (\alpha^2-\beta^2)(Z-\mu))f(Z)\} = 0. \quad (3.8)$$

Proof. To simplify the calculations we prove the result for the special case $\mu = 0$, $\alpha = 1$, $-1 < \beta < 1$. We then extend the result to the general case, stated in the lemma, by a simple linear transformation.

Necessity. Suppose that $W \sim \mathrm{VG}_2(\nu, 1, \beta, 0)$. The conditions in the lemma ensure that the expectations $\mathbb{E}Wf''(W)$, $\mathbb{E}f'(W)$, $\mathbb{E}Wf'(W)$, $\mathbb{E}f(W)$ and $\mathbb{E}Wf(W)$ are well defined. We denote the normalizing constant for the density function by M and split the range of integration to obtain

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

where

$$I_1 = \int_{-\infty}^{0} \{xf''(x) + (2\nu+1+2\beta x)f'(x) + ((2\nu+1)\beta - (1-\beta^2)x)f(x)\}\mathrm{e}^{\beta x}(-x)^{\nu}K_{\nu}(-x)\,\mathrm{d}x,$$
$$I_2 = \int_{0}^{\infty} \{xf''(x) + (2\nu+1+2\beta x)f'(x) + ((2\nu+1)\beta - (1-\beta^2)x)f(x)\}\mathrm{e}^{\beta x}x^{\nu}K_{\nu}(x)\,\mathrm{d}x.$$

We begin by considering I_2. Applying integration by parts we obtain

$$\begin{aligned}\int_0^{\infty} f''(x)x^{\nu+1}\mathrm{e}^{\beta x}K_{\nu}(x)\,\mathrm{d}x &= \Big[f'(x)x^{\nu+1}\mathrm{e}^{\beta x}K_{\nu}(x)\Big]_0^{\infty} - \int_0^{\infty} f'(x)\frac{\mathrm{d}}{\mathrm{d}x}\Big(x^{\nu+1}\mathrm{e}^{\beta x}K_{\nu}(x)\Big)\,\mathrm{d}x \\ &= -\int_0^{\infty} f'(x)\frac{\mathrm{d}}{\mathrm{d}x}\Big(x^{\nu+1}\mathrm{e}^{\beta x}K_{\nu}(x)\Big)\,\mathrm{d}x.\end{aligned}$$

Here we made use of the fact that $f'(x)$ is bounded, and we used the asymptotic properties of $K_{\nu}(x)$, given by (B.8) and (B.11), which ensure that $\lim_{x\to 0^+} x^{\nu+1}\mathrm{e}^{\beta x}K_{\nu}(x) = \lim_{x\to\infty} x^{\nu+1}\mathrm{e}^{\beta x}K_{\nu}(x) = 0$, for $-1 < \beta < 1$. We now apply integration by parts once more to obtain

$$\int_0^\infty f''(x)x^{\nu+1}\mathrm{e}^{\beta x}K_\nu(x)\,\mathrm{d}x = -\bigg[f(x)\frac{\mathrm{d}}{\mathrm{d}x}\Big(x^{\nu+1}\mathrm{e}^{\beta x}K_\nu(x)\Big)\bigg]_0^\infty + \int_0^\infty f(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\Big(x^{\nu+1}\mathrm{e}^{\beta x}K_\nu(x)\Big)\,\mathrm{d}x.$$

We also use integration by parts to obtain

$$\int_0^\infty (2\nu+1+2\beta x)f'(x)x^\nu\mathrm{e}^{\beta x}K_\nu(x)\,\mathrm{d}x = \Big[f(x)((2\nu+1)x^\nu+2\beta x^{\nu+1})\mathrm{e}^{\beta x}K_\nu(x)\Big]_0^\infty - \int_0^\infty f(x)\frac{\mathrm{d}}{\mathrm{d}x}\Big(((2\nu+1)x^\nu+2\beta x^{\nu+1})\mathrm{e}^{\beta x}K_\nu(x)\Big)\,\mathrm{d}x.$$

We therefore have

$$\begin{aligned} I_2 &= \bigg[f(x)\bigg\{((2\nu+1)x^\nu+2\beta x^{\nu+1})\mathrm{e}^{\beta x}K_\nu(x) - \frac{\mathrm{d}}{\mathrm{d}x}\Big(x^{\nu+1}\mathrm{e}^{\beta x}K_\nu(x)\Big)\bigg\}\bigg]_0^\infty \\ &\quad+\int_0^\infty f(x)\bigg\{\frac{\mathrm{d}^2}{\mathrm{d}x^2}\Big(x^{\nu+1}\mathrm{e}^{\beta x}K_\nu(x)\Big) - \frac{\mathrm{d}}{\mathrm{d}x}\Big(((2\nu+1)x^\nu+2\beta x^{\nu+1})\mathrm{e}^{\beta x}K_\nu(x)\Big) \\ &\quad+((2\nu+1)\beta-(1-\beta^2)x)x^\nu\mathrm{e}^{\beta x}K_\nu(x)\bigg\}\,\mathrm{d}x \\ &= \lim_{x\to0^+} f(x)\mathrm{e}^{\beta x}(x^{\nu+1}K_\nu'(x)-\nu x^\nu K_\nu(x)) + \int_0^\infty f(x)x^{\nu-1}\mathrm{e}^{\beta x}\{x^2K_\nu''(x) \\ &\quad+(2(\nu+1)x+\beta x^2)K_\nu'(x)+(\nu(\nu+1)+2(\nu+1)\beta x+\beta^2x^2)K_\nu(x) \\ &\quad-(2\nu+1)(xK_\nu'(x)+(\nu+\beta)K_\nu(x)-2\beta(xK_\nu'(x)+(\nu+1+\beta x)K_\nu(x) \\ &\quad+((2\nu+1)\beta x-(1-\beta^2)x^2)K_\nu(x)\}\,\mathrm{d}x, \end{aligned}$$

where we used that f is bounded and (B.8) and (B.11). We may simplify the integrand in the above expression to obtain

$$\begin{aligned} I_2 &= \lim_{x\to0^+} f(x)\mathrm{e}^{\beta x}(x^{\nu+1}K_\nu'(x)-\nu x^\nu K_\nu(x)) \\ &\quad+\int_0^\infty f(x)x^{\nu-1}\mathrm{e}^{\beta x}\left\{x^2K_\nu''(x)+xK_\nu'(x)-(x^2+\nu^2)K_\nu(x)\right\}\,\mathrm{d}x \\ &= -\lim_{x\to0^+} f(x)\mathrm{e}^{\beta x}\bigg(\frac{1}{2}x^{\nu+1}(K_{\nu+1}(x)+K_{\nu-1}(x))+\nu x^\nu K_\nu(x)\bigg), \end{aligned}$$

where the final equality was due to an application of (B.26) and the fact that $K_\nu(x)$ satisfies the modified Bessel differential equation (B.46). We now calculate the limit in the above expression. We first consider the case $\nu > 0$, using (B.8),

$$I_2 = -\lim_{x\to0^+}(2^{\nu-1}\Gamma(\nu+1)+2^{\nu-1}\nu\Gamma(\nu))f(x) = -\lim_{x\to0^+}2^\nu\Gamma(\nu+1)f(x),$$

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

$$I_2 = -\lim_{x\to 0^+} f(x)\mathrm{e}^{\beta x} x K_1(x) = -\lim_{x\to 0^+} \Gamma(1) f(x) = -\lim_{x\to 0^+} 2^0\Gamma(1+1) f(x),$$

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

$$I_2 = -\lim_{x\to 0^+} 2^{\nu}\Gamma(\nu+1) f(x), \quad \text{for all } \nu \geq 0.$$

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

$$\begin{aligned} I_2 &= -\lim_{x\to 0^+} f(x)\mathrm{e}^{\beta x}(x^{\nu+1}(K_{\nu+1}(x) + K_{1-\nu}(x))/2 + \nu x^{\nu} K_{-\nu}(x)) \\ &= -\lim_{x\to 0^+} (2^{\nu-1}\Gamma(\nu+1) + 2^{\nu-1}(\Gamma(1-\nu) - (-\nu)\Gamma(-\nu))x^{2\nu}) f(x) \\ &= -\lim_{x\to 0^+} 2^{\nu-1}\Gamma(\nu+1) f(x), \quad \text{for } -1/2 < \nu < 0. \end{aligned}$$

We now turn our attention to I_1. Changing variables, $x = -t$, gives

$$I_1 = -\int_0^{\infty} \{tf''(-t) + (2\nu+1+2(-\beta)t)f'(-t) + ((2\nu+1)(-\beta) - (1-\beta^2)t)f(-t)\}\mathrm{e}^{-\beta t} t^{\nu} K_{\nu}(t)\,\mathrm{d}t.$$

We observe that this has a similar form to I_2 and we therefore have that

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

Since f is continuous we have

$$I_1 + I_2 = \lim_{x\to 0^+} 2^{\nu}\Gamma(\nu+1)(f(-x) - f(x)) = 0, \quad \text{for } \nu \geq 0.$$

The argument for $-1/2 < \nu < 0$ is similar.

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

$$xf''(x) + (2\nu+1+2\beta x)f'(x) + ((2\nu+1)\beta - (1-\beta^2)x)f(x) = \chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z), \quad (3.9)$$

where $K_{\nu,\beta}(z)$ is the cumulative distribution function of the $\mathrm{VG}_2(\nu,1,\beta,0)$ distribution.

Our goal is to show that if the random variable X satisfies (3.8), with $\alpha = 1$ and $\mu = 0$, for all functions f with the properties as stated in the lemma, then $X \sim \mathrm{VG}_2(\nu,1,\beta,0)$. Hence, we need to show that for all $z \in \mathbb{R}$

$$\begin{aligned} 0 &= \mathbb{P}(X \leq z) - K_{\nu,\beta}(x) \\ &= \mathbb{E}(\mathbf{1}(X \leq z)) - K_{\nu,\beta}(x) \\ &= \mathbb{E}\{Xf_z''(X) + (2\nu + 1 + 2\beta X)f_z'(X) + ((2\nu+1)\beta - (1-\beta^2)X)f_z(X)\}. \end{aligned}$$

We therefore need to show that the expectations in the above the expression exist and we now do so.

Using Lemma 3.14 (below) with $h(x) = \chi_{(-\infty,z]}(x)$ we see that a bounded solution to (3.9) is given by

$$\begin{aligned} f_z(x) &= -\frac{\mathrm{e}^{-\beta x}K_\nu(|x|)}{|x|^\nu}\int_0^x \mathrm{e}^{\beta y}|y|^\nu I_\nu(|y|)[\chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z)]\,\mathrm{d}y \\ &\quad - \frac{\mathrm{e}^{-\beta x}I_\nu(|x|)}{|x|^\nu}\int_x^\infty \mathrm{e}^{\beta y}|y|^\nu K_\nu(|y|)[\chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z)]\,\mathrm{d}y \\ &= -\frac{\mathrm{e}^{-\beta x}K_\nu(|x|)}{|x|^\nu}\int_0^x \mathrm{e}^{\beta y}|y|^\nu I_\nu(|y|)[\chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z)]\,\mathrm{d}y \\ &\quad + \frac{\mathrm{e}^{-\beta x}I_\nu(|x|)}{|x|^\nu}\int_{-\infty}^x \mathrm{e}^{\beta y}|y|^\nu K_\nu(|y|)[\chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z)]\,\mathrm{d}y, \end{aligned}$$

and by Lemma 3.13 (below) this is the unique bounded solution.

Let $p(x)$ denote the density function of a $\mathrm{VG}_2(\nu, 1, \beta, 0)$ random variable. Integrating (3.9) over $\mathbb{R}$ gives

$$\begin{aligned} \int_{-\infty}^\infty |xf_z''(x)|p(x)\,\mathrm{d}x &\leq (2\nu+1)\int_{-\infty}^\infty |f_z'(x)|p(x)\,\mathrm{d}x + 2\beta\int_{-\infty}^\infty |xf_z'(x)|p(x)\,\mathrm{d}x \\ &\quad + (2\nu+1)\beta\int_{-\infty}^\infty |f_z(x)|p(x)\,\mathrm{d}x + (1-\beta^2)\int_{-\infty}^\infty |xf_z(x)|p(x)\,\mathrm{d}x \\ &\quad + \int_{-\infty}^\infty |\chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z)|p(x)\,\mathrm{d}x. \end{aligned}$$

The last integral is finite and therefore the above expression shows it is sufficient to show that $\mathbb{E}|f_z'(X)|$, $\mathbb{E}|Xf_z'(X)|$, $\mathbb{E}|f_z(X)|$ and $\mathbb{E}|Xf_z(X)|$ are finite. By Lemma 3.14, f and f' are bounded for all $x \in \mathbb{R}$, so there exist constants M_1 and M_2 such that $\|f_z\| \leq M_1$, $\|f_z'\| \leq M_2$. Therefore

$$\mathbb{E}|Xf_z(X)| \leq \|f_z\|\mathbb{E}|X| \leq M_1\{\mathbb{E}X^2\}^{1/2} < \infty.$$

We may bound the other expectations similarly. Also, f_z is a piecewise twice continuously differentiable function, and f_z and f_z' are bounded. Suppose that (3.8), in the case $\mu = 0$, $\alpha = 1$, holds for all piecewise twice continuously differentiable functions. Then it holds for f_z. By (3.9) we have

$$\begin{aligned}0 &= \mathbb{E}\{Wf_z''(W) + (2\nu+1+2\beta W)f_z'(W) + ((2\nu+1)\beta - (1-\beta^2)W)f_z(W)\}\\ &= \mathbb{E}\{\chi_{(-\infty,z]}(x) - K_{\nu,\beta}(z)\}\\ &= \mathbb{P}(W \leq z) - K_{\nu,\beta}(z).\end{aligned}$$

Therefore $\mathscr{L}(W) = \mathrm{VG}_2(\nu, 1, \beta, 0)$.

We now extend the result to the general case. Let $W = \alpha(Z-\mu)$, then

$$W \sim \mathrm{VG}_2(\nu, 1, \beta, 0) \qquad \text{if and only if} \qquad Z \sim \mathrm{VG}_2(\nu, \alpha, \alpha\beta, \mu).$$

Also

$$\mathbb{E}\{Wf''(W) + (2\nu+1+2\beta W)f'(W) + ((2\nu+1)\beta - (1-\beta^2)W)f(W)\} = 0,$$

if and only if

$$\begin{aligned}&\mathbb{E}\{\alpha(Z-\mu)f''(\alpha(Z-\mu)) + (2\nu+1+2\beta\alpha(Z-\mu))f'(\alpha(Z-\mu))\\ &+((2\nu+1)\beta - (1-\beta^2)\alpha(Z-\mu))f(\alpha(Z-\mu))\} = 0.\end{aligned} \tag{3.10}$$

Let $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(z) = f(\alpha(z-\mu)). \tag{3.11}$$

We have $g'(z) = \alpha f'(\alpha(z-\mu))$ and $g''(z) = \alpha^2 f''(\alpha(z-\mu))$. Then g will satisfy the conditions stated in the lemma provided f does. Substituting these expressions for g into (3.10) gives

$$\mathbb{E}\{(Z-\mu)g''(Z) + (2\nu+1+2\alpha\beta(Z-\mu))g'(Z) + ((2\nu+1)\beta - (\alpha^2-(\alpha\beta)^2)(Z-\mu))g(Z)\} = 0.$$

Putting this together we see that $Z \sim \mathrm{VG}_2(\nu, \alpha, \alpha\beta, \mu)$ if and only if for all functions g, which satisfy the conditions in the lemma, the following holds

$$\mathbb{E}\{(Z-\mu)g''(Z) + (2\nu+1+2\alpha\beta(Z-\mu))g'(Z) + ((2\nu+1)\beta - (\alpha^2-(\alpha\beta)^2)(Z-\mu))g(Z)\} = 0.$$

Therefore the result has been proved in the general case. $\square$

Lemma 3.10 suggests the following Stein equation for the $\mathrm{VG}_2(\nu, \alpha, \beta, \mu)$ distribution:

$$(x-\mu)f''(x)+(2\nu+1+2\beta(x-\mu))f'(x)+((2\nu+1)\beta-(\alpha^2-\beta^2)(x-\mu))f(x) = h(x)-\tilde{\mathrm{VG}}^{\nu,\alpha}_{\beta,\mu}h, \tag{3.12}$$

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

In order to simplify the calculations of Section 3.3, we will make use of the Stein equation for the $\mathrm{VG}_2(\nu,1,\beta,0)$ distribution, where $-1<\beta<1$. Results for the full parametrisation can then be recovered by applying the linear transformation (3.11). For the $\mathrm{VG}_2(\nu,1,\beta,0)$ distribution, the Stein equation (3.12) reduces to

$$xf''(x)+(2\nu+1+2\beta x)f'(x)+((2\nu+1)\beta-(1-\beta^2)x)f(x)=h(x)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h. \qquad (3.13)$$

Changing parametrisation in (3.12) via (3.3) and multiplying through by σ^2 gives as Stein equation for the $\mathrm{VG}(r,\theta,\sigma,\mu)$ distribution

$$\sigma^2(x-\mu)f''(x)+(\sigma^2r+2\theta(x-\mu))f'(x)+(r\theta-(x-\mu))f(x)=h(x)-\mathrm{VG}^{r,\theta}_{\sigma,\mu}h, \qquad (3.14)$$

where $\mathrm{VG}^{r,\theta}_{\sigma,\mu}h$ denotes the quantity $\mathbb{E}h(X)$ for $X\sim\mathrm{VG}(r,\theta,\sigma,\mu)$.

Remark 3.11. Whilst the Gamma distribution is not covered by Lemma 3.10, we note that letting $r=2s$, $\theta=(2\lambda)^{-1}$, $\mu=0$ and taking the limit $\sigma\to 0$ in (3.14) gives the Stein equation

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

which, recalling (2.14), we recognise as the $\Gamma(s,\lambda)$ Stein equation (2.14) of Luk [44] (up to a multiple of λ^{-1}). As a Stein equation for a given distribution is not unique, (see Barbour [5]) the fact that in the appropriate limit the Variance-Gamma Stein equation (3.14) reduces to the Gamma Stein equation (2.14) is an attractive feature, and one we would hope for since the Gamma distribution is a limiting case of the Variance-Gamma distribution.

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

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

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

A remark is in order regarding the $\mathrm{VG}(r,\theta,\sigma,\mu)$ Stein equation (3.14).

Remark 3.12. We could have obtained a first order Stein operator for the $\mathrm{VG}(r,\theta,\sigma,\mu)$ distributions using the density approach of Stein et al. [74]. However, this approach would lead to an operator involving the modified function $K_\nu(x)$. Using such a Stein equation to prove approximation results with standard coupling techniques would be difficult. In contrast, our $\mathrm{VG}(r,\theta,\sigma,\mu)$ Stein equation is much more amenable to the use of couplings, as we shall see in

Chapter 4.

We now turn our attention to solving the Stein equation (3.13) and obtaining smoothness estimates for it. We begin by obtaining a bounded solution to the Stein equation and showing that this solution is the unique bounded solution. This is achieved through the following lemmas, the proofs of which are simple and are given in Appendix A.

Lemma 3.13. *Suppose $-1<\beta<1$ and $\nu\geq 0$, then there is at most one bounded solution to the Variance-Gamma Stein equation (3.13). Moreover, if $-1<\beta<1$, $\nu>-1/2$ and k is a positive integer then there is at most one solution which has a bounded k-th derivative.*

Lemma 3.14. *Let $h:\mathbb{R}\rightarrow\mathbb{R}$ be a measurable function with $\mathbb{E}|h(X)|<\infty$, where $X\sim\mathrm{VG}_2(\nu,1,\beta,0)$, and $\nu>-1/2$ and $-1<\beta<1$. Then a solution $f:\mathbb{R}\rightarrow\mathbb{R}$ to the Variance-Gamma Stein equation (3.13) is given by*

$$\begin{aligned} f(x) &= -\frac{\mathrm{e}^{-\beta x}K_{\nu}(|x|)}{|x|^{\nu}}\int_0^x \mathrm{e}^{\beta y}|y|^{\nu}I_{\nu}(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \\ &\quad -\frac{\mathrm{e}^{-\beta x}I_{\nu}(|x|)}{|x|^{\nu}}\int_x^{\infty} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \qquad (3.15)\\ &= -\frac{\mathrm{e}^{-\beta x}K_{\nu}(|x|)}{|x|^{\nu}}\int_0^x \mathrm{e}^{\beta y}|y|^{\nu}I_{\nu}(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \\ &\quad +\frac{\mathrm{e}^{-\beta x}I_{\nu}(|x|)}{|x|^{\nu}}\int_{-\infty}^{x} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y. \end{aligned}$$

Suppose further that h is bounded, then f and f' and are bounded for all $x\in\mathbb{R}$. Moreover, by Lemma 3.13 this is the unique bounded solution for $\nu\geq 0$ and $-1<\beta<1$, and the unique solution with bounded k-th derivative for $\nu>-1/2$ and $-1<\beta<1$, where $k\geq 1$.

Remark 3.15. The equality

$$\int_{-\infty}^{x} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y = -\int_x^{\infty} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y$$

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

3.3 Smoothness Estimates of the Solution

3.3.1 Smoothness estimates by a direct approach

Here we obtain bounds for the solution of the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14) and its first four derivatives. To simplify the calculations, we obtain bounds using the second parametrisation of the Variance-Gamma distributions and then make a change of variables to obtain the desired bounds in terms of the first parametrisation.

In the following lemma we present formulas for the first four derivatives of the solution of the $\mathrm{VG}_2(\nu,1,\beta,0)$ Stein equation (3.13). We restrict our attention to the region $x>0$, as discussed in Remark 3.15.

Lemma 3.16. *Suppose $h\in C^2(\mathbb{R})$ and let $\tilde{h}(x)=h(x)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h$. Then the first four derivatives of the Variance Gamma Stein equation (3.13), in the region $x>0$, are given by*

$$\begin{aligned}
f'(x) &= -\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right]\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
&\quad -\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y; \\
f''(x) &= \frac{\tilde{h}(x)}{x}-\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right]\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
&\quad -\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y; \qquad (3.16) \\
f^{(3)}(x) &= \frac{h'(x)}{x}-\left(\frac{2\nu+2}{x^2}+\frac{2\beta}{x}\right)\tilde{h}(x)-\left[\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right]\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
&\quad -\left[\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y; \qquad (3.17) \\
f^{(4)}(x) &= \frac{h''(x)}{x}-\left(\frac{2\nu+3}{x^2}+\frac{2\beta}{x}\right)h'(x)+\left(\frac{(2\nu+2)(2\nu+3)}{x^3}+\frac{\beta(6\nu+5)}{x^2}+\frac{1}{x}+\frac{3\beta^2}{x}\right)\tilde{h}(x) \\
&\quad -\left[\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right]\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
&\quad -\left[\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y. \qquad (3.18)
\end{aligned}$$

Proof. We will make repeated us of the Leibniz's theorem for differentiation of an integral, which states that provided the functions $u(y,x)$ and $\frac{\partial u}{\partial x}(y,x)$ are continuous in both x and y in the region $a(x)\le y\le b(x)$, $x_0\le x\le x_1$, and the functions $a(x)$ and $b(x)$ are continuous and have continuous derivatives for $x_0 \leq x \leq x_1$, then for $x_0 \leq x \leq x_1$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{a(x)}^{b(x)} u(y,x)\,\mathrm{d}y = \int_{a(x)}^{b(x)} \frac{\partial}{\partial x}u(y,x)\,\mathrm{d}y + u(b,x)\frac{\mathrm{d}b}{\mathrm{d}x} - u(a,x)\frac{\mathrm{d}a}{\mathrm{d}x}. \tag{3.19}$$

We will also make use of the Wronskian formula for modified Bessel functions (see (B.22)):

$$I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x) = \frac{1}{x},$$

as well as the formulas (B.31) – (B.36) for the first three derivatives of the functions $x^{-\nu}I_\nu(x)$ and $x^{-\nu}K_\nu(x)$.

It easy to compute the first and second derivatives by applying (3.19) and (B.22). The calculation of the third derivative is still straightforward but a little longer. We differentiate the formula for the second derivative using (3.19) to obtain

$$\begin{aligned} f^{(3)}(x) = \frac{h'(x)}{x} - \frac{\tilde{h}(x)}{x^2} &- \left[\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right]\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\ &- \left[\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\ &+ \tilde{h}(x)\bigg\{-\mathrm{e}^{\beta x}x^\nu I_\nu(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right) + \mathrm{e}^{\beta x}x^\nu K_\nu(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\bigg\}. \end{aligned}$$

Using the formulas (B.31) – (B.34) for the first two derivatives of $x^{-\nu}I_\nu(x)$ and $x^{-\nu}K_\nu(x)$, and the Wronskian formula (B.22) allows us to calculate the term in the brackets $(*)$ from the above expression as

$$\begin{aligned} (*) &= -x^\nu I_\nu(x)\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{K_\nu(x)}{x^\nu}\right) - 2\beta\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{K_\nu(x)}{x^\nu}\right) + \beta^2 K_\nu(x)\right) \\ &\quad + x^\nu K_\nu(x)\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{I_\nu(x)}{x^\nu}\right) - 2\beta\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{I_\nu(x)}{x^\nu}\right) + \beta^2 I_\nu(x)\right) \\ &= -\left(\frac{2\nu+1}{x} + 2\beta\right)[I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x)] \\ &= -\left(\frac{2\nu+1}{x^2} + \frac{2\beta}{x}\right). \end{aligned}$$

Substituting $(*)$ into the expression for $f^{(3)}(x)$ gives the result.

Finally, we verify the formula for the fourth derivative. We differentiate the formula for the third derivative using (3.19) to obtain

$$\begin{aligned}
f^{(4)}(x) &= \frac{h''(x)}{x} - \left(\frac{2\nu+3}{x^2} + \frac{2\beta}{x}\right)h'(x) + \left(\frac{2(2\nu+2)}{x^3} + \frac{2\beta}{x^2}\right)\tilde{h}(x) \\
&\quad - \left[\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right]\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
&\quad - \left[\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
&\quad + \tilde{h}(x)\bigg\{ - \mathrm{e}^{\beta x}x^\nu I_\nu(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right) + \mathrm{e}^{\beta x}x^\nu K_\nu(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\bigg\}.
\end{aligned}$$

Using the formulas (B.31) – (B.36) for the first three derivatives of $x^{-\nu}I_\nu(x)$ and $x^{-\nu}K_\nu(x)$, and the Wronskian formula (B.22) allows us to calculate the term in the brackets $(**)$ from the above expression as

$$\begin{aligned}
(**) &= -x^\nu I_\nu(x)\left(\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right) - 3\beta\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{K_\nu(x)}{x^\nu}\right) + 3\beta^2\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{K_\nu(x)}{x^\nu}\right) - \beta^3 K_\nu(x)\right) \\
&\quad + x^\nu K_\nu(x)\left(\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{I_\nu(x)}{x^\nu}\right) - 3\beta\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{I_\nu(x)}{x^\nu}\right) + 3\beta^2\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{I_\nu(x)}{x^\nu}\right) - \beta^3 I_\nu(x)\right) \\
&= \left(\frac{(2\nu+1)(2\nu+2)}{x^3} + \frac{\beta(6\nu+5)}{x^2} + \frac{1}{x} + \frac{3\beta^2}{x}\right)[I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x)] \\
&= \frac{(2\nu+1)(2\nu+2)}{x^3} + \frac{\beta(6\nu+5)}{x^2} + \frac{1}{x} + \frac{3\beta^2}{x}.
\end{aligned}$$

Substituting $(**)$ into the expression for $f^{(4)}(x)$ gives the result. □

In their current forms the derivatives of the solution are not suitable for bounding, as they contain terms that are singular. In the next lemma we use integration by parts to group the singularities together and then apply the triangle inequality.

The following notation for the the repeated integral of the function $\mathrm{e}^{\beta x}x^\nu I_\nu(x)$ will be used in the next lemma,

$$I_{(\nu,\beta,0)}(x) = \mathrm{e}^{\beta x}x^\nu I_\nu(x), \qquad I_{(\nu,\beta,n+1)}(x) = \int_0^x I_{(\nu,\beta,n)}(y)\,\mathrm{d}y, \quad n = 0,1,2,3,\ldots. \tag{3.20}$$

Lemma 3.17. *Suppose $h \in C_b^3(\mathbb{R})$ and let $\tilde{h}(x) = h(x) - \tilde{\mathrm{V}}\mathrm{G}^{\nu,1}_{\beta,0}h$. Then the solution and its first four derivatives of the Variance Gamma Stein equation (3.13), in the region $x > 0$, may be* *be bounded as follows*

$$\begin{aligned}
|f(x)| &\leq \|\tilde{h}\|\bigg|I_{(\nu,\beta,1)}(x)\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg| + \|\tilde{h}\|\bigg|\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\,\mathrm{d}y\bigg|; \\
|f'(x)| &\leq \|\tilde{h}\|\bigg|I_{(\nu,\beta,1)}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| + \|\tilde{h}\|\bigg|\bigg[\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\bigg)\bigg]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\,\mathrm{d}y\bigg|; \\
|f''(x)| &\leq \|\tilde{h}\|\bigg|\frac{1}{x} - I_{(\nu,\beta,1)}(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| + \|h'\|\bigg|I_{(\nu,\beta,2)}(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| \\
&\quad + \|\tilde{h}\|\bigg|\bigg[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\bigg)\bigg]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\,\mathrm{d}y\bigg|; \\
|f^{(3)}(x)| &\leq \|\tilde{h}\|\bigg|\frac{2\nu+2}{x^2} + \frac{2\beta}{x} + I_{(\nu,\beta,1)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| \\
&\quad + \|h'\|\bigg|\frac{1}{x} + I_{(\nu,\beta,2)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| + \|h''\|\bigg|I_{(\nu,\beta,3)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| \\
&\quad + \|\tilde{h}\|\bigg|\bigg[\frac{\mathrm{d}^3}{\mathrm{d}x^3}\bigg(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\bigg)\bigg]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\,\mathrm{d}y\bigg|; \\
|f^{(4)}(x)| &\leq \|\tilde{h}\|\bigg|\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{\beta(6\nu+5)}{x^2} + \frac{1}{x} + \frac{3\beta^2}{x} - I_{(\nu,\beta,1)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| \\
&\quad + \|h'\|\bigg|\frac{2\nu+3}{x^2} + \frac{2\beta}{x} - I_{(\nu,\beta,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| \\
&\quad + \|h''\|\bigg|\frac{1}{x} - I_{(\nu,\beta,3)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| + \|h^{(3)}\|\bigg|I_{(\nu,\beta,4)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg| \\
&\quad + \|\tilde{h}\|\bigg|\bigg[\frac{\mathrm{d}^4}{\mathrm{d}x^4}\bigg(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\bigg)\bigg]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\,\mathrm{d}y\bigg|.
\end{aligned}$$

Proof. The first two bounds are immediate from the the formulas for $f(x)$ and $f'(x)$. Integrating by parts and using the notation for the repeated integral of $\mathrm{e}^{\beta x}x^\nu I_\nu(x)$, which is defined by (3.20), gives

$$\begin{aligned}
f''(x) = \frac{\tilde{h}(x)}{x} &- \bigg[\tilde{h}(x)\int_0^x \mathrm{e}^{\beta y}y^\nu I_\nu(y)\,\mathrm{d}y - \int_0^x \tilde{h}(y)\bigg(\int_0^y \mathrm{e}^{\beta u}u^\nu I_\nu(u)du\bigg)\,\mathrm{d}y\bigg]\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg) \\
&- \bigg[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\bigg)\bigg]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
= \tilde{h}(x)&\bigg[\frac{1}{x} - I_{(\nu,\beta,1)}(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg] - \bigg[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\bigg)\bigg]\int_0^x h'(y)I_{(\nu,\beta,1)}(y)\,\mathrm{d}y \\
&- \bigg[\frac{\mathrm{d}^2}{\mathrm{d}x^2}\bigg(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\bigg)\bigg]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y.
\end{aligned}$$

The bound now follows from the triangle inequality and (3.20). The bounds for the third and fourth derivatives are obtained in a similar manner, in which use integration by parts on the integrals $I_{(\nu,\beta,n)}(x)$. □

We can see that there are four different types of terms that need bounding, which we list below:

(i). $\left\|\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}I_{(\nu,\beta,1)}(x)\right\|$;

(ii). $\left\|I_{(\nu,\beta,n)}(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right\|$;

(iii). $\left\|\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y\right\|$;

(iv). $\left\|I_{(\nu,\beta,k)}(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)+\text{singular terms}\right\|$,

where k and n are non negative integers and $k<n$. We require bounds for the expressions of type (ii) for $n=1,2,3,4$, bounds for the terms of type (iii) for $n=0,1,2,3,4$, and bounds for the terms of type (iv) for $1\leq k<n\leq 4$.

In Appendix D we show that the term (i) is bounded in the region $x\geq 0$ for all $\nu>-1/2$ and $-1<\beta<1$, and the terms of type (ii) and (iii) are bounded for $n=0,1$ and for all $\nu>-1/2$ and $-1<\beta<1$, in the region $x\geq 0$. This allows us to show that the solution of the $\mathrm{VG}_2(\nu,1,\beta,0)$ Stein equation and its first derivative are bounded – see Lemma 3.14.

However, we are unable to obtain uniform bounds for the terms of type (iv) for the general case $-1<\beta<1$. When we specialise to the case $\beta=0$ we are able to obtain uniform bounds for the terms of type (ii) and (iii) for all $n\geq 0$ and $\nu>-1/2$ in the region $x\geq 0$. We can obtain uniform bounds for the terms of type (iv), but only in the region $0\leq x\leq 1$. Although, for the case $\nu=\beta=0$ we obtain uniform bounds for all $x\geq 0$ – see Lemma D.20.

In the following lemma we use these bounds together with Lemma 3.17 to bound the solution of the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation and its first four derivatives in the region $0\leq x\leq 1$. In Lemma 3.19 we shall obtain bounds in the region $x\geq 1$ and combining these bounds gives a bound for all $x\geq 0$. Recalling Remark 3.15, we are then able to deduce bounds for the derivatives for all $x\in\mathbb{R}$. Changing parameters will then lead to uniform bounds for the solution of the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14).

It is an open problem to obtain uniform bounds for the terms of type (ii), (iii) and (iv) in the interval $0\leq x<\infty$ for all $n\geq 0$, $\nu>-1/2$ and $-1<\beta<1$. If such bounds were established, we would be able to achieve the goal of obtaining uniform bounds for the first four derivatives of the solution of the $\mathrm{VG}(r,\theta,\sigma,\mu)$ Stein equation, which would follow from a change of parameters via (3.3).

Our bounds for the solution of the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation (3.13) in the region $0\leq x\leq 1$ are given in the following lemma.

Lemma 3.18. *Suppose that $h \in C_b^3(\mathbb{R})$ and $\nu > -1/2$. Then, for all $x \geq 0$ the following bounds on the solution, and its first derivative, of the $\mathrm{VG}_2(\nu, 1, 0, 0)$ Stein equation (3.13) hold*

$$\begin{aligned}
\|f\| &\leq \left(\frac{1}{2\nu+1} + \frac{\pi\Gamma(\nu+1/2)}{2\Gamma(\nu+1)}\right)\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|, \\
\|f'\| &\leq \left(\frac{1}{2\nu+1} + \frac{1}{2\nu+2}\right)\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|.
\end{aligned}$$

Suppose now that $x \leq 1$, then the following bounds hold for the second, third and fourth derivative of the solution to the Symmetric Variance-Gamma Stein equation:

$$\begin{aligned}
\sup_{0\leq x\leq 1}|f''(x)| &\leq \frac{3\|h'\|}{2\nu+1} + \frac{(8+\sqrt{\pi})\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|}{2(\nu+1)}, \\
\sup_{0\leq x\leq 1}|f^{(3)}(x)| &\leq \frac{4\|h''\|}{2\nu+1} + \frac{5\|h'\|}{2\nu+1} + \left(\frac{1}{2(\nu+1)} + \frac{25}{12(2\nu+1)} + \frac{1}{v_1(\nu)}\right)\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|, \\
\sup_{0\leq x\leq 1}|f^{(4)}(x)| &\leq \frac{8\|h^{(3)}\|}{2\nu+1} + \frac{9\|h''\|}{2\nu+1} + \left(\frac{77}{20(2\nu+1)} + \frac{1}{v_2(\nu)}\right)\|h'\| + \left(\frac{1}{2(\nu+2)}\right. \\
&\quad \left. + \left(\frac{2779}{768} + \frac{\sqrt{\pi}}{2}\right)\frac{1}{2\nu+1} + \frac{1}{v_3(\nu)}\right)\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|,
\end{aligned}$$

where

$$\begin{aligned}
v_1(\nu) &= \begin{cases} 2^{2\nu+1}\nu!\,(\nu+2)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}. \end{cases} \\
v_2(\nu) &= \begin{cases} 2^{2\nu+2}\nu!\,(\nu+3)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu+1}\Gamma(\nu+1)\Gamma(\nu+5)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}. \end{cases} \\
v_3(\nu) &= \begin{cases} 2^{2\nu+3}(\nu+1)!\,(\nu+3)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu+2}\Gamma(\nu+2)\Gamma(\nu+5)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}. \end{cases}
\end{aligned}$$

When $\nu = 0, 1/2, 1, 3/2, \ldots$, the bounds for $f^{(3)}(x)$ and $f^{(4)}(x)$ simplify as follows:

$$\begin{aligned}
\sup_{0\leq x\leq 1}|f^{(3)}(x)| &\leq \frac{4\|h''\|}{2\nu+1} + \frac{5\|h'\|}{2\nu+1} + \left(\frac{4}{2\nu+1} + \frac{1}{2(\nu+1)}\right)\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|; \\
\sup_{0\leq x\leq 1}|f^{(4)}(x)| &\leq \frac{8\|h^{(3)}\|}{2\nu+1} + \frac{9\|h''\|}{2\nu+1} + \frac{4\|h'\|}{2\nu+1} + \left(\frac{8+\sqrt{\pi}}{2(2\nu+1)} + \frac{1}{2(\nu+2)}\right)\|h - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|.
\end{aligned}$$

Proof. We require uniform bounds for the supremum norms involving modified Bessel functions that appear in the bounds of Lemma 3.17, for the case $\beta = 0$. These expressions are bounded in Appendix D. To obtain our bound for f we apply inequalities (D.5) and (D.6); to bound f'

we use (D.4) and (D.7); to bound f'' we apply (D.4), (D.8) and (D.11). To obtain our first bound for $f^{(3)}$ we apply (D.4), (D.7), (D.11) and (D.17), and to obtain the second bound we use (D.18) instead of (D.17). To obtain our first bound for $f^{(4)}$ we apply (D.4), (D.8), (D.11), (D.19) and (D.21), and to obtain the second bound we use (D.20) and (D.22) instead of (D.19) and (D.21). □

By again using integration by parts we can obtain different formulas for the derivatives of the second, third and fourth derivatives of the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation. We are then able to bound these derivatives in the region $x\geq 1$. This gives:

Lemma 3.19. *Suppose that $h\in C_b^3(\mathbb{R})$ and $\nu>-1/2$. Then for $x\geq 1$ the following bound holds for the second, third and fourth derivative of the solution to the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation (3.13):*

$$\begin{aligned}\sup_{x\geq 1}|f''(x)| &\leq 3\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{1}{2\nu+2}\bigg)\|h'\|+4\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{1}{2\nu+1}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|;\\ \sup_{x\geq 1}|f^{(3)}(x)| &\leq 5\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\bigg)\|h''\|+18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\bigg)\|h'\|\\ &\quad+18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|;\\ \sup_{x\geq 1}|f^{(4)}(x)| &\leq 8\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+13}}+\frac{1}{2\nu+4}\bigg)\|h^{(3)}\|+52\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\bigg)\|h''\|\\ &\quad+123\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\bigg)\|h'\|+123\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|.\end{aligned}$$

Proof. We obtain the bound for $\|f''\|$. We obtain the bounds for $\|f^{(3)}\|$ and $\|f^{(4)}\|$ by similar calculations, which are given in Appendix A. Let $\tilde{h}(x)=h(x)-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h$. Substituting the second derivative formulas (B.37) and (B.38) for $x^{-\nu}I_\nu(x)$ and $x^{-\nu}K_\nu(x)$, respectively, into the formula (3.16) for $f''(x)$ gives

$$\begin{aligned}f''(x)=\frac{\tilde{h}(x)}{x}&-\bigg(\frac{K_{\nu+2}(x)}{x^\nu}-\frac{K_{\nu+1}(x)}{x^{\nu+1}}\bigg)\int_0^x y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y\\ &-\bigg(\frac{I_{\nu+2}(x)}{x^\nu}+\frac{I_{\nu+1}(x)}{x^{\nu+1}}\bigg)\int_x^\infty y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y,\end{aligned}$$

where we took $\beta=0$. Integrating by parts, using that $\frac{\mathrm{d}}{\mathrm{d}x}(x^{\nu+1}I_{\nu+1}(x))=x^{\nu+1}I_\nu(x)$ and

$\frac{\mathrm{d}}{\mathrm{d}x}(x^{\nu+1}K_{\nu+1}(x)) = -x^{\nu+1}K_\nu(x)$ (see (B.29) and (B.30)), gives

$$\begin{aligned} f''(x) = \frac{\tilde{h}(x)}{x} - \frac{K_{\nu+2}(x)}{x^\nu}\bigg[&x^\nu I_{\nu+1}(x)\tilde{h}(x) - x^{\nu+1}I_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &+ \int_0^x y^{\nu+2}I_{\nu+2}(y)\frac{\mathrm{d}}{\mathrm{d}y}\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg(\frac{\tilde{h}(y)}{y}\bigg)\bigg)\,\mathrm{d}y\bigg] \\ &+ \frac{K_{\nu+1}(x)}{x^{\nu+1}}\bigg[x^\nu I_{\nu+1}(x)\tilde{h}(x) - \int_0^x y^{\nu+1}I_{\nu+1}(y)\frac{\mathrm{d}}{\mathrm{d}y}\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\ &- \frac{I_{\nu+2}(x)}{x^\nu}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + x^{\nu+1}K_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &+ \int_0^x y^{\nu+2}K_{\nu+2}(y)\frac{\mathrm{d}}{\mathrm{d}y}\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg(\frac{\tilde{h}(y)}{y}\bigg)\bigg)\,\mathrm{d}y\bigg] \\ &+ \frac{I_{\nu+1}(x)}{x^{\nu+1}}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + \int_0^x y^{\nu+1}K_{\nu+1}(y)\frac{\mathrm{d}}{\mathrm{d}y}\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \end{aligned}$$

Cancelling terms and using the Wronskian formula (B.22) gives

$$\begin{aligned} f''(x) = &-\frac{K_{\nu+2}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\ &- \frac{K_{\nu+1}(x)}{x^{\nu+1}}\int_0^x y^\nu I_{\nu+1}(y)\bigg(h'(y) - \frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y \\ &- \frac{I_{\nu+2}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\ &- \frac{I_{\nu+1}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+1}(y)\bigg(h'(y) - \frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y. \end{aligned}$$

Note that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}(x^\nu I_{\nu+2}(x)) &= \frac{\mathrm{d}}{\mathrm{d}x}(x^{-2}\cdot x^{\nu+2}I_{\nu+2}(x)) = x^\nu I_{\nu+1}(x) - 2x^{\nu-1}I_{\nu+2}(x); \\ \frac{\mathrm{d}}{\mathrm{d}x}(x^\nu K_{\nu+2}(x)) &= \frac{\mathrm{d}}{\mathrm{d}x}(x^{-2}\cdot x^{\nu+2}K_{\nu+2}(x)) = -x^\nu K_{\nu+1}(x) - 2x^{\nu-1}K_{\nu+2}(x). \end{aligned}$$

Therefore, by integration by parts, we have

$$\begin{aligned} -\frac{K_{\nu+2}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+2}(y)h''(y)\,\mathrm{d}y &= h'(x)I_{\nu+2}(x)K_{\nu+2}(x) \\ &\quad+\frac{K_{\nu+2}(x)}{x^\nu}\int_0^x h'(y)[y^\nu I_{\nu+1}(y) - 2y^{\nu-1}I_{\nu+2}(y)]\,\mathrm{d}y \\ -\frac{I_{\nu+2}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+2}(y)h''(y)\,\mathrm{d}y &= -h'(x)I_{\nu+2}(x)K_{\nu+2}(x) \\ &\quad-\frac{I_{\nu+2}(x)}{x^\nu}\int_x^\infty h'(y)[y^\nu K_{\nu+1}(y) + 2y^{\nu-1}I_{\nu+2}(y)]\,\mathrm{d}y. \end{aligned}$$

Thus we have the following expression for $f''(x)$:

$$\begin{aligned}
f''(x) = &\frac{K_{\nu+2}(x)}{x^{\nu}}\int_0^x y^{\nu}I_{\nu+1}(y)h'(y)\,\mathrm{d}y - \frac{K_{\nu+2}(x)}{x^{\nu}}\int_0^x y^{\nu}I_{\nu+2}(y)\bigg(-\frac{h'(y)}{y}+\frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\
&-\frac{K_{\nu+1}(x)}{x^{\nu+1}}\int_0^x y^{\nu}I_{\nu+1}(y)\bigg(h'(y)-\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y \\
&-\frac{I_{\nu+2}(x)}{x^{\nu}}\int_x^{\infty} y^{\nu}K_{\nu+1}(y)h'(y)\,\mathrm{d}y - \frac{I_{\nu+2}(x)}{x^{\nu}}\int_x^{\infty} y^{\nu}K_{\nu+2}(y)\bigg(-\frac{h'(y)}{y}+\frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\
&-\frac{I_{\nu+1}(x)}{x^{\nu+1}}\int_x^{\infty} y^{\nu}K_{\nu+1}(y)\bigg(h'(y)-\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y,
\end{aligned}$$

We bound the integrals using Lemma D.8 to obtain

$$\begin{aligned}
\sup_{x\geq 1}|f''(x)| &\leq \bigg(\frac{1}{2\nu+2}+\frac{1}{2\nu+2}+\frac{1}{2\nu+2}+\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}\bigg)\|h'\| \\
&\quad+\bigg(\frac{3}{2\nu+1}+\frac{1}{2\nu+1}+\frac{3\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}\bigg)\|\tilde{h}\| \\
&\leq 3\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{1}{2\nu+2}\bigg)\|h'\|+4\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{1}{2\nu+1}\bigg)\|\tilde{h}\|,
\end{aligned}$$

as required. □

Combining the bounds for $0\leq x\leq 1$ and $x\geq 1$ that are given in Lemmas 3.18 and 3.19 we have:

Lemma 3.20. *Let $\nu > -1/2$, then the solution, and its first four derivatives, of the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation (3.13) are bounded as follows:*

$$\begin{aligned}
\|f\| &\leq \bigg(\frac{1}{2\nu+1}+\frac{\pi\Gamma(\nu+1/2)}{2\Gamma(\nu+1)}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|; \\
\|f'\| &\leq \bigg(\frac{1}{2\nu+1}+\frac{1}{2\nu+2}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|; \\
\|f''\| &\leq 3\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{1}{2\nu+1}\bigg)\|h'\|+4\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+5}}+\frac{1}{2\nu+1}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|; \\
\|f^{(3)}\| &\leq \max\bigg\{\frac{4}{2\nu+1},5\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\bigg)\bigg\}\|h''\| \\
&\quad+\max\bigg\{\frac{5}{2\nu+1},18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\bigg)\bigg\}\|h'\| \\
&\quad+\max\bigg\{\frac{25}{12(2\nu+1)}+\frac{1}{2(\nu+1)}+\frac{1}{v_1(\nu)},18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\bigg)\bigg\}\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}\|,
\end{aligned}$$

$$\begin{aligned}\|f^{(4)}\| \;\leq\; & \max\left\{\frac{8}{2\nu+1}, 8\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+13}}+\frac{1}{2\nu+4}\Big)\right\}\|h^{(3)}\| \\ & +\max\left\{\frac{9}{2\nu+1}, 52\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\Big)\right\}\|h''\| \\ & +\max\left\{\frac{77}{20(2\nu+1)}+\frac{1}{v_2(\nu)}, 123\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\Big)\right\}\|h'\| \\ & +\max\left\{\frac{1}{2\nu+2}+\frac{\sqrt{\pi}}{2(2\nu+1)}+\frac{2779}{768(2\nu+1)}+\frac{1}{v_3(\nu)}, 123\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\Big)\right\} \\ & \times\|h-\tilde{\mathrm{VG}}^{\nu,1}_{0,0}\|,\end{aligned}$$

where $v_1(\nu)$, $v_2(\nu)$ and $v_3(\nu)$ are defined as in Lemma 3.18. For $\nu\in\{0,1/2,1,3/2\ldots\}$, the bounds for $f^{(3)}(x)$ and $f^{(4)}(x)$ simplify as follows:

$$\begin{aligned}\|f^{(3)}\| \;\leq\; & 5\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\Big)\|h''\|+18\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\Big)\|h'\| \\ & +18\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\Big)\|h-\tilde{\mathrm{VG}}^{\nu,1}_{0,0}h\|; \\ \|f^{(4)}\| \;\leq\; & 8\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+13}}+\frac{1}{2\nu+1}\Big)\|h^{(3)}\|+52\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\Big)\|h''\| \\ & +123\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\Big)\|h'\|+123\Big(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\Big)\|h-\tilde{\mathrm{VG}}^{\nu,1}_{0,0}h\|.\end{aligned}$$

Proof. Combining Lemmas 3.18 and 3.19 we are able to bound f and its derivatives upto fourth order for all $x\geq 0$. In Remark 3.15 we noted that to obtain bounds on f and its derivatives for all $x\in\mathbb{R}$ it was sufficient to obtain bounds for all $x\geq 0$, and thus the proof is complete. □

Lemma 3.20 gives bounds for the first four derivatives of the solution of the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation (3.13) for all $\nu>-1/2$. The bounds are of order $\nu^{-1/2}$ as $\nu\to\infty$, except in the case that 2ν is very close to but not equal to an integer, in which case a term involving $1/\sin(\pi\nu)$ results in a poor bound. The $1/\sin(\pi\nu)$ term enters our analysis in Lemma 3.18, and the term is a result of using the series expansion (B.2) in proofs of Lemmas (D.16), (D.17) and (D.18) (these lemmas are used in the proof of Lemma 3.18), and we therefore conjecture that a bound of the form $C(2\nu+1)^{-1/2}$ holds for all $\nu>-1/2$, where C is a constant independent of ν.

The bounds simplify in the case that $\nu\in\{0,1/2,1,3/2,\ldots\}$ and from these bounds we use a simple change of variables to obtain uniform bounds on the first derivatives of the solution of the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14). These bounds are of order $r^{-1/2}$ for all $r\in\mathbb{Z}^+$, which is the same order as Pickett's [55] smoothness estimates (2.23) for the solution of the $\Gamma(r,\lambda)$ Stein equation (2.14).

We are now finally able to present our bounds for the derivatives of the solution of the

$\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14), which will be used in the limit theorems of Chapters 4 and 5.

Theorem 3.21. *Suppose that $h \in C_b^3(\mathbb{R})$, and that $r \in \mathbb{Z}^+$ and $\sigma > 0$, then the solution of the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14) and its derivatives upto fourth order satisfy*

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

where

$$\begin{aligned}
M_{r,\sigma}^0(h) &\leq \frac{1}{\sigma}\left(\frac{1}{r} + \frac{\pi\Gamma(r/2)}{2\Gamma(r/2+1/2)}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{r,0}h\|, \\
M_{r,\sigma}^1(h) &\leq \frac{1}{\sigma^2}\left(\frac{1}{r} + \frac{1}{r+1}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{r,0}h\|, \\
M_{r,\sigma}^2(h) &\leq \frac{3}{\sigma^2}\left(\frac{\sqrt{\pi}}{\sqrt{2r+3}} + \frac{1}{r}\right)\|h'\| + \frac{4}{\sigma^3}\left(\frac{\sqrt{\pi}}{\sqrt{2r+3}} + \frac{1}{r}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{r,0}h\|, \\
M_{r,\sigma}^3(h) &\leq \frac{5}{\sigma^2}\left(\frac{\sqrt{\pi}}{\sqrt{2r+7}} + \frac{1}{2r+2}\right)\|h''\| + \frac{18}{\sigma^3}\left(\frac{\sqrt{\pi}}{\sqrt{2r+7}} + \frac{1}{2r+1}\right)\|h'\| \\
&\quad + \frac{18}{\sigma^4}\left(\frac{\sqrt{\pi}}{\sqrt{2r+7}} + \frac{1}{r}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{r,0}h\|, \\
M_{r,\sigma}^4(h) &\leq \frac{8}{\sigma^2}\left(\frac{\sqrt{\pi}}{\sqrt{r+11}} + \frac{1}{r}\right)\|h^{(3)}\| + \frac{52}{\sigma^3}\left(\frac{\sqrt{\pi}}{\sqrt{2r+7}} + \frac{1}{r+2}\right)\|h''\| \\
&\quad + \frac{123}{\sigma^4}\left(\frac{\sqrt{\pi}}{\sqrt{2r+7}} + \frac{1}{r+1}\right)\|h'\| + \frac{123}{\sigma^5}\left(\frac{\sqrt{\pi}}{\sqrt{2r+7}} + \frac{1}{r}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{r,0}h\|.
\end{aligned}$$

Proof. Let $g_{\tilde{h}}(x)$ denote the solution (3.15) to the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation (3.13)

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

Then $f_h(x) = \frac{1}{\sigma}g_{\tilde{h}}(\frac{x-\mu}{\sigma})$ solves the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14)

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

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

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

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 $\|\tilde{h} - \tilde{\mathrm{VG}}_{0,0}^{\nu,1}\tilde{h}\| = \|h - \mathrm{VG}_{\sigma,\mu}^{r,0}h\|$ and $\|\tilde{h}^{(k)}\| = \sigma^k\|h^{(k)}\|$ for $k \geq 1$, and the result now follows from the bounds of Lemma 3.20. □

For the case $r = 1$ these bounds can be improved as follows:

Theorem 3.22. *Suppose that $h \in C_b^3(\mathbb{R})$. Then the unique bounded solution f of the* $\mathrm{VG}(1,0,1,0)$ *Stein equation (3.13) and its first four derivatives satisfy*

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

where

$$\begin{aligned} N_\sigma^0(h) &= \frac{1}{\sigma}\left(1+\frac{\pi}{2}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{1,0}h\|, \\ N_\sigma^1(h) &= \frac{3}{2\sigma^2}\|h - \mathrm{VG}_{\sigma,\mu}^{1,0}h\|, \\ N_\sigma^2(h) &= \frac{2}{\sigma^2}\|h'\| + \frac{1}{\sigma^3}\left(\frac{13}{4}+\frac{\sqrt{\pi}}{2}\right)\|h - \mathrm{VG}_{\sigma,\mu}^{1,0}h\|, \\ N_\sigma^3(h) &= \frac{4}{\sigma^2}\|h''\| + \frac{5}{\sigma^3}\|h'\| + \frac{4.89}{\sigma^4}\|h - \mathrm{VG}_{\sigma,\mu}^{1,0}h\|, \\ N_\sigma^4(h) &= \frac{8}{\sigma^2}\|h^{(3)}\| + \frac{9}{\sigma^3}\|h''\| + \frac{6.81}{\sigma^4}\|h'\| + \frac{15.75}{\sigma^5}\|h - \mathrm{VG}_{\sigma,\mu}^{1,0}h\|. \end{aligned}$$

Proof. The proof is similar to that of Lemma 3.18, the only difference being that we are able to obtain bounds for all $x \in \mathbb{R}$ (recall Remark 3.15) by using the bounds of Lemma D.20, which hold for all $x \geq 0$, instead of the bounds of Lemmas D.15, D.16, D.17 and D.18, which only hold for $0 \leq x \leq 1$. We then rescale as we did in the proof of Theorem 3.21. □

To summarise the work of this section, we have been able to establish uniform bounds for the solution of the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation. Our bounds arose from detailed calculations, in which we bounded the derivatives in two different regions and then combined these bounds to obtain our final bounds. This approach lead to bounds that are of order $r^{-1/2}$ for $r \in \mathbb{Z}^+$, which is the same order as the smoothness estimates for the solution of the $\Gamma(r,\lambda)$ Stein equation that were given by Pickett [55].

These smoothness estimates enable us to use Stein's method to study the asymptotic behaviour of many important statistics. In particular, under certain parameter values the D_2^* statistic will fall in $\mathrm{VG}(r,0,1,0)$ regime. In Chapter 4 we apply the theory of Stein method for Variance-Gamma distributions that we have developed in this chapter to obtain bounds on the error in approximating a statistic that is of similar form to D_2^*, but with the simpler assumption of i.i.d.

random variables, by its limiting Variance-Gamma distribution.

However, we have been unable to obtain smoothness estimates for the full class of Variance-Gamma distributions. It would be desirable to achieve this goal, which would allow us to use Stein's method to study the limiting distributions of many other interesting statistics – see Corollary 3.7 and Proposition 3.8 for examples of distributions that fall in the $\mathrm{VG}(r,\theta,\sigma,\mu)$ regime for the case that both θ and σ are non zero.

In Chapter 6 we demonstrate how the multivariate normal Stein equation (2.12) can be used to obtain bounds on the error in approximating statistics that are asymptotically distributed as a function of the multivariate normal distribution. Since many Variance-Gamma distributions can be characterised in terms of normal and χ^2 (a function of normal random variables) distributions (see Corollary 3.7), we would still be able to obtain bounds for the approximation error; we would just use the multivariate normal Stein equation instead of the $\mathrm{VG}(r,\theta,\sigma,\mu)$ Stein equation.

3.3.2 Smoothness estimates by a generator approach

Here we consider how the generator approach could be used to obtain smoothness estimates for the solution of the Variance-Gamma Stein equation. We begin by stating some standard results for generators and Bessel processes, which we need later.

Theorem 3.23. *Let $\mathcal{A}$ be the generator of a stochastic process $(X_t)_{t\geq 0}$, and suppose that g and c are bounded measurable functions with $c(x)\geq 0$. Let T_t denote the transition semigroup operator, defined by $T_tg(x)=\mathbb{E}\{g(X_t)\,|\,X_0=x\}$. Then the solution to the inhomogeneous differential equation*

$$\mathcal{A}f(x)-c(x)f(x)=g(x)$$

is given by

$$f(x)=-\int_0^\infty \mathrm{e}^{-c(x)t}T_tg(x)\,\mathrm{d}t,$$

if the integral exists.

Proof. The result is proved for the special case $c(x)=\lambda$, where $\lambda>0$, on p. 249 of Durrett [20]. The proof can be easily adapted to obtain the result for general $c(x)\geq 0$. □

The Bessel process with drift is defined in Linetsky [42] as follows:

Definition 3.24. *Let $x\geq 0$ and $\nu>-1/2$. The unique strong solution to the stochastic differential equation*

$$\mathrm{d}X_t=\left(\frac{2\nu+1}{2X_t}+\beta\right)\mathrm{d}t+\mathrm{d}B_t,\quad X_0=x,$$

where B_t denotes standard Brownian motion, is called the Bessel process of index ν and drift β started at x.

Linetsky [42] gives the following formula for the transition density of a Bessel process with drift.

Proposition 3.25. *The transition density for a Bessel process of index ν and drift β started at x is given by*

$$\begin{aligned} p(y,x,t,\nu,\beta) = \frac{1}{2\pi}\int_0^\infty & \mathrm{e}^{-(\beta^2+u^2)t/2}\left(\frac{y}{x}\right)^{\nu+1/2}\mathrm{e}^{\beta(y-x)+\pi\eta/u}M_{i\eta/u,\nu}(-2ixu) \\ & \times M_{-i\eta/u,\nu}(2iyu)\left|\frac{\Gamma(\nu+1/2+i\eta/u)}{\Gamma(\nu+1)}\right|\,\mathrm{d}u, \qquad y\geq 0, \end{aligned} \tag{3.21}$$

where $\eta=(\nu+1/2)\beta$ and $M_{\chi,\nu}(\cdot)$ is the Whittaker function (see Olver et al. [52]).

In the case $\beta=0$ the transition density simplifies as follows (see Theorem 9.1 of Kent [39]):

Proposition 3.26. *The transition density for a Bessel process of index ν started at x is given by*

$$p(y,x,t,\nu) = x^{-\nu}y^{\nu+1}t^{-1}\exp\left(-\frac{x^2+y^2}{2t}\right)I_\nu\left(\frac{xy}{t}\right), \qquad y\geq 0. \tag{3.22}$$

Therefore the transition semigroup operator of a Bessel process with index ν acting on the function $g(x)$ is given by

$$T_t^{(\nu)}g(x) = \int_0^\infty g(y)p(y,x,t,\nu)\,\mathrm{d}y, \tag{3.23}$$

where $p(y,x,t,\nu)$ is the transition density of a Bessel process of index ν, which is given by equation (3.22).

We can apply Theorem 3.23 and the transition density formula (3.21) to write down the solution of the $\mathrm{VG}_2(\nu,1,\beta,0)$ Stein equation (3.13). However, this formula for the transition density is very complicated, so we restrict our attention to the case $\beta=0$, for which the transition density takes a much simpler form.

Lemma 3.27. *Let $\tilde{h}(x)=h(x)-\mathrm{VG}_{0,0}^{\nu,1}h$, where $h:\mathbb{R}\rightarrow\mathbb{R}$ is a bounded function. Then the solution to the Variance-Gamma Stein equation (3.13) in the region $x>0$ is*

$$f(x) = -\frac{1}{2}\int_0^\infty \mathrm{e}^{-t/2}T_t^{(\nu)}\left(\frac{\tilde{h}(x)}{x}\right)\mathrm{d}t. \tag{3.24}$$

Proof. Dividing both sides of the Stein equation (3.13), with $\beta=0$, by $2x$ gives

$$\frac{1}{2}f''(x)+\frac{2\nu+1}{2x}f'(x)-\frac{1}{2}f(x)=\frac{\tilde{h}(x)}{2x}.$$

We recognise $\frac{1}{2}f''(x)+\frac{2\nu+1}{2x}f'(x)$ as the generator of a Bessel process of index ν and therefore applying Theorem 3.23 shows that (3.24) is a solution to the Stein equation. $\square$

Remark 3.28. We can verify that solution (3.24) is equal to solution (3.15) for $x>0$ (this is to be expected due to Lemma 3.13). Substituting (3.23) into (3.24) gives

$$f(x)=-\frac{1}{2}\int_0^\infty\int_0^\infty x^{-\nu}y^\nu t^{-1}\mathrm{e}^{-t/2}\exp\bigg(-\frac{x^2+y^2}{2t}\bigg)I_\nu\bigg(\frac{xy}{t}\bigg)\tilde{h}(y)\,\mathrm{d}y\,\mathrm{d}t.$$

Interchanging the order of integration by Tonelli's Theorem and integrating with respect to t by formula (B.44) verifies that the solutions are equal.

The solution (3.24) is of a very convenient form for our task of obtaining a formula for its k–th derivative. Provided we can differentiate under the integral sign, we would have

$$f^{(k)}(x)=-\frac{1}{2}\int_0^\infty\mathrm{e}^{-t/2}\frac{\mathrm{d}^k}{\mathrm{d}x^k}\bigg(T_t^{(\nu)}\bigg(\frac{\tilde{h}(x)}{x}\bigg)\bigg)\,\mathrm{d}t. \tag{3.25}$$

Hence our task of obtaining a formula for the k-th derivative of the solution (3.24) reduces to finding a formula for the k-th derivative of the semigroup $T_t^{(\nu)}w(x)$ for a Bessel process. Before obtaining such a formula we introduce some notation. We define

$$\mathbf{C}_b^k(I)=\{w\in C^k(I)\,:\,\exists M_w>0\text{ such that }\|w^{(j)}\|\leq M_w\text{ for }j=0,1,\ldots,k\},$$

where $I\subset\mathbb{R}$ is an interval.

In the following lemmas we present formulas for k-th derivative of semigroup $T_t^{(\nu)}w(x)$ for a Bessel process. The proof of Lemma 3.29 rests on the assumption that way can differentiate under the integral sign. Further work is needed to justify the interchange of the operations of differentiation and integration.

Lemma 3.29. *Suppose $w\in\mathbf{C}_b^1((0,\infty))$. Then for $x>0$ the semigroup $T_t^{(\nu)}w(x)$ of a Bessel process with index ν satisfies the following identity*

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(T_t^{(\nu)}w(x)\Big)=xT_t^{(\nu+1)}\bigg(\frac{1}{x}\frac{\mathrm{d}w(x)}{\mathrm{d}x}\bigg). \tag{3.26}$$

Proof. Provided we can differentiate under the integral sign, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big(T_t^{(\nu)}w(x)\Big)=\int_0^\infty w(y)\frac{\partial}{\partial x}p(y,x,t,\nu)\,\mathrm{d}y. \tag{3.27}$$

We now obtain an expression for the partial derviative of the density with respect to x, which will allows us to obtain a simple expression for a derivative of the semigroup. Using (B.31) we have

$$\begin{aligned}\frac{\partial}{\partial x}p(y,x,t,\nu) &= y^{\nu+1}t^{-1}\exp\left(-\frac{x^2+y^2}{2t}\right)\left(\frac{\partial}{\partial x}\left(x^{-\nu}I_\nu\left(\frac{xy}{t}\right)\right)-\frac{x}{t}\cdot x^{-\nu}I_\nu\left(\frac{xy}{t}\right)\right)\\ &= y^{\nu+1}t^{-1}\exp\left(-\frac{x^2+y^2}{2t}\right)\left(\frac{y}{t}x^{-\nu}I_{\nu+1}\left(\frac{xy}{t}\right)-\frac{x}{t}\cdot x^{-\nu}I_\nu\left(\frac{xy}{t}\right)\right)\\ &= t^{-1}x(p(y,x,t,\nu+1)-p(y,x,t,\nu)).\end{aligned}$$

A similar calculation gives

$$\begin{aligned}\frac{\partial}{\partial y}p(y,x,t,\nu) &= y^{-1}p(y,x,t,\nu)+yt^{-1}p(y,x,t,\nu-1)-yt^{-1}p(y,x,t,\nu)\\ &= y^{-1}p(y,x,t,\nu)-\frac{y}{x}\frac{\partial}{\partial x}p(y,t,x,\nu-1).\end{aligned}$$

Putting these two calculations together gives

$$\frac{\partial}{\partial x}p(y,x,t,\nu)=\frac{x}{y^2}p(y,x,t,\nu+1)-\frac{x}{y}\frac{\partial}{\partial y}p(y,x,t,\nu+1).$$

Substituting this expression into (3.27) and integrating by parts yields

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(T_t^{(\nu)}w(x)\right) &= \int_0^\infty w(y)\left(\frac{x}{y^2}p(y,x,t,\nu+1)-\frac{x}{y}\frac{\partial}{\partial y}p(y,x,t,\nu+1)\right)\mathrm{d}y\\ &= \int_0^\infty \frac{x\tilde{h}(y)}{y^2}p(y,x,t,\nu+1)\,\mathrm{d}y-\left[\frac{x\tilde{h}(y)}{y}p(y,x,t,\nu+1)\right]_0^\infty\\ &\quad+\int_0^\infty x\frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{\tilde{h}(y)}{y}\right)p(y,x,t,\nu+1)\,\mathrm{d}y\\ &= \int_0^\infty \frac{x}{y}\frac{\mathrm{d}w(y)}{\mathrm{d}y}p(y,x,t,\nu+1)\,\mathrm{d}y\\ &= xT_t^{(\nu+1)}\left(\frac{1}{x}\frac{\mathrm{d}w(x)}{\mathrm{d}x}\right).\end{aligned}$$

The term in square brackets vanishes as $\|g\|\leq M_g$, for some constant M_g and because the function $xy^{-1}p(y,x,t,\nu+1)$ tends to zero in the limits $y\rightarrow 0$ and $y\rightarrow\infty$, which can be seen from the asymptotic expansions (B.7) and (B.10) of $I_\nu(x)$. □

We may use Lemma 3.29 to obtain a formula for the k-derivative of $T_t^{(\nu)}w(x)$. Before stating the lemma we introduce some notation.

For $n=1,2,3,\ldots$ we define the operator $(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x})^n$ recursively by

$$\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^1 f(x)=\frac{f'(x)}{x},\qquad \left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^n=\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{n-1}f(x)\right],\quad n=2,3,4,\ldots. \tag{3.28}$$

With this notation we have:

Lemma 3.30. *Let $k\in\mathbb{N}$. Suppose that $w\in\mathbf{C}_b^{2k}((0,\infty))$, then for $x>0$ the semigroup $T_t^{(\nu)}w(x)$ of a Bessel process with index ν satisfies the identity*

$$\frac{\mathrm{d}^{2k}}{\mathrm{d}x^{2k}}\left(T_t^{(\nu)}w(x)\right)=\sum_{j=0}^{k}\frac{2^{j-k}(2k)!}{(k-j)!\,(2j)!}x^{2j}T_t^{(\nu+k+j)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+j}w(x)\right). \tag{3.29}$$

Now suppose that $g\in\mathbf{C}_b^{2k+1}((0,\infty))$, then for $x>0$ we have

$$\frac{\mathrm{d}^{2k+1}}{\mathrm{d}x^{2k+1}}\left(T_t^{(\nu)}w(x)\right)=\sum_{j=0}^{k}\frac{2^{j-k}(2k+1)!}{(k-j)!\,(2j+1)!}x^{2j+1}T_t^{(\nu+k+j+1)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+j+1}w(x)\right). \tag{3.30}$$

Proof. We prove the result by induction on k. Formula (3.29) is true for $k=0$ and (3.30) holds for $k=0$ by Lemma 3.30. Suppose that (3.30) is true for $k\geq 0$. Then, using Lemma 3.30, we have

$$\begin{aligned}\frac{\mathrm{d}^{2k+2}}{\mathrm{d}x^{2k+2}}\left(T_t^{(\nu)}w(x)\right)&=\sum_{j=0}^{k}\frac{2^{j-k}(2k+1)!}{(k-j)!\,(2j+1)!}\bigg\{(2j+1)x^{2j}T_t^{(\nu+k+j+1)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+j+1}w(x)\right)\\&\quad+x^{2j+2}T_t^{(\nu+k+j+2)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+j+2}w(x)\right)\bigg\}\\&=\sum_{i=0}^{k}\left[\frac{2^{i+1-k}(2k+1)!}{(k-i)!\,(2i+3)!}\cdot(2i+3)+\frac{2^{i-k}(2k+1)!}{(k-i)!\,(2i+1)!}\right]\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+i+1}w(x)\right)\\&\quad+x^{2k+2}T_t^{(\nu+2k+2)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{2k+2}w(x)\right).\end{aligned}$$

A straightforward calculation shows that

$$\frac{2^{i-k}(2k+1)!}{(k-i)!\,(2i+1)!}+\frac{2^{i+1-k}(2k+1)!}{(k-i)!\,(2i+2)!}=\frac{2^{i-k-1}(2k+2)!}{(k+1-i)!\,(2i)!},$$

and therefore

$$\frac{\mathrm{d}^{2k+2}}{\mathrm{d}x^{2k+2}}\left(T_t^{(\nu)}w(x)\right)=\sum_{j=0}^{k+1}\frac{2^{j-k-1}(2k+2)!}{(k+1-j)!\,(2j)!}x^{2j}T_t^{(\nu+k+j+1)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+j+1}w(x)\right),$$

as required. Suppose now that (3.29) is true for $k \geq 1$. A similar calculation shows that

$$\frac{\mathrm{d}^{2k+1}}{\mathrm{d}x^{2k+1}}\left(T_t^{(\nu)}w(x)\right) = \sum_{j=0}^{k} \frac{2^{j-k}(2k+1)!}{(k-j)!\,(2j+1)!} x^{2j+1} T_t^{(\nu+k+j+1)}\left(\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\right)^{k+j+1} w(x)\right).$$

Hence the result has been proved by induction on k. $\square$

Therefore, provided we can differentiate under the integral sign, we can obtain a formula for the k-th derivative of the solution to the Symmetric Variance-Gamma Stein equation by substituting the formulas from Lemma 3.30 into (3.25). However, these formulas for the derivatives of $T_t^{(\nu)}(\tilde{h}(x)/x)$ are complicated and bounding them does not appear to be straightforward. If we could simplify these formulas for the k-th derivative of $T_t^{(\nu)}(\tilde{h}(x)/x)$ then it may be possible to obtain bounds for the k-th derivative of the solution to the Symmetric Variance-Gamma Stein equation. This is left as an interesting open problem.

Chapter 4

Limit Theorems for Symmetric Variance-Gamma distributions

In this chapter we apply Stein's method for Variance-Gamma distributions in two different settings. Let $\mathbf{X}$ be a $m \times r$ matrix of i.i.d. random variables X_{ik} with zero mean and unit variance. Similarly, we let $\mathbf{Y}$ be a $n \times r$ matrix of i.i.d. random variables Y_{jk} with zero mean and unit variance, where the random variables Y_{jk} are independent of the random variables X_{ik}. Then the statistics

$$U_r = \frac{1}{m}\sum_{k=1}^{r}\left(\sum_{i=1}^{m}X_{ik}\right)^2 - \frac{1}{n}\sum_{k=1}^{r}\left(\sum_{j=1}^{n}Y_{jk}\right)^2, \tag{4.1}$$

$$V_r = \frac{1}{\sqrt{mn}}\sum_{i,j,k=1}^{m,n,r}X_{ik}Y_{jk} \tag{4.2}$$

are asymptotically $\mathrm{VG}(r,0,2,0)$ and $\mathrm{VG}(r,0,1,0)$ distributed respectively, by simple application of the central limit theorem. Indeed, if $X_{ik} \sim N(0,1)$ and $Y_{jk} \sim N(0,1)$ for all i,j,k, then by Proposition 3.8 and Corollary 3.7, $U_r \sim \mathrm{VG}(r,0,2,0)$ and $V_r \sim \mathrm{VG}(r,0,1,0)$ exactly. Pickett [55] showed that the statistic $W_d = \frac{1}{m}\sum_{k=1}^{d}(\sum_{i=1}^{m}X_{ik})^2$, where the X_{ik} are i.i.d. random variables with zero mean, unit variance and bounded eighth moment, converges to a $\chi^2_{(d)}$ random variable at a rate of order m^{-1} for smooth test functions. This rate of convergence is faster than a convergence rate of order $m^{-1/2}$ that might have been expected from the Berry-Esséen Theorem, and was achieved through the use of novel symmetry arguments. We now exhibit proofs which gives a bound for the rate of convergence of the statistics U_r and V_r to $\mathrm{VG}(r,0,2,0)$ and $\mathrm{VG}(r,0,1,0)$ random variables respectively, under additional moment assumptions, which are shown to be of order $m^{-1}+n^{-1}$ for smooth test functions, using similar symmetry arguments, as in Pickett [55] to obtain this rate of convergence. However, for non-smooth test functions we still expect a Berry-Esséen type $m^{-1/2}+n^{-1/2}$ to be of optimal order – see Remark 4.15 for why this is the case.

We are able to apply our bounds for the error in approximating V_r by its limiting the $\mathrm{VG}(r,0,1,0)$ distribution to an application of binary sequence comparison. This is a special case of the word sequence comparison problem that was discussed in Section 1.1: we consider an alphabet size $d=2$ and a statistic based on 1-tuples.

We end this chapter by providing a Variance-Gamma limit theorem in which the random variables have a local dependence structure, and in doing so demonstrate the tractability of the Variance-Gamma Stein equation even when the assumption of i.i.d. random variables is dropped.

4.1 A Limit Theorem for the difference of two χ^2 distributions

The following theorem, which is Theorem 4.7 of Pickett [55], allows us to easily obtain a bound on the error in approximating the statistic (4.1) by its limiting $\mathrm{VG}(r,0,2,0)$ distribution.

Theorem 4.1. *Suppose $\mathbf{X}$ is a be a $m\times r$ matrix of i.i.d. random variables X_{ik} with zero mean, unit variance and bounded eighth moment. Let X be a random variable with the same distribution as the X_{ik}. Let $W_d=\frac{1}{m}\sum_{k=1}^{d}(\sum_{i=1}^{m}X_{ik})^2$. Then, for $h\in\mathcal{C}_{1/2}\cap C_b^3(\mathbb{R})$, we have*

$$|\mathbb{E}h(W_d)-\chi^2_{(d)}h|\leq\frac{1}{m}\bigg((\sqrt{2\pi}+\mathrm{e}^{-1})\sqrt{\frac{d}{2}}+1\bigg)(\alpha_1(X)\|h'\|+\alpha_2(X)\|h''\|+\alpha_3(X)\|h^{(3)}\|),\quad(4.3)$$

where $\chi^2_{(d)}h$ denotes the expectation of $h(Z)$, for $Z\sim\chi^2_{(d)}$, and

$$\begin{aligned}\alpha_1(X)&=6\bigg(\mathbb{E}|X|+\frac{1}{2}+\frac{\mathbb{E}X^4}{6}\bigg)+3|\mathbb{E}X^3|\bigg[(\mathbb{E}|X|+\mathbb{E}|X|^3)\bigg(\bigg(3+\frac{\mathbb{E}X^4}{m}\bigg)^{3/4}+5\bigg)\\&\quad+\frac{4\mathbb{E}X^4}{\sqrt{m}}(1+\mathbb{E}X^4)+\frac{5}{m}(\mathbb{E}|X|^3+\mathbb{E}|X|^5)+\frac{1}{m\sqrt{m}}(\mathbb{E}X^4+\mathbb{E}X^6)\bigg],\\ \alpha_2(X)&=24\bigg[\frac{1}{2}+\mathbb{E}|X|+\frac{\mathbb{E}|X|^3}{m}+\mathbb{E}X^4\bigg(\frac{1}{6}+\frac{1}{2m}\bigg)+\frac{\mathbb{E}X^6}{6m}\bigg]+2|\mathbb{E}X^3|\bigg[6\mathbb{E}|X|\\&\quad+3\mathbb{E}|X|^3\bigg(2+\frac{1}{m}\bigg)+\frac{3\mathbb{E}|X|^5}{m}=2\bigg(10+\frac{\mathbb{E}X^4}{m}\bigg)(\mathbb{E}|X|+\mathbb{E}|X|^3)\\&\quad+\frac{1}{m^2}(\mathbb{E}|X|^5+\mathbb{E}|X|^7)+\frac{4|\mathbb{E}X^3|}{m}(I\{\mathbb{E}(|X|X)\mathbb{E}X^3\geq 0\}+I\{\mathbb{E}(|X^3|X)\mathbb{E}X^3\geq 0\})\bigg],\end{aligned}$$

$$\alpha_3(X) = 48\left[\frac{1}{2} + \mathbb{E}|X| + \frac{\mathbb{E}|X|^3}{m} + \mathbb{E}X^4\left(\frac{1}{6} + \frac{1}{2m}\right) + \frac{\mathbb{E}X^6}{6m}\right] + \frac{32}{m}\left[\mathbb{E}X^4\left(\mathbb{E}|X| + \frac{1}{2} + \frac{\mathbb{E}X^4}{6}\right)\right.$$
$$+ |\mathbb{E}X^3|(I\{\mathbb{E}(|X|X)\mathbb{E}X^3 \geq 0\}\mathbb{E}(|X|X) + \frac{1}{6}I\{\mathbb{E}X^3\mathbb{E}X^5 \geq 0\}\mathbb{E}X^5)\Big] + \frac{8}{m^2}\left(\mathbb{E}|X|^5\right.$$
$$\left. + \frac{\mathbb{E}X^6}{2} + \frac{\mathbb{E}X^8}{6}\right) + 4|\mathbb{E}X^3|\left[\left(6 + \frac{\mathbb{E}X^4}{m}\right)(\mathbb{E}|X| + \mathbb{E}|X|^3) + \frac{6}{m}(\mathbb{E}|X|^3 + \mathbb{E}|X|^5)\right.$$
$$\left. + \frac{1}{m^2}(\mathbb{E}|X|^5 + \mathbb{E}|X|^7) + \frac{4|\mathbb{E}X^3|}{m}(I\{\mathbb{E}(|X|X)\mathbb{E}X^3 \geq 0\} + I\{\mathbb{E}(|X^3|X)\mathbb{E}X^3 \geq 0\})\right].$$

Remark 4.2. The bound (4.3) we present in Theorem 4.1 differs slightly from the one given in Pickett [55]. This is because we used the smoothness estimates (2.35) for the solution of the $\chi^2_{(d)}$ Stein equation (2.15), rather than the smoothness estimates (2.23) that were used by Pickett. This only changes the bound by a multiplicative constant that does not depend on m or d.

Examining Pickett's proof we see that he had the following bound:

$$|2g'(W_d) + (d - W_d)g(W_d)| \leq 2d(\alpha_1(X)\|g'\| + \alpha_2(X)\|g''\| + \alpha_3(X)\|g^{(3)}\|). \tag{4.4}$$

The bound given in Theorem 4.1 was then obtained by using the Stein equation for Gamma distributions (2.14) to write the left-hand side of (4.4) in terms of the test function h, and the bound (2.23) on the derivatives of the solution of the (2.14) was used to write the right-hand side of (4.4) in terms of the derivatives of h. We now apply this result to prove the following theorem.

Theorem 4.3. *Let* $\mathbf{X}$ *be a* $m \times r$ *matrix of i.i.d. random variables* X_{ik} *with zero mean, unit variance and bounded eighth moment. Similarly, let* $\mathbf{Y}$ *be a* $n \times r$ *matrix of i.i.d. random variables* Y_{jk} *with zero mean, unit variance and bounded eighth moment. Suppose further that the* σ*-fields* $\sigma\{X_{ik} : i = 1, \ldots, m,\ k = 1, \ldots, r\}$ *and* $\sigma\{Y_{jk} : j = 1, \ldots, n,\ k = 1, \ldots, r\}$ *are independent. Let* $W_r = \frac{1}{m}\sum_{k=1}^r(\sum_{i=1}^m X_{ik})^2 - \frac{1}{n}\sum_{k=1}^r(\sum_{i=1}^n Y_{ik})^2$*. Then for* $h \in C_b^3(\mathbb{R})$*, we have*

$$|\mathbb{E}h(W_r) - \mathrm{VG}^{r,0}_{2,0}| \leq r(\beta_1 M^1_{r,2}(h) + \beta_2 M^2_{r,2}(h) + \beta_3 M^3_{r,2}(h) + \beta_4 M^4_{r,2}(h)).$$

Here the $M^k_{r,2}(h)$ *are defined as in Theorem (3.21), and* $\mathrm{VG}^{r,0}_{2,0}h$ *denotes the expectation of* $h(Z)$*, for* $Z \sim \mathrm{VG}(r, 0, 2, 0)$*, and*

$$\begin{aligned}
\beta_1 &= 2\alpha_1(X) + 2\alpha_1(Y) \\
\beta_2 &= 4\alpha_1(X) + 4\alpha_1(Y) + 2\alpha_2(X) + 2\alpha_2(Y) \\
\beta_3 &= 4\alpha_2(X) + 4\alpha_2(Y) + 2\alpha_3(X) + 2\alpha_3(Y) \\
\beta_4 &= 4\alpha_3(X) + 4\alpha_3(Y),
\end{aligned}$$

where the α_i are defined as in Theorem 4.1.

Proof. Using the VG$(r,0,2,0)$ Stein equation (3.14), we require a bound on the the expression $\mathbb{E}\{4W_rf''(W_r)+4rf'(W_r)+(r-W_r)f(W_r)\}$. Let $U_r=\frac{1}{m}\sum_{k=1}^r(\sum_{i=1}^m X_{ik})^2$ and $V_r=\frac{1}{n}\sum_{k=1}^r(\sum_{j=1}^m Y_{jk})^2$. Then $W_r=U_r-V_r$. Let f_h be the solution of the VG$(r,0,2,0)$ Stein equation Using (4.4) we have

$$\begin{aligned}\mathbb{E}W_rf(W_r)&=\mathbb{E}U_rf(W_r)-\mathbb{E}V_rf(W_r)\\&=\mathbb{E}(\mathbb{E}U_rf(W_r)|V_r)-\mathbb{E}(\mathbb{E}V_rf(W_r)|U_r))\\&=\{2\mathbb{E}U_rf'(W_r)+r\mathbb{E}f(W_r)+R_1\}-\{-2\mathbb{E}V_rf'(W_r)+r\mathbb{E}f(W_r)+R_2\}\\&=2\mathbb{E}U_rf'(W_r)+2\mathbb{E}V_rf'(W_r)+R_1-R_2,\end{aligned}$$

where

$$\begin{aligned}|R_1|&\leq 2r\Big[\alpha_1(X)\|f_h'\|+\alpha_2(X)\|f_h''\|+\alpha_3(X)\|f_h^{(3)}\|\Big],\\|R_2|&\leq 2r\Big[\alpha_1(Y)\|f_h'\|+\alpha_2(Y)\|f_h''\|+\alpha_3(Y)\|f_h^{(3)}\|\Big].\end{aligned}$$

Here we used that $\|g(x+c)\|=\|g(x)\|$ to obtain the third equality. We now use (4.4) once more to obtain

$$\begin{aligned}\mathbb{E}W_rf(W_r)&=2\mathbb{E}U_rf'(W_r)+2\mathbb{E}V_rf'(W_r)+R_1-R_2\\&=2\mathbb{E}(\mathbb{E}U_rf'(W_r)|V_r)+2\mathbb{E}(\mathbb{E}V_rf'(W_r)|U_r)+R_1-R_2\\&=\{4\mathbb{E}U_rf''(W_r)+2r\mathbb{E}f'(W_r)+R_3\}+\{-4\mathbb{E}V_rf''(W_r)+2r\mathbb{E}f'(W_r)+R_4\}\\&\quad+R_1-R_2\\&=4\mathbb{E}W_rf''(W_r)+4r\mathbb{E}f'(W_r)+R_1-R_2+R_3+R_4,\end{aligned}$$

where

$$\begin{aligned}|R_3|&\leq 4r\Big[\alpha_2(X)\|f_h''\|+\alpha_3(X)\|f_h^{(3)}\|+\alpha_4(X)\|f_h^{(4)}\|\Big],\\|R_4|&\leq 4r\Big[\alpha_2(Y)\|f_h''\|+\alpha_3(Y)\|f_h^{(3)}\|+\alpha_4(Y)\|f_h^{(4)}\|\Big].\end{aligned}$$

This completes the proof. $\square$

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

We now prove a limit theorem for the statistic $W_r = \frac{1}{\sqrt{mn}} \sum_{i,j,k=1}^{m,n,r} X_{ik}Y_{jk}$. To elucidate the proof, we first consider the case $r = 1$; the general r case follows easily as W_r is a linear sum of independent W_1. Throughout the rest of this chapter we shall set $X_i \equiv X_{i1}$, $Y_j \equiv Y_{j1}$ and $W \equiv W_1$. Then we have the following:

Theorem 4.4. *Suppose $X, X_1, \ldots, X_m$ is a collection of i.i.d. random variables with zero mean, unit variance and bounded sixth moment, suppose also that $Y, Y_1, \ldots, Y_n$ is a collection of i.i.d. random variables with zero mean, unit variance and bounded sixth moment. Suppose further that the σ-fields $\sigma\{X_i : i = 1, \ldots, m\}$ and $\sigma\{Y_j : j = 1, \ldots, n\}$ are independent. Let $W = \frac{1}{\sqrt{mn}} \sum_{i,j=1}^{m,n} X_iY_j$. Then for $h \in C_b^3(\mathbb{R})$, we have*

$$|\mathbb{E}h(W) - \mathrm{VG}^{1,0}_{1,0}h| \leq \gamma_1(X,Y)N_1^2(h) + \gamma_2(X,Y)N_1^3(h) + \gamma_3(X,Y)N_1^4(h), \tag{4.5}$$

where the $N_1^k(h)$ are defined as in Theorem 3.22, and $\mathrm{VG}^{1,0}_{1,0}h$ denotes the expectation of $h(Z)$, for $Z \sim \mathrm{VG}(1,0,1,0)$, and

$$\begin{aligned}
\gamma_1(X,Y) &= \frac{\pi|\mathbb{E}Y^3|}{n}(\mathbb{E}|Y^3|+2) + \frac{3\pi|\mathbb{E}X^3|}{8\sqrt{mn}}\bigg(4 + 2\mathbb{E}|Y^3| + \frac{2}{\sqrt{n}} + \frac{\mathbb{E}Y^4}{\sqrt{n}}\bigg), \\
\gamma_2(X,Y) &= \frac{1}{6m}(9 + 3|\mathbb{E}X^3| + \mathbb{E}X^4)\bigg(3 + \frac{\mathbb{E}Y^4}{n}\bigg) + \frac{\pi|\mathbb{E}Y^3|}{n}\bigg(4 + 2\mathbb{E}|Y^3| + \frac{2}{\sqrt{n}} + \frac{\mathbb{E}Y^4}{\sqrt{n}}\bigg) \\
&\quad \frac{|\mathbb{E}Y^3|}{2n}(1 + \mathbb{E}|Y^3|) + \frac{3\pi|\mathbb{E}X^3|}{\sqrt{mn}}\bigg(4 + 2\mathbb{E}|Y^3| + \frac{2|\mathbb{E}Y^3|}{n} + \frac{\mathbb{E}|Y^5|}{n}\bigg) \\
&\quad + \frac{\pi|\mathbb{E}Y^3|}{\sqrt{mn}}(2 + \mathbb{E}|X^3|)\bigg(2 + \bigg(3 + \frac{\mathbb{E}Y^4}{n}\bigg)^{3/4}\bigg), \\
\gamma_3(X,Y) &= \frac{1}{24n}\bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg(4\mathbb{E}Y^4 + 6|\mathbb{E}Y^3|\bigg(4 + \mathbb{E}|Y^3| + \frac{4}{\sqrt{n}} + \frac{\mathbb{E}Y^4}{\sqrt{n}}\bigg)\bigg) \\
&\quad + \frac{\pi|\mathbb{E}X^3|}{m}(\mathbb{E}|X^3|+2)\bigg(8 + \bigg(15 + \frac{10(\mathbb{E}Y^3)^2}{n} + \frac{15\mathbb{E}Y^4}{n^2} + \frac{\mathbb{E}Y^6}{n^3}\bigg)^{5/6}\bigg) \\
&\quad + \frac{\pi|\mathbb{E}Y^3|}{4n}\bigg(2 + \bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg)\bigg(18 + \bigg(9 + \frac{8}{n}\bigg)\mathbb{E}Y^3| + \frac{4\mathbb{E}|Y^5|}{n}\bigg) \\
&\quad + \frac{\pi|\mathbb{E}X^3|}{\sqrt{mn}}(\mathbb{E}|Y^3|+2)\bigg(3 + 4\bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg) \\
&\quad + \frac{5\pi|\mathbb{E}Y^3|}{4\sqrt{mn}}\bigg(4 + 2\mathbb{E}|X^3| + \frac{2}{\sqrt{m}} + \frac{\mathbb{E}X^4}{\sqrt{m}}\bigg)\bigg(6 + \frac{\mathbb{E}Y^4}{n}\bigg).
\end{aligned}$$

Remark 4.5. Notice that the statistic $W = \frac{1}{\sqrt{mn}}\sum_{i,j=1}^{m,n} X_iY_j$ is symmetric in m and n and the random variables X_i and Y_j, and yet the bound (4.5) of Theorem 4.4 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 Section 5.3 we use zero bias couplings to arrive at bounds for W that are symmetric in m and n and the moments of X and Y – see Theorems 5.20 and 5.21.

In practice, when applying Theorem 4.4, we would compute $\gamma_k(X,Y)$ for $k=1,2,3$, and $\gamma_k(Y,X)$ for $k=1,2,3$, which would yields two bounds for the quantity $|\mathbb{E}h(W) - \mathrm{VG}^{1,0}_{1,0}h|$. We would then take the minimum of these two bounds. We proceed in this manner when applying bound (4.5) to prove Theorem 4.16.

Let us now prove Theorem 4.4.

Let $S = \frac{1}{\sqrt{m}}\sum_{i=1}^m X_i$ and $T = \frac{1}{\sqrt{n}}\sum_{j=1}^n Y_j$, so that $W = ST$. Let $S_i = S - \frac{1}{\sqrt{m}}X_i$ and $T_j = T - \frac{1}{\sqrt{n}}Y_j$, so that S_i is independent of X_i, and T_j is independent of Y_j. We therefore have the following formulas

$$\begin{aligned} W - S_iT &= ST - S_iT = \frac{1}{\sqrt{m}}X_iT \\ W - ST_j &= ST - ST_j = \frac{1}{\sqrt{n}}Y_jS. \end{aligned} \tag{4.6}$$

We will make repeated use of Taylor expansions, so let us denote $S_i^{[p]} = S_i + \theta_p(S - S_i)$ and $T_j^{[p]} = T_j + \theta_p(T - T_j)$, where $\theta_p \in (0,1)$.

Before beginning proper, we shall need a small lemma, which can be found in Pickett [55], Lemma 4.3:

Lemma 4.6. *Let $X, X_1, \dots, X_m$ be a collection of i.i.d. random variables with mean zero and unit variance. Define $S = \frac{1}{\sqrt{m}}\sum_{i=1}^m X_i$. Then, $\mathbb{E}S^p = O(1)$ for all $p \geq 1$. Specifically,*

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

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. Similar bounds hold for $\mathbb{E}T^p$.

We will also use the following lemma.

Lemma 4.7. *Suppose $p \geq 0$, then $\mathbb{E}|S_i|^p \leq \mathbb{E}|S|^p$.*

Proof. Applying Jensen's inequality gives

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

as required. □

Proof Part I: Expansions and Bounding

Using the VG(1, 0, 1, 0) Stein equation (3.14), we require a bound on the the expression $\mathbb{E}\{Wf''(W) + f'(W) - Wf(W)\}$. Due to the i.i.d. property 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

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

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

$$\mathbb{E}\{Wf''(W) + f'(W) - Wf(W)\} = R_1 + R_2 + R_3,$$

where

$$\begin{aligned}
R_1 &= -\frac{1}{2m^{3/2}}\sum_{i=1}^{m}\mathbb{E}X_i^3T^3f''(S_iT), \\
R_2 &= -\frac{1}{6m^2}\sum_{i=1}^{m}\mathbb{E}X_i^4T^4f^{(3)}(S_i^{[1]}T), \\
R_3 &= \mathbb{E}STf''(W) + \mathbb{E}f'(W) - \frac{1}{m}\sum_{i=1}^{m}\mathbb{E}X_i^2T^2f'(S_iT).
\end{aligned}$$

We begin by bounding R_1. Taylor expanding $f''(S_iT)$ about W and using (4.6) gives

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

where we used that $\mathbb{E}T^4 < 3 + \frac{\mathbb{E}Y^4}{n}$ and that $\mathbb{E}|X_i| \leq \sqrt{\mathbb{E}X_i^2} = 1$ to obtain the final equality. The term $\frac{1}{\sqrt{m}}|\mathbb{E}T^3 f''(W)|$ will be bounded to the desired order of $O(m^{-1}+n^{-1})$ by symmetry in the next section.

The bound for R_2 is immediate. We have

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

We now concentrate on bounding R_3. We use independence and that $\mathbb{E}X_i^2 = 1$ and then Taylor expand $f'(S_iT)$ about W to obtain

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

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

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

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

$$R_3 = \mathbb{E}STf''(W) + \mathbb{E}f'(W) - \mathbb{E}T^2 f'(W) + R_4,$$

where

$$|R_4| \leq \frac{1}{2m^2}\bigg|\sum_{i=1}^{m}\mathbb{E}X_i^2T^4f^{(3)}(S_i^{[3]}T)\bigg| + \frac{1}{m^2}\bigg|\sum_{i=1}^{m}\mathbb{E}X_i^2T^4f^{(3)}(S_i^{[4]}T)\bigg| \leq \frac{3\|f^{(3)}\|}{2m}\bigg(3+\frac{\mathbb{E}Y^4}{n}\bigg).$$

Noting that $T^2 = \frac{1}{\sqrt{n}}\sum_{j=1}^n Y_jT = \frac{1}{\sqrt{n}}\sum_{j=1}^n Y_j(\frac{1}{\sqrt{n}}Y_j + T_j)$, we may write R_3 as follows

$$R_3 = R_4 + R_5 + R_6,$$

where

$$\begin{aligned} R_5 &= \mathbb{E}f'(W) - \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}Y_j^2f'(W), \\ R_6 &= \mathbb{E}STf''(W) - \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\mathbb{E}Y_jT_jf'(W). \end{aligned}$$

We first consider R_5. Taylor expanding $f'(W)$ about ST_j and using that $ST - ST_j = \frac{1}{\sqrt{n}}Y_jS$ gives

$$\begin{aligned} R_5 &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(1-Y_j^2)f'(W) \\ &= \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(1-Y_j^2)\bigg(f'(ST_j) + \frac{1}{\sqrt{n}}Y_jSf''(ST_j) + \frac{1}{2n}Y_j^2S^2f^{(3)}(ST_j^{[5]})\bigg) \\ &= -\frac{\mathbb{E}Y^3}{n^{3/2}}\sum_{j=1}^{n}\mathbb{E}Sf''(ST_j) + \frac{1}{2n^2}\sum_{j=1}^{n}\mathbb{E}(Y_j^2 - Y_j^4)S^2f^{(3)}(ST_j^{[5]}), \end{aligned}$$

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_5

$$|R_5| \leq \frac{|\mathbb{E}Y^3|}{\sqrt{n}}|\mathbb{E}Sf''(W)| + \frac{\|f^{(3)}\|}{2n}(1 + 2|\mathbb{E}Y^3| + \mathbb{E}Y^4).$$

The term $\frac{1}{\sqrt{n}}|\mathbb{E}Sf''(W)|$ is bounded to the desired order of $O(m^{-1}+n^{-1})$ by symmetry in the next section.

We now consider R_6. 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

$$\begin{aligned} R_6 &= \mathbb{E}STf''(W) - \frac{1}{\sqrt{n}}\sum_{j=1}^n \mathbb{E}Y_jT_j\bigg(f'(ST_j) + \frac{1}{\sqrt{n}}Y_jSf''(ST_j) \\ &\quad + \frac{1}{2n}Y_j^2S^2f^{(3)}(ST_j) + \frac{1}{6n^{3/2}}Y_j^3S^3f^{(4)}(ST_j^{[7]})\bigg) \\ &= R_7 + R_8 + R_9, \end{aligned}$$

where

$$\begin{aligned} R_7 &= \frac{\mathbb{E}Y^3}{2n^{3/2}}\sum_{j=1}^n \mathbb{E}S^2T_jf^{(3)}(ST_j), \\ R_8 &= \frac{1}{6n^2}\sum_{j=1}^n \mathbb{E}Y_j^4S^3T_jf^{(4)}(ST_j^{[7]}), \\ R_9 &= \mathbb{E}STf''(W) - \frac{1}{n}\sum_{j=1}^n \mathbb{E}ST_jf''(ST_j). \end{aligned}$$

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

$$\begin{aligned} |R_7| &= \frac{|\mathbb{E}Y^3|}{2n^{3/2}}\bigg|\sum_{j=1}^n \mathbb{E}S^2Tf^{(3)}(ST_j)\bigg| \\ &= \frac{|\mathbb{E}Y^3|}{2n^{3/2}}\bigg|\sum_{j=1}^n \mathbb{E}\bigg(S^2Tf^{(3)}(W) - \frac{1}{\sqrt{n}}Y_jS^3\bigg(T_j + \frac{1}{\sqrt{n}}Y_j\bigg)f^{(4)}(ST_j^{[8]})\bigg)\bigg| \\ &\leq \frac{|\mathbb{E}Y^3|}{2\sqrt{n}}|\mathbb{E}S^2Tf^{(3)}(W)| + \frac{\|f^{(4)}\||\mathbb{E}Y^3|}{2n}\bigg(1 + \frac{1}{\sqrt{n}}\bigg)\bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}. \end{aligned}$$

The term $\frac{1}{\sqrt{n}}|\mathbb{E}S^2Tf^{(3)}(W)|$ is bounded to the desired order of $O(m^{-1} + n^{-1})$ by symmetry in the next section.

The bound for R_8 is immediate:

$$|R_8| \leq \frac{\|f^{(4)}\|}{6n}\mathbb{E}Y^4\bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}.$$

To bound R_9 we Taylor expand $f''(W)$ about ST_j and use independence and that the Y_j have zero mean to obtain

$$|R_9| = \frac{1}{n}\bigg|\sum_{j=1}^n \mathbb{E}ST[f''(ST) - f''(ST_j)]\bigg|$$

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

Before using symmetry arguments to bound the terms $\mathbb{E}Sf''(W)$, $\mathbb{E}S^2Tf^{(3)}(W)$ and $\mathbb{E}T^3f''(W)$ to the desired order, we obtain a useful bound for $\mathbb{E}S^2Tf^{(3)}(W)$ that will ensure that bound we seek for $|\mathbb{E}h(W)-\mathrm{VG}_{1,0}^{1,0}h|$ will only involve bounds of the first four derivatives of the $\mathrm{VG}(1,0,1,0)$ Stein equation (3.14), 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.

Lemma 4.8. *Let $f:\mathbb{R}\rightarrow\mathbb{R}$ be four times continuously differentiable, then*

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

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

$$
\begin{aligned}
\mathbb{E}ST^2f''(W) &= \frac{1}{\sqrt{n}}\sum_{j=1}^n \mathbb{E}Y_j\bigg(T_j+\frac{1}{\sqrt{n}}Y_j\bigg)Sf''(W) \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^n \mathbb{E}Y_j\bigg(T_j+\frac{1}{\sqrt{n}}Y_j\bigg)S\bigg(f''(ST_j) \\
&\quad +\frac{1}{\sqrt{n}}Y_jSf^{(3)}(ST_j)+\frac{1}{2n}Y_j^2S^2f^{(4)}(ST_j^{[10]})\bigg) \\
&= R_{10}+R_{11}+R_{12},
\end{aligned}
$$

where

$$
\begin{aligned}
R_{10} &= \frac{1}{n}\sum_{j=1}^n \mathbb{E}Sf''(ST_j), \\
R_{11} &= \frac{1}{n}\sum_{j=1}^n \mathbb{E}S^2\bigg(Y_j^2T_j+\frac{1}{\sqrt{n}}Y_j^3\bigg)f^{(3)}(ST_j), \\
R_{12} &= \frac{1}{2n^{3/2}}\sum_{j=1}^n \mathbb{E}S^3\bigg(Y_j^3T_j+\frac{1}{\sqrt{n}}Y_j^4\bigg)f^{(4)}(ST_j^{[10]}).
\end{aligned}
$$

Here we used that $\mathbb{E}Y_j = 0$ and $\mathbb{E}Y_j^2 = 1$ to simplify R_{10}. To bound R_{10} we Taylor expand $f''(W_j)$ about W to obtain

$$|R_{10}| \le |\mathbb{E}Sf''(W)| + \frac{\|f^{(3)}\|}{\sqrt{n}}.$$

The bound for R_{12} is immediate. We have

$$|R_{12}| \le \frac{\|f^{(4)}\|}{2\sqrt{n}}\left(\mathbb{E}|Y^3| + \frac{\mathbb{E}Y^4}{\sqrt{n}}\right)\left(3 + \frac{\mathbb{E}X^4}{m}\right)^{3/4}.$$

Finally, we bound R_{11}. By independence and that $\mathbb{E}Y_j = 0$ and $\mathbb{E}Y_j^2 = 1$, we have

$$R_{11} = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}S^2Tf^{(3)}(W_j) + \frac{\mathbb{E}Y^3}{n^{3/2}}\sum_{j=1}^{n}\mathbb{E}S^2f^{(3)}(ST_j).$$

Taylor expanding the $f^{(3)}(ST_j)$ about W allows us to write R_{11} as

$$R_{11} = \mathbb{E}S^2Tf^{(3)}(W) - \frac{1}{n^{3/2}}\sum_{j=1}^{n}\mathbb{E}Y_jS^3\left(T_j + \frac{1}{\sqrt{n}}Y_j\right)f^{(4)}(ST_j^{[11]}) + \frac{\mathbb{E}Y^3}{n^{3/2}}\sum_{j=1}^{n}\mathbb{E}S^2f^{(3)}(ST_j).$$

The final two terms of the above expression can be bounded by

$$\frac{\|f^{(3)}\|\mathbb{E}|Y^3|}{\sqrt{n}} + \frac{\|f^{(4)}\|}{\sqrt{n}}\left(1 + \frac{1}{\sqrt{n}}\right)\left(3 + \frac{\mathbb{E}X^4}{m}\right)^{3/4}.$$

Adding these terms gives the result. □

Proof Part II: Symmetry Argument for Optimal Rate

We now obtain a bound for $\frac{1}{\sqrt{n}}\mathbb{E}Sf''(W)$, $\frac{1}{\sqrt{n}}\mathbb{E}ST^2f''(W)$ and $\frac{1}{\sqrt{m}}\mathbb{E}T^3f''(W)$. To obtain the desired rate of convergence we shall need to use symmetry arguments, as were used in Section 4.1.2. of Pickett [55], to achieve the optimal rate of convergence for χ^2 limit theorems.

For large m and n we have $S \approx N(0,1)$ and $T \approx N(0,1)$, so we can apply the $O(n^{-1/2})$ bivariate central limit convergence rate (see, for example, Reinert and Röllin [61]) to the bivariate standard normal Stein equation with test functions $g_1(s,t) = sf''(st)$, $g_2(s,t) = st^2f''(st)$ and $g_3(s,t) = t^3f''(st)$. Intuitively, this will give a bound of the required order. The Stein equation with test function $g_k(s,t)$, $k = 1,2,3$, and solution ψ_k for a bivariate standard normal random variable is given by

$$\frac{\partial^2\psi_k}{\partial s^2}(s,t) + \frac{\partial^2\psi_k}{\partial t^2}(s,t) - s\frac{\partial\psi_k}{\partial s}(s,t) - t\frac{\partial\psi_k}{\partial t}(s,t) = g_k(s,t) - \mathbb{E}g_k(Z_1,Z_2), \quad (4.7)$$

where Z_1 and Z_2 are independent standard normal random variables. Due to the symmetry of the test function g_i, the following lemma holds, simplifying the right-hand side of (4.7).

Lemma 4.9. *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 $\mathbb{E}g_k(Z_1Z_2) = 0$, for $k = 1,2,3$.*

Proof. Let $Z_1' = -Z_1$, $Z_2' = -Z_2$, 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 $\mathbb{E}g(Z_1,Z_2) = 0$. $\square$

Now, providing that a solution ψ_k exists for the test function g_k, we have, performing Taylor expansions as in Chapter 2,

$$\begin{aligned}\mathbb{E}g_k(S,T) &= \mathbb{E}\left\{\frac{\partial^2\psi_k}{\partial s^2}(S,T) + \frac{\partial^2\psi_k}{\partial t^2}(S,T) - S\frac{\partial\psi_k}{\partial s}(S,T) - T\frac{\partial\psi_k}{\partial t}(S,T)\right\} \\ &= R_{13}^k + R_{14}^k + R_{15}^k + R_{16}^k,\end{aligned}$$

where

$$\begin{aligned}R_{13}^k &= \frac{1}{2m^{3/2}}\sum_{i=1}^m \mathbb{E}X_i^3\frac{\partial^3\psi_k}{\partial s^3}\left(S_i + \phi_1\frac{X_i}{\sqrt{m}}, T\right), \\ R_{14}^k &= \frac{1}{2n^{3/2}}\sum_{j=1}^n \mathbb{E}Y_j^3\frac{\partial^3\psi_k}{\partial t^3}\left(S, T_j + \phi_2\frac{Y_j}{\sqrt{n}}\right), \\ R_{15}^k &= \frac{1}{m^{3/2}}\sum_{i=1}^m \mathbb{E}X_i\frac{\partial^3\psi_k}{\partial s^3}\left(S_i + \phi_3\frac{X_i}{\sqrt{m}}, T\right), \\ R_{16}^k &= \frac{1}{n^{3/2}}\sum_{j=1}^n \mathbb{E}Y_j\frac{\partial^3\psi_k}{\partial t^3}\left(S, T_j + \phi_4\frac{Y_j}{\sqrt{n}}\right),\end{aligned}$$

with ϕ_1, ϕ_2, ϕ_3, $\phi_4 \in (0,1)$.

Before we bound the remainder terms, we need an expression for the third order partial derivatives of the solution ψ_k in terms of the test functions ψ_k. We achieve this task by using the following lemma.

Lemma 4.10. *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be twice differentiable, then the third order partial derivatives of the solution ψ to the standard bivariate normal Stein equation (4.7) are given by*

$$\begin{aligned}\frac{\partial^3\psi}{\partial s^3} &= \int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-u}}{\sqrt{1-\mathrm{e}^{-2u}}}\frac{\partial^2}{\partial s^2}g(z_s,z_t)\phi'(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u, \qquad (4.8)\\ \frac{\partial^3\psi}{\partial t^3} &= \int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-u}}{\sqrt{1-\mathrm{e}^{-2u}}}\frac{\partial^2}{\partial t^2}g(z_s,z_t)\phi(x)\phi'(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u,\end{aligned}$$

where

$$
\begin{aligned}
z_s &= \mathrm{e}^{-u}s + \sqrt{1-\mathrm{e}^{-2u}}x, && (4.9)\\
z_t &= \mathrm{e}^{-u}t + \sqrt{1-\mathrm{e}^{-2u}}y, && (4.10)
\end{aligned}
$$

and

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}x^2}, \qquad \phi'(x) = -x\phi(x).$$

Proof. This is a straightforward generalisation of the proof of Lemma 3.2 of Raic [57]. □

The following lemmas allow us to bound the third order partial derivatives of the ψ_k in terms of the derivatives of f. Before stating the lemma, we define the double factorial function. The double factorial of a positive integer n is given by

$$n!! = \begin{cases} 1\cdot 3\cdot 5\cdot\dots\cdot(n-2)\cdot n, & n>0 \text{ odd},\\ 2\cdot 4\cdot 6\cdots(n-2)\cdot n, & n>0 \text{ even},\end{cases} \tag{4.11}$$

and we define $(-1)!! = 0!! = 1$ (Arfken [3], p.547).

Lemma 4.11. *Suppose that $f : \mathbb{R}^2 \to \mathbb{R}$ is four times differentiable, and let $g(s,t) = z_s^a z_t^b f''(z_s z_t)$, where $a, b \in \mathbb{N}$, and z_s, z_t are defined as in equations (4.9) and (4.10), respectively. Then, the third order partial derivatives of the solution ψ to the standard bivariate normal Stein equation (4.7) with test function g are bounded as follows*

$$
\begin{aligned}
\left|\frac{\partial^3\psi}{\partial s^3}\right| &\le \frac{\pi}{4}\{2^{a+b}\|f^{(4)}\|(|s|^a + a!!)(|t|^{b+2}+(b+1)!!)\\
&\quad +2a2^{a+b-2}\|f^{(3)}\|(|s|^{a-1}+(a-1)!!)(|t|^{b+1}+b!!)\\
&\quad +a(a-1)2^{a+b-4}\|f''\|(|s|^{a-2}+(a-2)!!)(|t|^b+(b-1)!!)\}, \qquad (4.12)\\
\left|\frac{\partial^3\psi}{\partial t^3}\right| &\le \frac{\pi}{4}\{2^{a+b}\|f^{(4)}\|(|s|^{a+2} + (a+1)!!)(|t|^{b}+b!!)\\
&\quad +2b2^{a+b-2}\|f^{(3)}\|(|s|^{a+1}+a!!)(|t|^{b-1}+(b-1)!!)\\
&\quad +b(b-1)2^{a+b-4}\|f''\|(|s|^{a-2}+(a-1)!!)(|t|^{b-2}+(b-2)!!)\},
\end{aligned}
$$

where $n!!$ denotes the double factorial of $n \in \mathbb{N}$.

Proof. We prove that the first bound holds; the second bound then follows by symmetry. We begin by calculating the second order partial derivative of g with respect to s. Since

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

we have that

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

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

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

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

$$\begin{aligned}\left|\frac{\partial^2 g}{\partial s^2}\right| \leq \mathrm{e}^{-2u}\{&2^{a+b}\|f^{(4)}\|(|s|^a + |x|^a)(|t|^{b+2} + |y|^{b+2}) \\ &+ 2a2^{a+b-2}\|f^{(3)}\|(|s|^{a-1} + |x|^{a-1})(|t|^{b+1} + |y|^{b+1}) \\ &+ a(a-1)2^{a+b-4}\|f''\|(|s|^{a-2} + |x|^{a-2})(|t|^b + |y|^b)\}.\end{aligned}$$

Applying this bound to equation (4.8) gives the following bound on the third order partial derivative of ψ with respect to s:

$$\begin{aligned}\left|\frac{\partial^3 \psi}{\partial s^3}\right| \leq & \int_0^\infty \int_{\mathbb{R}^2} \frac{\mathrm{e}^{-u}}{\sqrt{1-\mathrm{e}^{-2u}}} \left|\frac{\partial^2 g}{\partial s^2}\right| |x|\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u \\ \leq & \int_0^\infty \int_{\mathbb{R}^2} \frac{\mathrm{e}^{-3u}}{\sqrt{1-\mathrm{e}^{-2u}}}\{2^{a+b}\|f^{(4)}\|(|s|^a + |x|^a)(|t|^{b+2} + |y|^{b+2}) \\ & + 2a2^{a+b-2}\|f^{(3)}\|(|s|^{a-1} + |x|^{a-1})(|t|^{b+1} + |y|^{b+1}) \\ & + a(a-1)2^{a+b-4}\|f''\|(|s|^{a-2} + |x|^{a-2})(|t|^b + |y|^b)\}|x|\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}u \\ = & \frac{\pi}{4}\int_{\mathbb{R}^2}\{2^{a+b}\|f^{(4)}\|(|s|^a + |x|^a)(|t|^{b+2} + |y|^{b+2}) \\ & + 2a2^{a+b-2}\|f^{(3)}\|(|s|^{a-1} + |x|^{a-1})(|t|^{b+1} + |y|^{b+1}) \\ & + a(a-1)2^{a+b-4}\|f''\|(|s|^{a-2} + |x|^{a-2})(|t|^b + |y|^b)\}|x|\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y, \end{aligned} \tag{4.13}$$

where the final equality follows from the formula $\int_0^\infty \frac{\mathrm{e}^{-3u}}{\sqrt{1-\mathrm{e}^{-2u}}}\,\mathrm{d}u = \frac{\pi}{4}$ (see (2.21)). We can now obtain the bound (4.8) by using the following formula (see formula 17 of Winkelbauer [75]) to evaluate (4.13):

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

which completes the proof. □

With these bounds it is easy to bound the remainder terms. The following lemma allows us to easily deduce bounds for the remaining remainder terms.

Lemma 4.12. *Suppose that $f : \mathbb{R}^2 \to \mathbb{R}$ is four times differentiable, and let $g(s,t) = s^a t^b f''(st)$, where $a, b \in \mathbb{N}$, then*

$$\begin{aligned}
|R_{13}^k| \leq\ & \frac{\pi}{8\sqrt{m}}\bigg\{2^{a+b}\|f^{(4)}\|\bigg(2^{a-1}\bigg(\mathbb{E}|X|^3\mathbb{E}|S|^a + \frac{\mathbb{E}|X|^{a+3}}{m^{a/2}}\bigg) + a!!\,\mathbb{E}|X|^3\bigg) \\
&\times(\mathbb{E}|T|^{b+2} + (b+1)!!) \\
&+2a2^{a+b-2}\|f^{(3)}\|\bigg(2^{a-2}\bigg(\mathbb{E}|X|^3\mathbb{E}|S|^{a-1} + \frac{\mathbb{E}|X|^{a+2}}{m^{(a-1)/2}}\bigg) + (a-1)!!\,\mathbb{E}|X|^3\bigg) \\
&\times(\mathbb{E}|T|^{b+1} + b!!) \\
&+a(a-1)2^{a+b-4}\|f''\|\bigg(2^{a-3}\bigg(\mathbb{E}|X|^3\mathbb{E}|S|^{a-2} + \frac{\mathbb{E}|X|^{a+1}}{m^{(a-2)/2}}\bigg) + (a-2)!!\,\mathbb{E}|X|^3\bigg) \\
&\times(\mathbb{E}|T|^{b} + (b-1)!!)\bigg\}, \\
|R_{14}^k| \leq\ & \frac{\pi}{8\sqrt{n}}\bigg\{2^{a+b}\|f^{(4)}\|\bigg(2^{b-1}\bigg(\mathbb{E}|Y|^3\mathbb{E}|T|^b + \frac{\mathbb{E}|Y|^{b+3}}{n^{b/2}}\bigg) + b!!\,\mathbb{E}|Y|^3\bigg) \\
&\times(\mathbb{E}|S|^{a+2} + (a+1)!!) \\
&+2b2^{a+b-2}\|f^{(3)}\|\bigg(2^{b-2}\bigg(\mathbb{E}|Y|^3\mathbb{E}|T|^{b-1} + \frac{\mathbb{E}|Y|^{b+2}}{n^{(b-1)/2}}\bigg) + (b-1)!!\,\mathbb{E}|Y|^3\bigg) \\
&\times(\mathbb{E}|S|^{a+1} + a!!) \\
&+b(b-1)2^{a+b-4}\|f''\|\bigg(2^{b-3}\bigg(\mathbb{E}|Y|^3\mathbb{E}|T|^{b-2} + \frac{\mathbb{E}|Y|^{b+1}}{n^{(b-2)/2}}\bigg) + (b-2)!!\,\mathbb{E}|Y|^3\bigg) \\
&\times(\mathbb{E}|S|^{a} + (a-1)!!)\bigg\}.
\end{aligned}$$

The bound for R_{15}^k is similar to the bound for R_{13}^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_{16}^k is similar to the bound for R_{14}^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_{13}^k holds; the bound for R_{14}^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 $p \geq 1$:

$$|S_i^*|^p = \left|S_i + \frac{\phi_1}{\sqrt{m}}X_i\right|^p \leq 2^{p-1}\bigg(|S_i|^p + \frac{\phi_1^p}{m^{p/2}}|X_i|^p\bigg) \leq 2^{p-1}\bigg(|S_i|^p + \frac{|X_i|^p}{m^{p/2}}\bigg). \tag{4.14}$$

Using our bound (4.12) for the third order partial derivative of ψ with respect to s, we have

$$\begin{aligned}
|R_{13}^k| &= \frac{1}{2m^{3/2}}\bigg|\sum_{i=1}^m \mathbb{E}X_i^3 \frac{\partial^3\psi}{\partial s^3}(S_i^*, T)\bigg| \\
&\leq \frac{\pi}{8m^{3/2}}\sum_{i=1}^m \mathbb{E}\bigg|X_i^3\Big\{2^{a+b}\|f^{(4)}\|(|S_i^*|^a + a!\,!)(|T|^{b+2} + (b+1)!\,!) \\
&\quad + 2a2^{a+b-2}\|f^{(3)}\|(|S_i^*|^{a-1} + (a-1)!\,!)(|T|^{b+1} + b!\,!) \\
&\quad + a(a-1)2^{a+b-4}\|f''\|(|S_i^*|^{a-2} + (a-2)!\,!)(|T|^b + (b-1)!\,!)\}| \\
&\leq \frac{\pi}{8m^{3/2}}\sum_{i=1}^m \mathbb{E}\bigg|X_i^3\bigg\{2^{a+b}\|f^{(4)}\|\bigg(2^{a-1}\bigg(|S_i|^a + \frac{|X_i|^a}{m^{a/2}}\bigg) + a!\,!\bigg)(|T|^{b+2} + (b+1)!\,!) \\
&\quad + 2a2^{a+b-2}\|f^{(3)}\|\bigg(2^{a-2}\bigg(|S_i|^{a-1} + \frac{|X_i|^{a-1}}{m^{(a-1)/2}}\bigg) + a!\,!\bigg)(|T|^{b+1} + b!\,!) \\
&\quad + a(a-1)2^{a+b-4}\|f''\|\bigg(2^{a-3}\bigg(|S_i|^{a-2} + \frac{|X_i|^{a-2}}{m^{(a-2)/2}}\bigg) + a!\,!\bigg)(|T|^b + (b-1)!\,!)\bigg\}\bigg|,
\end{aligned}$$

where we used (4.14) to obtain the final inequality. Applying the triangle inequality, that X_i and S_i are independent and that, by Lemma 4.7, $\mathbb{E}|S_i|^p \leq \mathbb{E}|S|^p$ gives the desired bound. The final statement of the Lemma is clear. □

We can bound R_{13}^k, R_{14}^k, R_{15}^k and R_{16}^k by using the bounds in Lemma 4.12. We illustrate the argument by bounding R_{13}^1. In this case we have $g_1(s,t) = sf''(st)$, that is $a = 1$ and $b = 0$. We have

$$\begin{aligned}
|R_{13}^1| &\leq \frac{\pi}{8\sqrt{m}}\bigg\{2\|f^{(4)}\|\bigg(2|X^3| + \frac{\mathbb{E}X^4}{\sqrt{m}}\bigg)(\mathbb{E}T^2 + 1!\,!) + \|f^{(3)}\|(2\mathbb{E}|X^3|)(\mathbb{E}|T| + 0!\,!)\bigg\} \\
&= \frac{\pi}{2\sqrt{m}}\bigg\{\|f^{(4)}\|\bigg(2\mathbb{E}|X^3| + \frac{\mathbb{E}X^4}{\sqrt{m}}\bigg) + \|f^{(3)}\|\mathbb{E}|X^3|\bigg\},
\end{aligned}$$

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 bounds.

$$\begin{aligned}
|R_{14}^1| &\leq \frac{\pi\|f^{(4)}\|}{2\sqrt{n}}\mathbb{E}|Y^3|\bigg(2 + \bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg), \\
|R_{13}^2| &\leq \frac{\pi}{\sqrt{m}}\bigg\{\|f^{(4)}\|\bigg(2\mathbb{E}|X^3| + \frac{\mathbb{E}X^4}{\sqrt{m}}\bigg)\bigg(6 + \frac{\mathbb{E}Y^4}{n}\bigg) \\
&\quad + \|f^{(3)}\|\mathbb{E}|X^3|\bigg(2 + \bigg(3 + \frac{\mathbb{E}Y^4}{n}\bigg)^{3/4}\bigg)\bigg\},
\end{aligned}$$

$$
\begin{aligned}
|R_{14}^2| &\leq \frac{2\pi}{\sqrt{n}}\bigg\{\|f^{(4)}\|\bigg(2\mathbb{E}|Y^3|+\frac{\mathbb{E}|Y^5|}{n}\bigg)\bigg(2+\bigg(3+\frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg) \\
&\quad +\|f^{(3)}\|\bigg(2\mathbb{E}|Y^3|+\frac{\mathbb{E}Y^4}{\sqrt{n}}\bigg)+\|f''\|\mathbb{E}|Y^3|\bigg\}, \\
|R_{13}^3| &\leq \frac{2\pi\|f^{(4)}\|}{\sqrt{m}}\mathbb{E}|X^3|\bigg(8+\bigg(15+\frac{10(\mathbb{E}Y^3)^2}{n}+\frac{15\mathbb{E}Y^4}{n^2}+\frac{\mathbb{E}Y^6}{n^3}\bigg)^{5/6}\bigg), \\
|R_{14}^3| &\leq \frac{\pi}{4\sqrt{n}}\bigg\{8\|f^{(4)}\|\bigg(\mathbb{E}|Y^3|\bigg(3+4\bigg(3+\frac{\mathbb{E}Y^4}{m}\bigg)^{3/4}\bigg)+\frac{\mathbb{E}Y^6}{m^{3/2}}\bigg) \\
&\quad +24\|f^{(3)}\|\bigg(2\mathbb{E}|Y^3|+\frac{\mathbb{E}|Y^5|}{n}\bigg)+3\|f''\|\bigg(2\mathbb{E}|Y^3|+\frac{\mathbb{E}Y^4}{\sqrt{n}}\bigg)\bigg\}.
\end{aligned}
$$

The bound for R_{15}^k is similar to the bound for R_{13}^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_{16}^k is similar to the bound for R_{14}^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 can now sum up the remainder terms to obtain a bound for $|\mathbb{E}h(W)-\mathrm{VG}_{1,0}^{1,0}h|$, thus completing the proof of Theorem 4.4.

4.3 Extension to the Multi-Dimensional Case

For the case of $r>1$, we have the following multivariate generalisation of Theorem 4.4:

Theorem 4.13. *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*

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

where the $M_{r,1}^i(h)$ are defined as in Theorem 3.21, $\mathrm{VG}_{1,0}^{r,0}h$ denotes the expectation of $h(Z)$ for $Z\sim\mathrm{VG}(r,0,1,0)$, and the γ_i are as in Theorem 4.4.

Proof. We begin by introducing some notation. We define $W^{(k)}=\frac{1}{\sqrt{mn}}\sum_{i,j=1}^{m,n}X_{ik}Y_{jk}$. With this notation we have $W_r=\sum_{k=1}^r W^{(k)}$. Using the $\mathrm{VG}(r,0,1,0)$ Stein equation (3.14) we have

$$
\begin{aligned}
|\mathbb{E}h(W_r)-\mathrm{VG}_{1,0}^{r,0}h| &= |\mathbb{E}\{W_rf''(W_r)+rf'(W_r)-W_rf(W_r)\}| \\
&= \bigg|\sum_{k=1}^r\mathbb{E}\{W^{(k)}f''(W_r)+f'(W_r)-W^{(k)}f(W_r)\}\bigg| \\
&= \bigg|\sum_{k=1}^r\mathbb{E}\Big[\mathbb{E}(W^{(k)}f''(W_r)+f'(W_r)-W^{(k)}f(W_r)\mid W^{(l)},\, l\neq k)\Big]\bigg|.
\end{aligned}
$$

Letting $w_{(k)}$ denote the set $\{w_l\in\mathbb{R}\ :\ l=1,\dots,k-1,k+1,\dots,r\}$, and defining $g_k(x)=$

$f(x+\sum_{l\neq k}^{r} w_l)$, where $\sum_{l\neq k}^{r} w_l$ represents the sum of the variables that we have conditioned on, we may bound the conditional expectation in the above expression to obtain

$$|\mathbb{E}h(W_r) - \mathrm{VG}^{r,0}_{1,0}h| \leq \sum_{k=1}^{r} \sup_{w_{(k)}} \mathbb{E}|W^{(k)}g_k''(W^{(k)}) + g_k'(W^{(k)}) - W^{(k)}g_k(W^{(k)})|.$$

Since $\|g_k^{(n)}\| = \|f^{(n)}\|$ for all n and k, we may use our bound (4.5) from Theorem 4.4 to bound the above the expression, thus completing the proof. □

Remark 4.14. The terms $M^k_{r,1}(h)$, for $k = 2, 3, 4$, are of order $r^{-1/2}$ as $r \to \infty$ (recall Theorem 3.21), and therefore the bound of Theorem 4.13 is of order $r^{1/2}(m^{-1} + n^{-1})$. This in agreement with bound (4.3) of Pickett [55] for χ^2 approximation, which is of order $d^{1/2}m^{-1}$.

Remark 4.15. 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-Esséen theorem). Consider the following example in the case of a $\mathrm{VG}(1, 0, 1, 0)$ random variable with test function $h \equiv \chi_{\{0\}}$. Let X_i, $i = 1, \ldots, m = 2k$ and Y_j, $j = 1, \ldots, n = 2l$, be random variables taking values in the set $\{-1, 1\}$ with equal probability. Then $\mathbb{E}X_i = \mathbb{E}Y_j = 0$, $\mathrm{Var}X_i = \mathrm{Var}Y_j = 1$ and

$$\begin{aligned}\mathbb{E}h(W) &= \mathbb{P}\Big(\sum\nolimits_{i,j} X_iY_j = 0\Big) \leq \mathbb{P}\Big(\sum\nolimits_{i} X_i = 0\Big) + \mathbb{P}\Big(\sum\nolimits_{j} Y_j = 0\Big) \\ &= \binom{2k}{k}\left(\frac{1}{2}\right)^{2k} + \binom{2l}{l}\left(\frac{1}{2}\right)^{2l} \approx \frac{1}{\sqrt{\pi k}} + \frac{1}{\sqrt{\pi l}} = \sqrt{\frac{2}{\pi m}} + \sqrt{\frac{2}{\pi n}},\end{aligned}$$

by Stirling's approximation. Furthermore, $\mathrm{VG}^{1,0}_{1,0}h = \mathbb{P}(\mathrm{VG}(1, 0, 1, 0) = 0) = 0$, and hence the univariate bound (4.5) fails.

4.4 Application: Binary Sequence Comparison

We now consider the word sequence comparison problem, which was described in the Introduction, in the case of an alphabet of size 2 with comparison based on the content of 1-tuples. We consider a limit theorem for a statistic that is closely related to D_2^*, the D_2 statistic:

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

where, as in Section 1.1, $X_{\mathbf{w}}$ and $Y_{\mathbf{w}}$ denote the occurrence of word $\mathbf{w}$ in the the first and second sequences, respectively. Lippert et al. [43] studied the D_2 statistic and showed that for certain parameter values D_2 is approximately normally or Poisson distributed.

Using Theorem 4.4 we are able to obtain explicit bounds for the error in approximating D_2 by its limiting distribution for the case of an alphabet of size 2 with comparison based one 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

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

We now consider the standardised D_2 statistic,

$$\begin{aligned} W &= \frac{D_2 - \mathbb{E}D_2}{\sqrt{\mathrm{Var}D_2}} \\ &= \frac{2}{\sqrt{mn}}\left(XY + (m-X)(n-Y) - \frac{mn}{2}\right) \\ &= \frac{2}{\sqrt{mn}}\left(2XY - mX - nY + \frac{mn}{2}\right) \\ &= \left(\frac{X-\frac{m}{2}}{\sqrt{\frac{m}{4}}}\right)\left(\frac{Y-\frac{n}{2}}{\sqrt{\frac{n}{4}}}\right). \end{aligned} \tag{4.15}$$

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 $\mathrm{VG}(1,0,1,0)$ distribution. We now apply Theorem 4.4 to obtain a bound on the error, in a weak convergence setting, in approximating the standardised D_2 statistic by its limiting $\mathrm{VG}(1,0,1,0)$ distribution.

Theorem 4.16. *For uniform i.i.d. binary sequences of lengths m and n, the standardised D_2 statistic W, defined as in equation (4.15), based on 1-tuple content is approximately $\mathrm{VG}(1,0,1,0)$ distributed as $m \to \infty$ and $n \to \infty$ simultaneously. Moreover, for $h \in C_b^3(\mathbb{R})$ the following bound on the error in approximating the distribution of W by a $\mathrm{VG}(1,0,1,0)$ distribution,*

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

where

$$A = \frac{5}{3m}\left(3 + \frac{1}{n}\right)N_1^3(h) + \frac{1}{6n}\left(3 + \frac{1}{m}\right)^{3/4} N_1^4(h),$$

$$B = \frac{5}{3n}\left(3+\frac{1}{m}\right)N_1^3(h) + \frac{1}{6m}\left(3+\frac{1}{n}\right)^{3/4} N_1^4(h),$$

and the $N_1^k(h)$ *is defined as in Theorem 3.22, and* $\mathrm{VG}^{1,0}_{1,0}h$ *denotes the expectation of* $h(Z)$*, for* $Z \sim \mathrm{VG}(1,0,1,0)$*.*

Proof. Let $\mathbb{I}_i$ be the indicator random variable that letter 0 appears at position i in the first sequence. Similarly we let $\mathbb{J}_j$ be the indicator random variable that letter 0 appears at position j in the second sequence. The $\mathbb{I}_i$ and $\mathbb{J}_j$ are independent Bernoulli random variables with mean $\frac{1}{2}$. Then the number of occurrences of letter 0 in the first and second sequences X and Y, respectively, can be written in term of the indicators $\mathbb{I}_i$, $\mathbb{J}_j$ as $X = \sum_{i=1}^m \mathbb{I}_i$ and $Y = \sum_{j=1}^n \mathbb{J}_j$. Therefore we may write the standardised D_2 statistic W as follows

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

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 directly apply Theorem 4.4 to achieve a bound on $|\mathbb{E}h(W) - \mathrm{VG}^{1,0}_{1,0}h|$. Applying bound (4.5) with $\mathbb{E}X_i^3 = \mathbb{E}(2\mathbb{I}_i - 1)^3 = 0$, $\mathbb{E}Y_j^3 = \mathbb{E}(2\mathbb{J}_j - 1)^3 = 0$, $\mathbb{E}X_i^4 = \mathbb{E}(2\mathbb{I}_i - 1)^4 = 1$ and $\mathbb{E}Y_j^4 = \mathbb{E}(2\mathbb{J}_j - 1)^4 = 1$ yields bound (4.16). $\square$

We now summarise the results of Sections 4.2, 4.3 and 4.4, in the context of the word sequence comparison problem that we considered in the Introduction. We have used Stein's method for Variance-Gamma distributions to study the asymptotic behaviour of the statistic (4.2), which is of a similar form to D_2^* but with the simpler assumption of i.i.d. random variables. In Theorem 4.16 we applied the bounds of Theorem 4.4 to study the limiting behaviour of a statistic that is closely related to D_2^*, for the case of an alphabet of size 2 with comparison based on 1-tuples. We have therefore made substantial progress towards to goal of using Stein's method to study the asymptotic behaviour of D_2^*. With further research we would hope to obtain bounds for the case of larger alphabet sizes and comparisons based on k-tuples with length greater than 1.

4.5 A Variance-Gamma limit theorem for random variables with a local dependence structure

In this section, we use the $\mathrm{VG}(1,0,1,0)$ Stein equation (3.13), together with Taylor expansions and local approach couplings, to prove a limit theorem for the statistic $\sum_{i,j=1}^{m,n} X_i Y_j$, where the X_i have zero mean and are locally dependent, the Y_j have zero mean and are also locally dependent, but the collections $\{X_1, \dots X_m\}$ and $\{Y_1, \dots, Y_n\}$ are independent of each other. Our calculations demonstrate that the generalisation from i.i.d. to locally dependent random variables is straightforward, as is the case for the normal distribution (see Example 2.8).

Before stating the theorem, we introduce some notation and precisely specify the dependence structure of the X_i and Y_j. Suppose for each $i = 1, \dots, n$, there exist index sets $\{i\} \subseteq A_i^{(1)} \subseteq A_i^{(2)} \subseteq \{1, \dots, m\}$ such that $X_i \perp\!\!\!\perp \sigma\{X_j : j \in A_i^{(1)}\}$, and if $X_p \in \sigma\{X_j : j \in A_i^{(1)}\}$ then $X_p \perp\!\!\!\perp \sigma\{X_j : j \in A_i^{(2)}\}$, where $\sigma\{X_j\}$ denotes the σ–algebra generated by X_j. Similarly, we suppose that there exist index sets $\{j\} \subseteq B_j^{(1)} \subseteq B_j^{(2)} \subseteq B_j^{(3)} \subseteq \{1, \dots, n\}$ such that $Y_j \perp\!\!\!\perp \sigma\{Y_k : k \in B_j^{(1)}\}$, if $Y_p \in \sigma\{Y_k : k \in B_j^{(1)}\}$ then $Y_p \perp\!\!\!\perp \sigma\{Y_k : k \in B_j^{(2)}\}$, and if $Y_q \in \sigma\{Y_k : k \in B_j^{(2)}\}$ then $Y_q \perp\!\!\!\perp \sigma\{Y_k : k \in B_j^{(3)}\}$. Note that the sequences of random variables $X_1, \dots, X_m$ and $Y_1, \dots, Y_n$ have the same dependence structure as the sequence of random variables we considered in Example 2.8.

Theorem 4.17. *Let $X_1, X_2, \dots, X_m$ be a collection of mean zero random variables with $\mathbb{E}|X_i|^3 < \infty$, for $1 \leq i \leq m$, that have a dependence structure as outlined above, and are normalized so that $\mathbb{E}S^2 = 1$, where $S = \sum_{i=1}^m X_i$. Similarly, let $Y_1, Y_2, \dots, Y_n$ be a collection of mean zero random variables with $\mathbb{E}Y_j^4 < \infty$, for $1 \leq j \leq n$, that have a dependence structure as outlined above, and are normalized so that $\mathbb{E}T^2 = 1$, where $T = \sum_{j=1}^n Y_j$. Suppose further that the σ-fields $\sigma\{X_i : i = 1, \dots, m\}$ and $\sigma\{Y_j : j = 1, \dots, n\}$ are independent. For $k = 1, 2$ we define $X_i^{(k)} = \sum_{u \in A_i^{(k)}} X_u$ and for $l = 1, 2, 3$ we define $Y_j^{(l)} = \sum_{v \in B_j^{(l)}} Y_v$. Let $W = ST$. Then for $h \in C_b^2(\mathbb{R})$, we have*

$$|Eh(W) - \mathrm{VG}^{1,0}_{1,0}h| \leq A_1 N_1^2(h) + A_2 N_1^3(h),$$

where the $N_1^k(h)$ are defined as in Theorem 3.22,

$$\begin{aligned} A_1 &= \mathbb{E}|T|^3 \sum_{i=1}^m \Big\{ \frac{1}{2}\mathbb{E}\Big|X_i (X_i^{(1)})^2\Big| + \mathbb{E}\Big|X_i X_i^{(1)}(X_i^{(2)} - X_i^{(1)})\Big| + \Big|\mathbb{E}X_i X_i^{(1)}\Big|\mathbb{E}\Big|X_i^{(2)}\Big|\Big\} \\ &\quad + \sum_{j=1}^n \Big\{ \mathbb{E}\Big|Y_j Y_j^{(1)}\Big| + 2\Big|\mathbb{E}Y_j Y_j^{(1)}\Big|\mathbb{E}\Big|Y_j^{(2)}\Big| + \mathbb{E}\Big|Y_j Y_j^{(1)}(Y_j^{(2)} - Y_j^{(1)})\Big|\Big\}, \\ A_2 &= \sum_{j=1}^n \Big\{ \Big[1 + \mathbb{E}\Big|Y_j^{(2)}\Big|\Big]\mathbb{E}Y_j (Y_j^{(1)})^2\Big| + \Big|\mathbb{E}Y_j Y_j^{(1)}\Big|\mathbb{E}\Big|Y_j^{(2)}\Big| \\ &\quad + \Big[2 + \mathbb{E}\Big|Y_j^{(2)}\Big|\Big]\mathbb{E}\Big|Y_j Y_j^{(1)}(Y_j^{(2)} - Y_j^{(1)})\Big| + \mathbb{E}\Big|Y_j Y_j^{(1)}(Y_j^{(2)} - Y_j^{(1)})(Y_j^{(3)} - Y_j^{(2)})\Big|\Big\}, \end{aligned}$$

and

$$\mathbb{E}|T|^3 \leq \left(\sum_{j=1}^n \mathbb{E}Y_j Y_j^{(1)} Y_j^{(2)} Y_j^{(3)} + 3 \sum_{j=1}^n \sum_{k \notin B_j^{(2)}} \mathbb{E}Y_j Y_j^{(1)} \mathbb{E}Y_k Y_k^{(1)} \right)^{3/4}.$$

Proof. Using the VG$(1,0,1,0)$ Stein equation (3.14), we require a bound on the the expression $\mathbb{E}\{Wf''(W)+f'(W)-Wf(W)\}$. We will arrive at such a bound by using Taylor expansions in conjunction with local approach couplings. For $k=1,2$ we define $S_i^{(k)}=S-X_i^{(k)}$ and for $l=1,2,3$ we define $T_j^{(l)}=T-Y_j^{(l)}$. Throughout the proof we will make use of the fact that $X_i \perp\!\!\!\perp S_i^{(1)}$ and $X_i^{(1)} \perp\!\!\!\perp S_i^{(2)}$, with similar relations holding for the Y_j.

We Taylor expand $f(W)$ about $S_i^{(1)}T$, and use that $W-S_i^{(1)}T=X_i^{(1)}T$, to obtain

$$\begin{aligned}\mathbb{E}Wf(W)&=\mathbb{E}STf(ST)\\&=\sum_{i=1}^m\mathbb{E}X_iTf(ST)\\&=\sum_{i=1}^m\mathbb{E}X_iTf(S_i^{(1)}T)+\sum_{i=1}^m\mathbb{E}X_iX_i^{(1)}T^2f'(S_i^{(1)}T)+R_1,\end{aligned}$$

where

$$|R_1|\le\frac{1}{2}\|f''\|\sum_{i=1}^m\mathbb{E}\left|X_i(X_i^{(1)})^2T^3\right|=\frac{1}{2}\|f''\|\mathbb{E}|T|^3\sum_{i=1}^m\mathbb{E}\left|X_i(X_i^{(1)})^2\right|.$$

We bound $\mathbb{E}|T|^3$ in terms of the $Y_j^{(k)}$ by using Hölder's inequality, $\mathbb{E}|T|^3\le\{\mathbb{E}T^4\}^{3/4}$, and a straightforward calculation shows that

$$\{\mathbb{E}T^4\}^{3/4}=\left(\sum_{j=1}^n\mathbb{E}Y_jY_j^{(1)}Y_j^{(2)}Y_j^{(3)}+3\sum_{j=1}^n\sum_{k\notin B_j^{(2)}}\mathbb{E}Y_jY_j^{(1)}\mathbb{E}Y_kY_k^{(1)}\right)^{3/4}.$$

We now use independence and that the X_i have zero mean to obtain

$$\begin{aligned}\mathbb{E}Wf(W)&=\sum_{i=1}^m\mathbb{E}X_i\mathbb{E}Tf(S_i^{(1)}T)+\sum_{i=1}^m\mathbb{E}X_iX_i^{(1)}T^2f'(S_i^{(1)}T)+R_1\\&=\sum_{i=1}^m\mathbb{E}X_iX_i^{(1)}T^2f'(S_i^{(1)}T)+R_1.\end{aligned}$$

Taylor expanding $f'(S_i^{(1)}T)$ about $S_i^{(2)}T$, and using independence and that $S_i^{(1)}T-S_i^{(2)}T=(X_i^{(2)}-X_i^{(1)})T$, gives

$$\mathbb{E}Wf(W)=\sum_{i=1}^m\mathbb{E}X_iX_i^{(1)}\mathbb{E}T^2f'(S_i^{(2)}T)+R_1+R_2,$$

where

$$|R_2| \leq \|f''\| \sum_{i=1}^{m} \mathbb{E}\Big|X_i X_i^{(1)}(X_i^{(2)} - X_i^{(1)})T^3\Big| = \|f''\|\mathbb{E}|T|^3 \sum_{i=1}^{m} \mathbb{E}\Big|X_i X_i^{(1)}(X_i^{(2)} - X_i^{(1)})\Big|.$$

A Taylor expansion of $f'(S_i^{(2)}T)$ about W gives

$$\mathbb{E}Wf(W) = \sum_{i=1}^{m} \mathbb{E}X_i X_i^{(1)} \mathbb{E}T^2 f'(W) + R_1 + R_2 + R_3,$$

where

$$|R_3| \leq \|f''\| \sum_{i=1}^{m} \left|\mathbb{E}X_i X_i^{(1)}\right| \mathbb{E}\left|X_i^{(2)}T^3\right| = \|f''\|\mathbb{E}|T|^3 \sum_{i=1}^{m} \left|\mathbb{E}X_i X_i^{(1)}\right| \mathbb{E}\left|X_i^{(2)}\right|.$$

Now, since $\sum_{i=1}^{m} \mathbb{E}X_i X_i^{(1)} = \mathbb{E}S^2 = 1$, we have

$$\begin{aligned}
EWf(W) &= \mathbb{E}T^2 f'(W) + R_1 + R_2 + R_3 \\
&= \sum_{j=1}^{n} \mathbb{E}Y_j T f'(W) + R_1 + R_2 + R_3 \\
&= \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} f'(W) + \sum_{j=1}^{n} \mathbb{E}Y_j T_j^{(1)} f'(W) + R_1 + R_2 + R_3. \qquad (4.16)
\end{aligned}$$

We now deal with the first term of (4.16). Taylor expanding $f'(W)$ about $ST_j^{(2)}$, and using independence and that $W - ST_j^{(2)} = Y_j^{(2)}S$, gives

$$\sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} f'(W) = \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} \mathbb{E}f'(ST_j^{(2)}) + R_4,$$

where

$$|R_4| \leq \|f''\| \sum_{j=1}^{n} \mathbb{E}\Big|Y_j Y_j^{(1)} Y_j^{(2)} S\Big| \leq \|f''\| \sum_{j=1}^{n} \mathbb{E}\Big|Y_j Y_j^{(1)} Y_j^{(2)}\Big|,$$

and we used the Cauchy–Schwarz inequality to bound $\mathbb{E}|S| \leq \{\mathbb{E}S^2\}^{1/2} = 1$. We now Taylor expand $f'(ST_j^{(2)})$ about W, and then use that $\sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} = \mathbb{E}T^2 = 1$, to obtain

$$\sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} f'(W) = \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} \mathbb{E}f'(W) + R_4 + R_5 = \mathbb{E}f'(W) + R_4 + R_5,$$

where

$$|R_5| \leq \|f''\| \sum_{j=1}^{n} \left|\mathbb{E}Y_j Y_j^{(1)}\right| \mathbb{E}\Big|Y_j^{(2)} S\Big| \leq \|f''\| \sum_{j=1}^{n} \left|\mathbb{E}Y_j Y_j^{(1)}\right| \mathbb{E}\Big|Y_j^{(2)}\Big|,$$

and we again used that $\mathbb{E}|S| \leq 1$.

Finally, we deal with the second term of (4.16). We Taylor expand, and use independence and that the Y_j have zero mean, to obtain

$$\begin{aligned}\sum_{j=1}^{n} \mathbb{E}Y_j T_j^{(1)} f'(W) &= \sum_{j=1}^{n} \mathbb{E}Y_j T_j^{(1)} \mathbb{E}f'(ST_j^{(1)}) \\ &\quad + \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} ST_j^{(1)} f''(ST_j^{(1)}) + R_6 \\ &= \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} ST_j^{(1)} f''(ST_j^{(1)}) + R_6,\end{aligned}$$

where

$$|R_6| \leq \|f^{(3)}\| \sum_{j=1}^{n} \mathbb{E}\Big|Y_j (Y_j^{(1)})^2 S^2 T_j^{(1)}\Big| = \|f^{(3)}\| \sum_{j=1}^{n} \mathbb{E}\Big|Y_j (Y_j^{(1)})^2 T_j^{(1)}\Big|.$$

We bound $\mathbb{E}|Y_j (Y_j^{(1)})^2 T_j^{(1)}|$ as follows,

$$\begin{aligned}\mathbb{E}\Big|Y_j (Y_j^{(1)})^2 T_j^{(1)}\Big| &= \mathbb{E}\Big|Y_j (Y_j^{(1)})^2 (T_j^{(2)} + Y_j^{(2)} - Y_j^{(1)})\Big| \\ &\leq \mathbb{E}\Big|Y_j (Y_j^{(1)})^2\Big| \mathbb{E}\Big|T_j^{(2)}\Big| + \mathbb{E}\Big|Y_j Y_j^{(1)} (Y_j^{(2)} - Y_j^{(1)})\Big| \\ &\leq \mathbb{E}\Big|Y_j (Y_j^{(1)})^2\Big| \Big[\mathbb{E}|T| + \mathbb{E}\Big|Y_j^{(2)}\Big|\Big] + \mathbb{E}\Big|Y_j Y_j^{(1)} (Y_j^{(2)} - Y_j^{(1)})\Big| \\ &\leq \mathbb{E}\Big|Y_j (Y_j^{(1)})^2\Big| \Big[1 + \mathbb{E}\Big|Y_j^{(2)}\Big|\Big] + \mathbb{E}\Big|Y_j Y_j^{(1)} (Y_j^{(2)} - Y_j^{(1)})\Big|,\end{aligned}$$

and therefore

$$|R_6| \leq \|f^{(3)}\| \sum_{j=1}^{n} \bigg\{ \Big[1 + \mathbb{E}\Big|Y_j^{(2)}\Big|\Big] \mathbb{E}\Big|Y_j (Y_j^{(1)})^2\Big| + \mathbb{E}\Big|Y_j Y_j^{(1)} (Y_j^{(2)} - Y_j^{(1)})\Big| \bigg\}.$$

Writing $T_j^{(1)} = T_j^{(2)} + Y_j^{(2)} - Y_j^{(1)}$, we see that

$$\sum_{j=1}^{n} \mathbb{E}Y_j T_j^{(1)} f'(W) = \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} ST_j^{(2)} f''(ST_j^{(1)}) + R_6 + R_7,$$

where

$$|R_7| = \bigg| \sum_{j=1}^{n} \mathbb{E}Y_j Y_j^{(1)} (Y_j^{(2)} - Y_j^{(1)}) S f''(ST_j^{(1)}) \bigg| \leq \|f''\| \sum_{j=1}^{n} \mathbb{E}\Big|Y_j Y_j^{(1)} (Y_j^{(2)} - Y_j^{(1)})\Big|.$$

Making repeated use of Taylor expansions and our usual independence arguments gives

$$\begin{aligned}\sum_{j=1}^{n}\mathbb{E}Y_jT_j^{(1)}f'(W)&=\sum_{j=1}^{n}\mathbb{E}Y_jY_j^{(1)}\mathbb{E}ST_j^{(2)}f''(ST_j^{(2)})+R_6+R_7+R_8\\&=\sum_{j=1}^{n}\mathbb{E}Y_jY_j^{(1)}\mathbb{E}STf''(ST_j^{(2)})+R_6+R_7+R_8+R_9\\&=\sum_{j=1}^{n}\mathbb{E}Y_jY_j^{(1)}\mathbb{E}STf''(W)+R_6+R_7+R_8+R_9+R_{10}\\&=\mathbb{E}Wf''(W)+R_6+R_7+R_8+R_9+R_{10},\end{aligned}$$

where

$$\begin{aligned}|R_8|&\leq\|f^{(3)}\|\sum_{j=1}^{n}\mathbb{E}\Big|Y_jY_j^{(1)}(Y_j^{(2)}-Y_j^{(1)})S^2T_j^{(2)}\Big|,\\|R_9|&=\left|\sum_{j=1}^{n}\mathbb{E}Y_jY_j^{(1)}\mathbb{E}Y_j^{(2)}Sf''(ST_j^{(2)})\right|\leq\|f''\|\sum_{j=1}^{n}\Big|\mathbb{E}Y_jY_j^{(1)}\Big|\mathbb{E}\Big|Y_j^{(2)}\Big|,\\|R_{10}|&\leq\|f^{(3)}\|\sum_{j=1}^{n}\Big|\mathbb{E}Y_jY_j^{(1)}\Big|\mathbb{E}\Big|Y_j^{(2)}S^2T\Big|\leq\|f^{(3)}\|\sum_{j=1}^{n}\Big|\mathbb{E}Y_jY_j^{(1)}\Big|\mathbb{E}\Big|Y_j^{(2)}\Big|.\end{aligned}$$

Writing $T_j^{(2)}=T_j^{(3)}+Y_j^{(3)}-Y_j^{(2)}$ and using independence leads to the following bound for $|R_8|$:

$$\begin{aligned}|R_8|&\leq\|f^{(3)}\|\sum_{j=1}^{n}\mathbb{E}S^2\mathbb{E}\Big|Y_jY_j^{(1)}(Y_j^{(2)}-Y_j^{(1)})(T_j^{(3)}+Y_j^{(3)}-Y_j^{(2)})\Big|\\&\leq\|f^{(3)}\|\sum_{j=1}^{n}\Big\{\mathbb{E}\Big|Y_jY_j^{(1)}(Y_j^{(2)}-Y_j^{(1)})\Big|\mathbb{E}\Big|T_j^{(3)}\Big|+\mathbb{E}\Big|Y_jY_j^{(1)}(Y_j^{(2)}-Y_j^{(1)})(Y_j^{(3)}-Y_j^{(2)})\Big|\Big\}\\&\leq\|f^{(3)}\|\sum_{j=1}^{n}\Big\{\Big[1+\mathbb{E}\Big|Y_j^{(3)}\Big|\Big]\mathbb{E}\Big|Y_jY_j^{(1)}(Y_j^{(2)}-Y_j^{(1)})\Big|+\mathbb{E}\Big|Y_jY_j^{(1)}(Y_j^{(2)}-Y_j^{(1)})(Y_j^{(3)}-Y_j^{(2)})\Big|\Big\},\end{aligned}$$

where we used that $\mathbb{E}|T_j^{(3)}|=\mathbb{E}|T-Y_j^{(3)}|\leq\mathbb{E}|T|+\mathbb{E}|Y_j^{(3)}|\leq1+\mathbb{E}|Y_j^{(3)}|$ to obtain the last inequality.

Adding the bounds, we have shown that

$$\mathbb{E}Wf(W)=\mathbb{E}f'(W)+\mathbb{E}Wf''(W)+\sum_{k=1}^{10}R_k.$$

We therefore have

$$|\mathbb{E}h(W) - \mathrm{VG}^{1,0}_{1,0}h| = |\mathbb{E}\{Wf''(W) + f'(W) - Wf(W)\}| \leq \sum_{k=1}^{10} |R_k|,$$

which on summing up the remainder terms completes the proof. □

Remark 4.18. Provided we imposed stronger moment conditions on the X_i and Y_j, we could use symmetry arguments, similar to those used in the proof of Theorem 4.4, to obtain a (quite complicated) bound with a faster convergence rate. For space considerations we did not carry out this task.

Chapter 5

Stein's method for Product Normal distributions

In the previous chapter we used the local coupling approach to obtain bounds on the error in approximating the statistic (4.2), which is given by

$$W_r = \sum_{i,j,l=1}^{m,n,r} X_{ik}Y_{jk},$$

by its limiting $\mathrm{VG}(r,0,1,0)$ distribution. In this chapter we use zero-bias and exchangeable pair couplings to establish Variance-Gamma limit theorems for the statistic W_1, where the X_{i1} and Y_{j1} are, again, independent random variables with mean zero and unit variance.

We also consider the more general problem of of obtaining a bound on the error in approximating the statistic

$$S_r = \prod_{k=1}^{r}\bigg(\frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k} X_{ik}\bigg), \tag{5.1}$$

where the X_{ik} are independent random variables with mean zero and unit variance, by its limiting distribution as $n_k \to \infty$ for $k = 1, \ldots, r$, which by the central limit theorem is the product of r independent standard normal distributions. We proceed by obtaining two of the three key ingredients of Stein's method for the distribution of the product of independent standard normals: a Stein equation and a coupling technique which involves a generalisation of the zero-bias transformation. Unfortunately, we have been unable to obtain the final ingredient, which is bounds on the derivatives of the solution to the Stein equation. This is left as an interesting, but possibly difficult, open problem.

5.1 A Stein equation for products of independent central normal and $\chi^2_{(1)}$ variables

In this section we obtain a Stein equation for the distribution of the the product of r independent mean zero normal, and $\chi^2_{(1)}$ variables. The probability density functions of the products of independent central normal, and $\chi^2_{(1)}$ variables are shown by Springer and Thompson [70] to be a Meijer G-function. The Meijer G-function is defined by the contour integral:

$$G^{m,n}_{p,q}\left(z \;\middle|\; \begin{matrix} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{matrix}\right) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} z^{-s}\frac{\prod_{j=1}^m \Gamma(s+b_j)\prod_{j=1}^n \Gamma(1-a_j-s)}{\prod_{j=n+1}^p \Gamma(s+a_j)\prod_{j=m+1}^q \Gamma(1-b_j-s)}\,\mathrm{d}s,$$

where c is a real constant defining a Bromwich path separating the poles of $F(s+b_j)$ from those of $F(1-a_j-s)$ and where we use the convention that $\prod_{j=r}^{-1} a_j = 1$. A more detailed discussion of the Meijer G-function and examples are given in Bateman [12], pp. 374–379.

For the case $a_1 = \ldots = a_p = b_1 = \ldots b_q = 0$ we write $G^{m,n}_{p,q}(z \mid 0)$, and for the cases $p=0$ and $q=0$ we write $G^{m,n}_{p,q}(z \mid b_1,\ldots,b_q)$ and $G^{m,n}_{p,q}(z \mid a_1,\ldots,a_q)$, respectively. With this notation and the above definition of the Meijer G-function the theorems of Springer and Thompson are as follows:

Theorem 5.1. *The probability density function of the product $Z_r = X_1X_2\cdots X_r$ independent normal random variables $N(0,\sigma^2_{x_i})$, $i=1,2,\ldots,r$, is a Meijer G-function multiplied by a normalizing constant, namely,*

$$p_{Z_r}(x) = \frac{1}{(2\pi)^{r/2}\sigma_r} G^{r,0}_{0,r}\left(\frac{x^2}{2^r\sigma_r^2} \;\middle|\; 0\right), \tag{5.2}$$

where $\sigma_r = \sigma_{x_1}\sigma_{x_2}\cdots\sigma_{x_r}$. If (5.2) holds then we write $Z_r \sim \mathrm{PN}(r,\sigma^2)$.

Theorem 5.2. *The probability density function of the product $Z_r = Y_1Y_2\cdots Y_r$ of r independent $\chi^2_{(d_i)}$, for $1 \le i \le r$, variables is a Meijer G-function multiplied by a normalizing constant, namely,*

$$p_{Z_r}(x) = \frac{1}{2^r\prod_{i=1}^r \Gamma(\frac{d_i}{2})} G^{r,0}_{0,r}\left(\frac{x}{2^r} \;\middle|\; \frac{d_1}{2}-1, \frac{d_2}{2}-1,\cdots,\frac{d_r}{2}-1\right). \tag{5.3}$$

We recover the familiar expressions for the probability density functions for $N(0,\sigma_x)$, $\mathrm{VG}(1,0,\sigma_x\sigma_y,0)$ and $\chi^2_{(d)}$ variables by using the following special cases of the Meijer G-function:

$$G^{1,0}_{0,1}(x \mid b) = x^b\mathrm{e}^{-x}, \qquad G^{2,0}_{0,2}(x \mid b,c) = 2x^{(b-c)/2}K_{b-c}(2\sqrt{x}).$$

Before we obtain a Stein equation for the Product Normal distribution, we introduce some notation and state a simple lemma that we will make use of later.

Recalling (3.28), for $n = 1, 2, 3, \ldots$ we define the operator $(x\frac{\mathrm{d}}{\mathrm{d}x})^n$ is defined recursively by

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^1 f(x) = xf'(x), \qquad \left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^n = x\frac{\mathrm{d}}{\mathrm{d}x}\left[\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{n-1} f(x)\right], \quad n = 2, 3, 4, \ldots.$$

Lemma 5.3. *Let r be a positive integer and suppose that $f \in C^r(\mathbb{R})$, then*

$$\frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r f(x) = \sum_{k=1}^{r} A_k^r x^{k-1} f^{(k)}(x), \tag{5.4}$$

where

$$A_1^r = A_r^r = 1 \qquad \textit{and} \qquad A_k^{r+1} = A_{k-1}^r + kA_k^r, \quad k = 2, 3, \ldots, r. \tag{5.5}$$

Proof. We prove the result by induction on r. The result is true for $r = 1$, so suppose $r \geq 2$. Using the inductive hypothesis to obtain the third equality we have

$$\begin{aligned}
\frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^{r+1} f(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r f(x)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left[x \cdot \frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r f(x)\right] \\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{k=1}^{r} A_k^r x^k f^{(k)}(x)\right) \\
&= \sum_{k=1}^{r} A_k^r [x^k f^{(k+1)}(x) + kx^{k-1} f^{(k)}(x)] \\
&= \sum_{j=1}^{r+1} \tilde{A}_j^{r+1} x^{j-1} f^{(j)}(x),
\end{aligned}$$

where $\tilde{A}_1^{r+1} = \tilde{A}_{r+1}^{r+1} = 1$ and $\tilde{A}_j^{r+1} = A_{j-1}^r + jA_j^r$, $j = 2, 3, 4, \ldots$. Hence, the result has been proved by induction on r. □

We can find A_k^r by using forward substitution in the recurrence equation (5.5). We have particularly simple expressions for A_{r-1}^r and A_2^r:

$$A_{r-1}^r = \frac{r(r-1)}{2}, \qquad A_2^r = 2^{r-1} - 1, \qquad \text{for } r \geq 2.$$

The following proposition suggests a Stein equation for the distribution of the product of independent mean zero normal variables.

Proposition 5.4. *Let Z be a real-valued random variable with mean zero and finite, non zero variance. Then $\mathscr{L}(Z) = \mathrm{PN}(r, \sigma^2)$ if and only if, for all $f : \mathbb{R} \to \mathbb{R}$ such that $f \in C^r(\mathbb{R})$, and $\mathbb{E}|Wf(W)| < \infty$ and $\mathbb{E}|W^{k-1}f^{(k)}(W)| < \infty$, $k = 1, 2, \ldots, r$, for $W \sim \mathrm{PN}(r, \sigma^2)$, we have*

$$\mathbb{E}\bigg\{\sigma^2\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=Z} - Zf(Z)\bigg\} = 0. \tag{5.6}$$

Proof. Necessity. We prove the result by induction on r. The result for the case $r = 1$ immediately follows from Lemma 2.3. Now suppose the result is true for $r = m \geq 1$. We write $W_m = \prod_{i=1}^m X_i$, where $X_i \sim N(0, \sigma_{x_i}^2)$ and the X_i are independent. Also, let $\sigma_m = \sigma_{x_1}\sigma_{x_2}\cdots\sigma_{x_m}$. Then

$$\begin{aligned}\mathbb{E}W_{m+1}f(W_{m+1}) &= \mathbb{E}(\mathbb{E}(W_{m+1}f(W_{m+1})|X_{m+1})) \\ &= \mathbb{E}\bigg[\mathbb{E}\bigg[\sigma_m^2 X_{m+1}^2 \sum_{k=1}^m A_k^m W_{m+1}^{k-1} f^{(k)}(W_{m+1})\bigg| X_{m+1}\bigg]\bigg] \\ &= \mathbb{E}\bigg[\sigma_m^2 \sum_{k=1}^m A_k^m X_{m+1}^{k+1} W_m^{k-1} f^{(k)}(W_{m+1})\bigg],\end{aligned}$$

where we used to the inductive hypothesis to obtain the second equality.

We now note that taking $f(x) = x^k g(\alpha x)$ in (2.2) leads to the following characterisation for $Z \sim N(0, \sigma^2)$:

$$\mathbb{E}Z^{k+1}g(\alpha Z) = \sigma^2\alpha\mathbb{E}Z^k g'(\alpha Z) + \sigma^2 k\mathbb{E}Z^{k-1}g(\alpha Z), \tag{5.7}$$

providing that the expectations exist.

Using (5.7) and independence gives

$$\begin{aligned}\mathbb{E}W_{m+1}f(W_{m+1}) &= \mathbb{E}\bigg[\mathbb{E}\bigg[\sigma_m^2 \sum_{k=1}^m A_k^m W_m^{k-1} X_{m+1}^{k+1} f^{(k)}(W_m X_{m+1})\bigg| X_1, X_2, \ldots, X_m\bigg]\bigg] \\ &= \mathbb{E}\bigg[\mathbb{E}\bigg[\sigma_m^2 \sum_{k=1}^m A_k^m W_m^{k-1} \cdot \sigma_{x_{m+1}}^2 [W_m X_{m+1}^k f^{(k+1)}(W_{m+1}) \\ &\quad + kX_{m+1}^{k-1} f^{(k)}(W_{m+1})]\bigg| X_1, X_2, \ldots, X_m\bigg]\bigg] \\ &= \sigma_{m+1}^2 \mathbb{E}\bigg[\sum_{k=1}^m A_k^m [W_{m+1}^k f^{(k+1)}(W_{m+1}) + kW_{m+1}^{k-1} f^{(k)}(W_{m+1})]\bigg] \\ &= \sigma_{m+1}^2 \mathbb{E}\bigg[\sum_{j=1}^{m+1} \tilde{A}_j^{m+1} W_{m+1}^j f^{(j)}(W_{m+1})\bigg],\end{aligned}$$

where $\tilde{A}_1^{m+1} = \tilde{A}_{m+1}^{m+1} = 1$ and $\tilde{A}_j^{m+1} = A_{j-1}^m + jA_j^m$, for $j = 2, 3, \ldots, m$. By Lemma 5.3 it follows that $\tilde{A}_j^{m+1} = A_j^{m+1}$, for $j = 1, 2, \ldots, m+1$, and so we have proved necessity by induction.

Sufficiency. This is established in the proof of Lemma 5.12, below. In the proof of Lemma 5.12 we establish that there is a unique probability distribution with zero mean and finite, non zero variance such that equation (5.6) holds, and since the $\mathrm{PN}(r, \sigma^2)$ distribution satisfies (5.6) the proof of sufficiency follows. □

Corollary 5.5. *Let $Y_i \sim \chi^2_{(1)}$, where $i = 1, 2, \ldots, r$, and the Y_i are independent. Define $W = \prod_{i=1}^r Y_i$. Then for all $f \in C^n(\mathbb{R})$ such that $\mathbb{E}|Wf(W)| < \infty$, $\mathbb{E}|f(W)| < \infty$ and $\mathbb{E}|W^k f^{(k)}(W)| < \infty$, for $k = 1, 2, \ldots, r$, we have*

$$\mathbb{E}\left\{\left[\frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r [xf(x^2)]\right]_{x=\sqrt{W}} - Wf(W)\right\} = 0.$$

Proof. Let X_i, for $i = 1, 2, \ldots, r$, be independent standard normal random variables. Define $Z_r = \prod_{i=1}^r X_i$. Then applying Proposition 5.4 with $f(x) = xg(x^2)$ gives

$$\mathbb{E}\left\{\left[\frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r [xg(x^2)]\right]_{x=Z_r} - Z_r^2 g(Z^2)\right\} = 0.$$

Since $X_i^2 \sim \chi^2_{(1)}$, it follows that $W = \prod_{i=1}^r Y_i \stackrel{\mathcal{D}}{=} \prod_{i=1}^r X_i^2 = Z_r^2$, which completes the proof. □

Proposition 5.4 suggests as Stein equation for the $\mathrm{PN}(r, \sigma^2)$ distribution:

$$\frac{\sigma^2}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r f(x) - xf(x) = h(x) - \mathrm{PN}_r^{\sigma^2} h, \tag{5.8}$$

where $\mathrm{PN}_r^{\sigma^2} h$ denotes the quantity $\mathbb{E}h(X)$ for $X \sim \mathrm{PN}(r, \sigma^2)$. Similarly, Corollary 5.5 suggests as Stein equation for the distribution of r independent $\chi^2_{(1)}$ variables:

$$\left[\frac{1}{y}\left(y\frac{\mathrm{d}}{\mathrm{d}y}\right)^r [yf(y^2)]\right]_{y=\sqrt{x}} - xf(x) = h(x) - \mathbb{E}h(Q), \tag{5.9}$$

where Q is distributed as the product of r independent $\chi^2_{(1)}$ random variables.

Taking $r = 1, 2$ in (5.8) we recover the familiar Stein equations for $N(0, \sigma^2)$ and $\mathrm{VG}(1, 0, \sigma, 0)$ distributions, and taking $r = 1$ in (5.9) gives the $\chi^2_{(1)}$ Stein equation (2.14) of Luk [44].

We end this section by making some remarks about the Product Normal Stein equation (5.8). Similar comments apply to the product $\chi^2_{(1)}$ Stein equation (5.9).

Remark 5.6. As was the case for the Variance-Gamma distributions, we could have obtained a first order Stein operator for the $\mathrm{PN}(r, \sigma^2)$ distributions using the density approach. However, this approach would lead to a complicated operator involving Meijer G-functions, which may not be amenable to the use of couplings. By Lemma 5.3, it follows that the $\mathrm{PN}(r,\sigma^2)$ Stein equation (5.8) is a r-th order linear differential equation, with very simple coefficients:

$$\sigma^2\sum_{k=1}^{r}A_k^r x^{k-1}f^{(k)}(x)-xf(x)=h(x)-\mathrm{PN}_r^{\sigma^2}h, \tag{5.10}$$

where the A_k^r can be computed using forward substitution in the recurrence relation (5.5). In Section 5.3 we see that the representation (5.10), together with zero bias couplings, leads to particularly simple proofs of limit theorems for statistics with an asymptotic $\mathrm{PN}(r,\sigma^2)$ distribution, provided we can bound the required derivatives of the solution to the Stein equation.

Remark 5.7. The problem of obtaining a (unique) bounded solution and bounds on its derivatives has been successfully achieved for the case $r=1$, and in Chapter 3 we obtained bounds on the first four derivatives of the solution for the case $r=2$. These smoothness estimates can be established through the use of either direct analytical calculations or a generator approach. Unfortunately, neither of these approaches have proved successful at yielding bounds on the derivatives of the solution of the Stein equation for the case $r\geq 3$. The generator method is not applicable to the Stein equation (5.8) in the case $r\geq 3$, because there does not exist a generator of degree greater than two for a stochastic process (see Theorem 1.4 and Example 1.1 of Chapter 7 of Durrett [20]).

Solving the Stein equation by a standard analytical method and then bounding the solutions derivatives via a direct calculation seems to be unrealistic. Indeed, consider the homogeneous differential equation

$$\frac{\sigma^2}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r f(x)-xf(x)=0.$$

Letting $f(x)=g\left((-1)^k\frac{x^2}{2^r\sigma^2}\right)$ leads to the homogeneous differential equation

$$\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r g(x)-(-1)^k xg(x)=0. \tag{5.11}$$

For $k=1,\ldots,r$, the Meijer G-function $G_{0,r}^{k,0}(x\mid 0)$ solves (5.11) (see Olver et al. [52], formula 16.21.1). Provided that $G_{0,r}^{1,0}(x\mid 0),\ldots,G_{0,r}^{r,0}(x\mid 0)$ are linearly independent, we could use variation of parameters to write down the general solution of the $\mathrm{PN}(2,\sigma^2)$ Stein equation, which would be in terms of the functions $G_{0,r}^{k,0}((-1)^k\frac{x^2}{2^r\sigma^2}\mid 0)$, for $k=1,\ldots,r$. However, this solution would take a complicated form involving Meijer G-functions and bounding would not be straightforward. In the case $r=2$, we have $G_{0,2}^{1,0}(-\frac{x^2}{4}\mid 0)=I_0(x)$ and $G_{0,2}^{2,0}(\frac{x^2}{4}\mid 0)=2K_0(x)$, (see Erdélyi et al. [23], Section 5.6, formulas 3 and 4) and therefore obtaining bounds for the derivatives of the solution via an analytical approach is possible. However, this author is not aware that such simplifications exist for $r \geq 3$. Obtaining smoothness estimates for the solution for $r \geq 3$ is therefore left as a difficult open problem.

5.2 A generalisation of the zero-bias transformation

The zero-bias coupling for use with Stein's method was introduced by Goldstein and Reinert [29] for normal approximation. It is defined as follows:

Definition 5.8. *Let W be a mean zero random variable with finite, non zero variance σ^2. We say that W^* has the W-zero biased distribution if for all $f \in C^1(\mathbb{R})$ for which $\mathbb{E}Wf(W)$ exists,*

$$\mathbb{E}Wf(W) = \sigma^2 \mathbb{E}f'(W^*). \tag{5.12}$$

The above definition shows why we might like to use a zero-biasing method for normal approximation: it gives a way of splitting apart an expectation, and reduces normal approximation to bounding the quantity $\sigma^2\mathbb{E}(f'(W) - f'(W^*))$. Goldstein and Reinert [29] also give a construction of the zero-bias for random variables that can be expressed as in the form of a sum of independent variables, which is particularly useful for normal approximation.

Lemma 5.9. *Let $X_1, \ldots, X_n$ be independent mean zero random variables with $\mathbb{E}X_i^2 = \sigma_i^2$. Set $W = \sum_{i=1}^n X_i$ and $\mathbb{E}W^2 = \sigma^2$. Let I be a random index independent of the X_i such that*

$$\mathbb{P}(I = i) = \frac{\sigma_i^2}{\sigma^2}.$$

Let

$$W_i = W - X_i = \sum_{j \neq i} X_j.$$

Let X_i^* have the X_i-zero bias distribution, then $W_I + X_I^*$ has the W-zero biased distribution.

That is, for W the sum of independent mean zero random variables, one achieves the W-zero biased distribution by replacing a variable chosen with probability proportional to its variance by one chosen independently from its zero bias distribution.

Motivated by the zero bias transformation and the multivariate normal Stein equation, Goldstein and Reinert [30] extended the concept of the zero bias transformation to any finite dimension.

Definition 5.10. *Let Γ be an arbitrary index set and let $\mathbf{X} = \{X_\gamma : \gamma \in \Gamma\}$ be a collection of* *mean zero random variables with covariances $\mathbb{E}X_\alpha X_\beta = \sigma_{\alpha\beta}$. For pairs α, β with $\sigma_{\alpha\beta} \neq 0$, we say that the collection of variables $\mathbf{X}^{(\alpha,\beta)} = \{X_\gamma^{(\alpha,\beta)} : \gamma \in \Gamma\}$ has the $\mathbf{X}$-zero biased distribution in coordinates (α, β) if for all finite $I \in \Gamma$,*

$$\mathbb{E}\sum_{\beta\in I} X_\beta \frac{\partial}{\partial x_\beta} f(\mathbf{X}) = \mathbb{E}\sum_{\alpha\in I}\sum_{\beta\in I} \sigma_{\alpha\beta}\frac{\partial^2}{\partial x_\alpha \partial x_\beta} f(\mathbf{X}^{(\alpha,\beta)})$$

for all twice differentiable functions f for which the above expectations exist.

We now consider a new generalisation of the zero-bias transformation, which is motivated by the the zero bias transformation and the Product Normal Stein equation (5.8). The transformation we introduce is a natural generalisation of the zero bias transformation to the study of products of independent normal distributions in the same way that the multivariate zero bias transformation is a natural extension to random vectors in $\mathbb{R}^d$. In the next section we shall use the transformation to prove limit theorems for Product Normal distributions. We have the following definition:

Definition 5.11. *Let W be a mean zero random variable with finite, non zero variance σ^2. We say that W_r^* has the W-zero biased distribution of order r if for all $f \in C^r(\mathbb{R})$ for which $\mathbb{E}Wf(W)$ exists,*

$$\mathbb{E}Wf(W) = \sigma^2\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=W^{*(r)}}\bigg\}. \tag{5.13}$$

Putting $r = 1$ in (5.13) gives the W-zero-bias transform. The existence of the zero bias distribution of order r for any such W is established by the following lemma.

Lemma 5.12. *Let W be a mean zero random variable with finite, non zero variance σ^2. Then there exists a unique random variable $W^{*(r)}$ such that for all $f \in C^r(\mathbb{R})$ for which the relevant expectations exist we have*

$$\mathbb{E}Wf(W) = \sigma^2\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=W^{*(r)}}\bigg\}.$$

Proof. For $f \in C_c$, the collection of continuous functions with compact support, let

$$G_1(w) = \int_0^w f(t_1)\,\mathrm{d}t_1,$$

and for $r \geq 2$,

$$G_r(w) = \int_0^w \int_0^{t_r} \cdots \int_0^{t_2} \frac{1}{t_2 t_3 \cdots t_r} f(t_1)\,\mathrm{d}t_1\,\mathrm{d}t_2 \cdots \mathrm{d}t_r.$$

We also define a linear operator T by

$$Tf = \sigma^{-2}\mathbb{E}WG_r(W).$$

Then Tf exists, since $\mathbb{E}W^2 < \infty$. To see, moreover, that T is positive, take $f \geq 0$. Then G_r is increasing, and therefore W and $G_r(W)$ are positively correlated. Hence $\mathbb{E}WG_r(W) \geq \mathbb{E}W\mathbb{E}G_r(W) = 0$, and T is positive. Using the Riesz representation theorem (see, for example [27]) we have $Tf = \int f\,\mathrm{d}\nu$, for some unique Radon measure ν, which is a probability measure as $T1 = 1$. We now take

$$f(x) = \frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r g(x),$$

where $g \in C^r(\mathbb{R})$, with derivatives up to r-th order being continuous with compact support. Then

$$\begin{aligned}
\mathbb{E}WG_r(W) &= \mathbb{E}W\int_0^W\int_0^{t_r}\cdots\int_0^{t_2}\frac{1}{t_1t_2\cdots t_r}\left(t_1\frac{\mathrm{d}}{\mathrm{d}t_1}\right)^r g(t_1)\,\mathrm{d}t_1\,\mathrm{d}t_2\cdots\mathrm{d}t_r \\
&= \mathbb{E}W\int_0^W\int_0^{t_r}\cdots\int_0^{t_2}\frac{1}{t_2\cdots t_r}\frac{\mathrm{d}}{\mathrm{d}t_1}\left[\left(t_1\frac{d}{dt_1}\right)^{r-1}g(t_1)\right]\mathrm{d}t_1\,\mathrm{d}t_2\cdots\mathrm{d}t_r \\
&= \mathbb{E}W\int_0^W\int_0^{t_r}\cdots\int_0^{t_3}\frac{1}{t_2t_3\cdots t_r}\left[\left(t_1\frac{d}{dt_1}\right)^{r-1}g(t_1)\right]_{t_1=0}^{t_1=t_2}\mathrm{d}t_2\,\mathrm{d}t_3\cdots\mathrm{d}t_r \\
&= \mathbb{E}W\int_0^W\int_0^{t_r}\cdots\int_0^{t_3}\frac{1}{t_2t_3\cdots t_r}\left(t_2\frac{d}{dt_2}\right)^{r-1}g(t_2)\,\mathrm{d}t_2\,\mathrm{d}t_3\cdots\mathrm{d}t_r.
\end{aligned}$$

Iterating gives

$$\mathbb{E}WG_r(W) = \mathbb{E}W\int_0^W g'(t_r)\,\mathrm{d}t_r = \mathbb{E}W(g(W)-g(0)) = \mathbb{E}Wg(W),$$

which completes the proof. □

The zero bias transformation of order r has many interesting properties that are collected in the following lemma. These properties generalise some of the important properties of the zero bias transformation (see Goldstein and Reinert [29]).

Lemma 5.13. *Let W be a mean zero variable with finite, non zero variance σ^2, and let $W^{*(r)}$ have the W-zero biased distribution of order r in accordance with Definition 5.11.*

(i) Let $X_i \sim N(0,\sigma_{x_i}^2)$ be independent and define $Z_r = \prod_{i=1}^r X_i$. Let $\sigma^2 = \sigma_{x_1}^2\sigma_{x_2}^2\cdots\sigma_{x_r}^2$. Then Z_r is the unique fixed point of the W-zero biased transformation of order r.

(ii) For $p \geq 1$ we have

$$\sigma^2 \mathbb{E}(W^{*(r)})^p = \frac{\mathbb{E}W^{p+2}}{(p+1)^r}, \tag{5.14}$$

$$\sigma^2 \mathbb{E}|W^{*(r)}|^p = \frac{\mathbb{E}|W|^{p+2}}{(p+1)^r}. \tag{5.15}$$

(iii) Suppose $W = \prod_{j=1}^r W_j$, where the W_j are independent random variables with zero mean and finite variance and $W_1^*, \ldots, W_r^*$ are independent random variables with W_j^* having the W_j-zero biased distribution. Then

$$W^{*(r)} = \prod_{j=1}^r W_j^* \tag{5.16}$$

has the W-zero biased distribution of order r.

(iv) For $c \in \mathbb{R}$, $cW^{*(r)}$ has the cW-zero biased distribution of order r.

Proof. (i) This is immediate from Proposition 5.4.

(ii) Note that

$$x\frac{\mathrm{d}}{\mathrm{d}x}(x^{p+1}) = (p+1)x^{p+1}.$$

We therefore have that

$$\frac{1}{x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right)^r x^{p+1} = \frac{1}{x}\cdot(p+1)^r x^{p+1} = (p+1)^r x^p.$$

Hence, substituting $x^{p+1}/(p+1)^r$ for $f(x)$ in the characterising equation (5.13) yields (5.14). Similarly, we can verify (5.15) by substituting $x^{p+1}\mathrm{sgn}(x^p)/(p+1)^r$ for $f(x)$ in the characterising equation (5.13). Note that, since $p \geq 1$, this function is differentiable at $x = 0$.

(iii) We prove that $\prod_{k=1}^r W_k^*$ has the W-zero bias distribution of order r by an inductive proof on r that is very similar to the one used to prove Proposition 5.4. This time, the base case is that $\mathbb{E}W_1 f(W_1) = \sigma^2 \mathbb{E}f'(W^{*(1)})$, which is the characterising equation (5.12) for the W-zero bias transformation.

Suppose the result is true for $r = m \geq 1$. We write $Y_m = \prod_{i=1}^m W_i$, for $m \geq 1$. We also suppose that $\mathrm{Var}W_i = \sigma_{w_i}^2$ and let $\sigma_m = \sigma_{w_1}\sigma_{w_2}\cdots\sigma_{w_m}$. Then

$$\begin{aligned}\mathbb{E}Y_{m+1}f(Y_{m+1}) &= \mathbb{E}[\mathbb{E}[Y_{m+1}f(Y_{m+1})|W_m]] \\ &= \mathbb{E}\bigg[\mathbb{E}\bigg[\sigma_m^2 W_{m+1}^2 \sum_{k=1}^m A_k^m W_{m+1}^{k-1}(Y_m^{*(m)})^{k-1} f^{(k)}(W_{m+1}Y_m^{*(m)})\bigg| W_m\bigg]\bigg]\end{aligned}$$

$$= \mathbb{E}\bigg[\sigma_m^2 \sum_{k=1}^m A_k^m W_{m+1}^{k+1} (Y_m^{*(m)})^{k-1} f^{(k)}(W_{m+1} Y_m^{*(m)})\bigg],$$

where we used to the characterising equation (5.13) to obtain the second equality.

We now note that taking $f(x) = x^k g(\alpha x)$ in the characterising equation (5.12) for the W-zero bias distribution leads to the following characterisation for a random variable Z with mean zero and non zero, finite variance:

$$\mathbb{E}Z^{k+1} g(\alpha Z) = \sigma^2 \alpha \mathbb{E} Z^k g'(\alpha Z^*) + \sigma^2 k \mathbb{E} Z^{k-1} g(\alpha Z^*), \qquad (5.17)$$

providing that the expectations exist.

Using (5.17) and independence gives

$$\begin{aligned}
\mathbb{E}Y_{m+1} f(Y_{m+1}) &= \mathbb{E}\bigg[\mathbb{E}\bigg[\sigma_m^2 \sum_{k=1}^m A_k^m (Y_m^{*(m)})^{k-1} W_{m+1}^{k+1} f^{(k)}(Y_m^{*(m)} W_{m+1}^*)\bigg| W_1, W_2, \ldots, W_m\bigg]\bigg] \\
&= \mathbb{E}\bigg[\mathbb{E}\bigg[\sigma_m^2 \sum_{k=1}^m A_k^m (Y_m^{*(m)})^{k-1} \cdot \sigma_{w_{m+1}}^2 [Y_m^{*(m)} (W_{m+1}^*)^k f^{(k+1)}(Y_m^{*(m)} W_{m+1}^*) \\
&\quad + k(W_{m+1}^*)^{k-1} f^{(k)}(W_{m+1})]\bigg| W_1, W_2, \ldots, W_m\bigg]\bigg] \\
&= \sigma_{m+1}^2 \mathbb{E}\bigg[\sum_{k=1}^m A_k^m [(Y_m^{*(m)} W_{m+1}^*)^k f^{(k+1)}(W_{m+1}) \\
&\quad + k(Y_m^{*(m)} W_{m+1}^*)^{k-1} f^{(k)}(Y_m^{*(m)} W_{m+1}^*)]\bigg] \\
&= \sigma_{m+1}^2 \mathbb{E}\bigg[\sum_{j=1}^{m+1} \tilde{A}_j^{m+1} (Y_m^{*(m)} W_{m+1}^*)^j f^{(j)}(Y_m^{*(m)} W_{m+1}^*)\bigg],
\end{aligned}$$

where $\tilde{A}_1^{m+1} = \tilde{A}_{m+1}^{m+1} = 1$ and $\tilde{A}_j^{m+1} = A_{j-1}^m + jA_j^m$, for $j = 2, 3, \ldots, m$. By Lemma 5.3, it follows that $\tilde{A}_k^{m+1} = A_k^{m+1}$, for $k = 1, 2, \ldots, m+1$, and therefore

$$\mathbb{E}Y_{m+1} f(Y_{m+1}) = \mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r+1} f(x)\bigg]_{x = Y_m^{*(m)} W_{m+1}^*}\bigg\}.$$

But by the inductive hypothesis $Y_m^{*(m)} \stackrel{\mathcal{D}}{=} \prod_{i=1}^m W_i^*$, and so

$$\mathbb{E}Y_{m+1} f(Y_{m+1}) = \mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r+1} f(x)\bigg]_{x = \prod_{k=1}^{m+1} W_k^*}\bigg\}.$$

By the characterising equation (5.13) we have that

$$\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r+1}f(x)\bigg]_{x=\prod_{k=1}^{m+1}W_k^*}\bigg\}=\mathbb{E}Y_{m+1}f(Y_{m+1})=\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r+1}f(x)\bigg]_{Y_{m+1}^{*(m+1)}}\bigg\}.$$

Suppose now that

$$\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r+1}f(x)=g(x),$$

where $g\in C_c$, the collection of continuous functions with compact support. Then, we have shown that $\mathbb{E}g(\prod_{k=1}^{m+1}W_k^*)=\mathbb{E}g(Y_{m+1}^{*(m+1)})$, and so $g(\prod_{k=1}^{m+1}W_k^*)$ and $g(Y_{m+1}^{*(m+1)})$ agree for all $g\in C_c$. Hence, the random variables $\prod_{k=1}^{m+1}W_k^*$ and $Y_{m+1}^{*(m+1)}$ must be equal in distribution. Therefore the result has been proven by induction on r.

(iv) Given any function g such that $\mathbb{E}Wg(W)$ exists, consider $\mathbb{E}cWg(cW)$. Let $\tilde{g}(x)=cg(cx)$. Then $\tilde{g}^{(k)}(x)=c^{k+1}g^{(k)}(cx)$. As $W^{*(r)}$ has the W-zero bias distribution of order r,

$$\begin{aligned}\mathbb{E}cWg(cW)&=\mathbb{E}W\tilde{g}(W)\\&=\sigma^2\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r}\tilde{g}(x)\bigg]_{x=W^{*(r)}}\bigg\}\\&=(c\sigma)^2\mathbb{E}\bigg\{\bigg[\frac{1}{y}\bigg(y\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^{r}g(y)\bigg]_{y=cW^{*(r)}}\bigg\}.\end{aligned}$$

Hence $cW^{*(r)}$ has the cW-zero bias distribution of order r. □

Some remarks are in order regarding Lemma 5.13.

Remark 5.14. Consider the problem of obtaining the zero bias transformation of order r for a random variable $W=\prod_{k=1}^{r}W_k$, where the W_k are independent. By property (iii) of Lemma 5.13, we have that the zero bias transformation of order r is $W^{*(r)}=\prod_{k=1}^{r}W_k^*$. It may then be possible to construct the zero bias transformations for each of the W_k by using one of the constructions given by Goldstein and Reinert [29]. For example, if $W_k=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_{ik}$, where the X_{ik} are independent random variables with mean zero and non zero, finite variance, then we can construct W_k^* using Lemma 5.9.

Remark 5.15. Another coupling, with similar properties to the zero bias coupling, that is commonly used with Stein's method to prove approximation results is the size bias coupling (for an application of this coupling to normal approximation see Baldi, Rinott and Stein [4]). It is defined as follows: if $W\geq 0$ has mean $\mu>0$, we say W^s has the W-size biased distribution if for all f such that $\mathbb{E}Wf(W)$ exists,

$$\mathbb{E}Wf(W) = \mu\mathbb{E}f(W^s).$$

Luk [44] showed that if we suppose $W = \prod_{k=1}^r W_k$, where the W_k are positive, independent random variables and let $W_1^s, \ldots, W_r^s$ be independent random variables with W_k^s having the W_k-size biased distribution, then

$$W^s = \prod_{k=1}^r W_k^s \tag{5.18}$$

has the W-size biased distribution. This construction is similar to the construction of part (iii) of Lemma 5.13, with the difference being that our construction shows that the product of r zero bias distributions has the W-zero bias distribution of order r, rather than the W-zero bias distribution.

We end this section by presenting an interesting relationship between the W-zero bias distribution of order r and the W-*square bias distribution*, which arose from a discussion with Larry Goldstein. For any random variable W with finite second moment, we say that $W^\square$ has the W-square bias distribution if for all f such that $\mathbb{E}W^2f(W)$ exists,

$$\mathbb{E}W^2f(W) = \mathbb{E}W^2\mathbb{E}f(W^\square). \tag{5.19}$$

Peköz et al. [54] observed that for a non-negative random variable W with positive mean, a random variable $W^\square$ having the W-square bias distribution can be constructed by taking the size bias distribution of W^s. For this reason, for non-negative random variables, the W-square bias distribution is also known as the W-*double size bias distribution.*

Square biasing, zero biasing and size biasing are special cases of a general form of distributional biasing that was given in Theorem 2.1 of Goldstein and Reinert [31]. Given a function $P(x)$ with $m \in \{0, 1, \ldots\}$ sign changes, and a distribution W which satisfies m orthogonality relations $\mathbb{E}W^iP(W) = 0$, $i = 0, \ldots, m-1$, there exists a distribution $W^{(P)}$ satisfying

$$\mathbb{E}P(W)f(W) = \alpha\mathbb{E}f^{(m)}(W^{(p)})$$

when $\alpha = \mathbb{E}P(W)W^m/m! > 0$. Following Chen et al. [18], for square biasing we take $P(x) = x^2$, and since $P(x)$ has no sign changes we see that the distribution of $W^\square$ exists for any distribution with finite second moment.

We now state our relationship, which is a natural generalisation of the relation between the zero bias distribution and the square bias distribution that is given in Proposition 2.3 of Chen et al.

[18].

Proposition 5.16. *Let W be a random variable with zero mean and finite, non zero variance σ^2, and let $W^\square$ have the W-square bias distribution. Let $U_1,\ldots,U_r$ be a sequence of independent $U[0,1]$ random variables, which are also independent of $W^\square$. Define $V_r=\prod_{k=1}^r U_k$. Then, the random variable*

$$W^{*(r)}\stackrel{\mathcal{D}}{=}V_rW^\square \tag{5.20}$$

has the W-zero bias distribution of order r.

Proof. Suppose $f\in C^r(\mathbb{R})$. For $r\geq 2$ we have

$$\begin{aligned}
\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=V_rW^\square}\bigg\}&=\mathbb{E}\int_{[0,1]^r}\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=u_1\cdots u_rW^\square}\mathrm{d}u_1\cdots\mathrm{d}u_r\\
&=\mathbb{E}\int_{[0,1]^r}\bigg[\frac{\mathrm{d}}{\mathrm{d}x}\bigg(x\cdot\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r-1}f(x)\bigg)\bigg]_{x=u_1\cdots u_rW^\square}\mathrm{d}u_1\cdots\mathrm{d}u_r\\
&=\mathbb{E}\int_{[0,1]^r}\frac{d}{du_r}\bigg(u_r\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r-1}f(x)\bigg]_{x=u_1\cdots u_rW^\square}\bigg)\mathrm{d}u_1\cdots\mathrm{d}u_r\\
&=\mathbb{E}\int_{[0,1]^{r-1}}\bigg[u_r\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r-1}f(x)\bigg]_{x=u_1\cdots u_rW^\square}\bigg]_{u_r=0}^{u_r=1}\mathrm{d}u_1\cdots\mathrm{d}u_{r-1}\\
&=\mathbb{E}\int_{[0,1]^{r-1}}\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r-1}f(x)\bigg]_{x=u_1\cdots u_{r-1}W^\square}\mathrm{d}u_1\cdots\mathrm{d}u_{r-1}\\
&=\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^{r-1}f(x)\bigg]_{x=V_{r-1}W^\square}\bigg\},
\end{aligned}$$

where we used the chain rule to obtain the third equality. Repeating this procedure gives

$$\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=V_rW^\square}\bigg\}=\mathbb{E}\bigg\{\bigg[\frac{1}{x}\cdot x\frac{\mathrm{d}}{\mathrm{d}x}f(x)\bigg]_{x=V_1W^\square}\bigg\}=\mathbb{E}f'(V_1W^\square)=\mathbb{E}f'(U_1W^\square).$$

Hence, for all $r\geq 1$, we have

$$\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=V_rW^\square}\bigg\}=\mathbb{E}f'(U_1W^\square). \tag{5.21}$$

Suppose now that

$$\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)=g(x),$$

where $g\in C_c$, the collection of continuous functions with compact support. Then, using (5.21) to obtain the second equality and the characterisation (5.19) to obtain the fifth equality, we have for all $r \geq 1$,

$$\begin{aligned}\sigma^2\mathbb{E}g(V_rW^{\square}) &= \sigma^2\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=V_rW^{\square}}\bigg\} \\ &= \sigma^2\mathbb{E}f'(U_1W^{\square}) \\ &= \sigma^2\mathbb{E}\int_0^1 f'(u_1W^{\square})\,\mathrm{d}u_1 \\ &= \sigma^2\mathbb{E}\bigg(\frac{f(W^{\square})-f(0)}{W^{\square}}\bigg) \\ &= \mathbb{E}\bigg(W^2\frac{f(W)-f(0)}{W}\bigg) \\ &= \mathbb{E}Wf(W) - f(0)\mathbb{E}W \\ &= \mathbb{E}Wf(W).\end{aligned}$$

Hence, if $W^{*(r)}$ has the W-zero bias distribution of order r, we have for all $r \geq 1$,

$$\sigma^2\mathbb{E}g(V_rW^{\square}) = \mathbb{E}Wf(W) = \sigma^2\mathbb{E}\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=W^{*(r)}} = \sigma^2\mathbb{E}g(W^{*(r)}).$$

Since the expectation of $g(V_rW^{\square})$ and $g(W^{*(r)})$ are equal for all $g \in C_c$, the random variables $V_rW^{\square}$ and $W^{*(r)}$ must be equal in distribution. □

From Proposition 5.16 and part (iii) of Lemma 5.13 we are able to deduce a construction of the W-square bias distribution, where W is a product of independent random variables. This construction is similar to the constructions (5.16) and (5.18) for the zero bias and size bias distributions, respectively.

Proposition 5.17. *Suppose $W = \prod_{k=1}^r W_k$, where the W_k are independent random variables and let $W_1^{\square}, \ldots, W_r^{\square}$ be independent random variables with $W_k^{\square}$ having the W_k-square biased distribution, then*

$$W^{\square} = \prod_{k=1}^r W_k^{\square} \tag{5.22}$$

has the W-square biased distribution.

Proof. Let $U_1, \ldots, U_r$ be a sequence of independent $U[0,1]$ random variables, which are also independent of $W_1^{\square}, \ldots, W_r^{\square}$, and define $V_r = \prod_{k=1}^r U_k$. Then, using construction (5.16) to obtain the first equality and construction (5.20) to obtain the second equality, we have

$$W^{*(r)} \stackrel{\mathcal{D}}{=} \prod_{k=1}^r W_k^* \stackrel{\mathcal{D}}{=} \prod_{k=1}^r U_kW_k^{\square} = V_r\prod_{k=1}^r W_k^{\square}.$$

But from Proposition 5.16 we have $W^{*(r)} \stackrel{\mathcal{D}}{=} V_r W^{\square}$ and, as V_r is non zero with probability 1, it follows $W^{\square}$ and $\prod_{k=1}^{r} W_k^{\square}$ must be equal in distribution. $\square$

Remark 5.18. Recall that for a non-negative random variable W with positive mean, a random variable $W^{\square}$ having the W-square bias distribution can be constructed by taking the size bias distribution of W^s. Hence, for the case $W \geq 0$ and $\mathbb{E}W > 0$, the construction (5.22) of Proposition 5.17 follows immediately from construction (5.18). We have

$$W^{\square} \stackrel{\mathcal{D}}{=} (W^s)^s \stackrel{\mathcal{D}}{=} \left(\prod_{k=1}^{r} W_k^s\right)^s \stackrel{\mathcal{D}}{=} \prod_{k=1}^{r} (W_k^s)^s \stackrel{\mathcal{D}}{=} \prod_{k=1}^{r} W_k^{\square},$$

as required.

5.3 Zero bias coupling bounds for Product Normal approximation

The construction of Lemma 5.13 allows us to obtain bounds on the error, in a weak convergence setting, in the approximations of certain statistics that have an asymptotic Product Normal distribution, in terms of the derivatives of the solution of the Stein equation (5.8). As we have bounds on the derivatives of the solution of the Stein equation (5.8) in the cases $r = 1, 2$ we are able to obtain simple proofs of limit theorems for normal and Symmetric Variance-Gamma distributions. Unfortunately, we have not been able to obtain bounds on the derivatives of the solution of the Stein equation for $r \geq 3$, and this remains an open problem.

Theorem 5.19. *Let W be a mean zero random variable with variance σ^2. Suppose that $(W, W^{*(r)})$ is given on a joint probability space so that $W^{*(r)}$ has the W-zero biased distribution of order r. Then*

$$|\mathbb{E}h(W) - \mathrm{PN}_r^{\sigma^2} h| \leq \sigma^2 \sum_{k=1}^{r} A_k^r \Big[\|f_h^{(k)}\| \mathbb{E}|W^{k-1} - (W^{*(r)})^{k-1}| + \|f_h^{(k+1)}\| \mathbb{E}|W^{k-1}(W - W^{*(r)})|\Big], \tag{5.23}$$

where f_h is the solution of the Stein equation (5.8) and the A_k^r are defined by equation (5.4).

Suppose now that $W=\prod_{k=1}^r W_k$, where the W_k are independent. Then

$$\begin{aligned}|\mathbb{E}h(W)-\mathrm{PN}_r^{\sigma^2}h| \leq \sigma^2\sum_{k=1}^r A_k^r\bigg\{&\|f_h^{(k)}\|\sqrt{\mathbb{E}\{\mathbb{E}(W^{k-1}-(W^{*(r)})^{k-1})|W_1,\ldots,W_r)^2\}} \\ &+\|f_h^{(k+1)}\|\sqrt{\mathbb{E}\{\mathbb{E}((W^{*(r)})^{k-1}(W-W^{*(r)})|W_1,\ldots,W_r)^2\}} \\ &+\frac{1}{2}\|f_h^{(k+2)}\|\mathbb{E}|(W^{*(r)})^{k-1}(W-W^{*(r)})^2|\bigg\}. \qquad (5.24)\end{aligned}$$

Proof. We prove the second bound. The the first bound is obtained by a similar but simpler calculation. Using equation (5.8) we have

$$\begin{aligned}\mathbb{E}h(W)-\mathrm{PN}_r^{\sigma^2}h &= \mathbb{E}\bigg\{\sigma^2\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=W}-Wf(W)\bigg\} \qquad (5.25)\\ &=\sigma^2\mathbb{E}\bigg\{\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=W}-\bigg[\frac{1}{x}\bigg(x\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^r f(x)\bigg]_{x=W^{*(r)}}\bigg\} \\ &=\sigma^2\sum_{k=1}^r A_k^r\mathbb{E}[W^{k-1}f^{(k)}(W)-(W^{*(r)})^{k-1}f^{(k)}(W^{*(r)})].\end{aligned}$$

By Taylor expansion we have

$$\begin{aligned}|\mathbb{E}[W^{k-1}f^{(k)}(W)-(W^{*(r)})^{k-1}f^{(k)}(W^{*(r)})]| \leq\; &|\mathbb{E}(W^{k-1}-(W^{*(r)})^{k-1})f^{(k)}(W)| \\ &+|\mathbb{E}(W^{*(r)})^{k-1}(W-W^{*(r)})f^{(k+1)}(W)| \\ &+\frac{1}{2}\|f_h^{(k+2)}\|\mathbb{E}|(W^{*(r)})^{k-1}(W-W^{*(r)})^2|.\end{aligned}$$

For the first term, condition on $W_1,\ldots,W_r$ and then apply the Cauchy-Schwarz inequality:

$$\begin{aligned}|\mathbb{E}(W-W^{*(r)})f^{(k+1)}(W)| &= |\mathbb{E}[f^{(k)}(W)\mathbb{E}(W^{k-1}-(W^{*(r)})^{k-1}|W_1,\ldots,W_r)]| \\ &\leq \|f_h^{(k)}\|\sqrt{\mathbb{E}\{\mathbb{E}(W^{k-1}-(W^{*(r)})^{k-1})|W_1,\ldots,W_r)^2\}}.\end{aligned}$$

The second term is dealt with in a similar way:

$$\begin{aligned}|\mathbb{E}(W^{*(r)})^{k-1}(W-W^{*(r)})f^{(k+1)}(W)| &= |\mathbb{E}[f^{(k+1)}(W)\mathbb{E}((W^{*(r)})^{k-1}(W-W^{*(r)})|W_1,\ldots,W_r)]| \\ &\leq \|f_h^{(k+1)}\|\sqrt{\mathbb{E}\{\mathbb{E}((W^{*(r)})^{k-1}(W-W^{*(r)})|W_1,\ldots,W_r)^2\}},\end{aligned}$$

and thus we obtain the desired bound. □

We now demonstrate how Theorem 5.19 can be used to easily obtain Symmetric Variance-Gamma limit theorems. The bounds of Corollaries 5.20 and 5.21 are symmetric in m and n and the moments of X and Y. This is in contrast to the asymmetric bounds of Theorems 4.4, 4.13 and 4.17, which were obtained using the local coupling approach. In Section 5.4 we shall use the exchangeable pair coupling to prove a Symmetric Variance-Gamma limit theorem, and that approach leads to an asymmetric bound.

Corollary 5.20. *Let $X_1, X_2, \ldots, X_m$ be a collection of i.i.d. random variables with mean zero and unit variance. Also, let $Y_1, Y_2, \ldots, Y_n$ be a collection of i.i.d. random variables with mean zero and unit variance. Suppose further that the σ-fields $\sigma\{X_i : i = 1, \ldots, m\}$ and $\sigma\{Y_j : i = 1, \ldots, n\}$ are independent. Let $U = \sum_{i=1}^m X_i$, $V = \sum_{j=1}^n Y_j$ and set $W = \frac{1}{\sqrt{mn}} UV$. Let X and Y be variables with the same distribution as the X_i and Y_j, and suppose that $\mathbb{E}|X^3|$, $\mathbb{E}|Y^3| < \infty$. Then for $h \in C_b^2(\mathbb{R})$, we have*

$$|\mathbb{E}h(W) - \mathrm{VG}_{1,0}^{1,0}h| \leq (2N_1^2(h) + N_1^3(h))\bigg(\frac{3\mathbb{E}|X^3|}{2\sqrt{m}} + \frac{3\mathbb{E}|Y^3|}{2\sqrt{n}} + \frac{9\mathbb{E}|X^3|\mathbb{E}|Y^3|}{4\sqrt{mn}}\bigg),$$

where the $N_1^k(h)$ are defined as in Theorem 3.22.

Proof. We make use of bound (5.23) of Theorem 5.19. We just need to bound $\mathbb{E}|W - W^{*(2)}|$ and $\mathbb{E}|W(W - W^{*(2)})|$. By part (iii) of Lemma 5.13 and Lemma 5.9, we have that $W^{*(2)} = \frac{1}{\sqrt{mn}} U^* V^*$ where $U^* = U - X_I + X_I^*$ and $V^* = V - Y_J + Y_J^*$. By the independence of the collections $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ we have

$$\begin{aligned}
\mathbb{E}|W - W^{*(2)}| &= \frac{1}{\sqrt{mn}}\mathbb{E}|UV - U^*V^*| \\
&= \frac{1}{\sqrt{mn}}\mathbb{E}|(X_I - X_I^*)V + (Y_J - Y_J^*)U - (X_I - X_I^*)(Y_J - Y_J^*)| \\
&\leq \frac{1}{\sqrt{mn}}\{(\mathbb{E}|X_I| + \mathbb{E}|X_I^*|)\mathbb{E}|V| + (\mathbb{E}|Y_J| + \mathbb{E}|Y_J^*|)\mathbb{E}|U| \\
&\quad + (\mathbb{E}|X_I| + \mathbb{E}|X_I^*|)(\mathbb{E}|Y_J| + \mathbb{E}|Y_J^*|)\}.
\end{aligned}$$

By part (ii) of Lemma 5.13 we have that $\mathbb{E}|X_I| \leq \frac{1}{2}\mathbb{E}|X^3|$. Using this fact and that $\mathbb{E}|U| \leq \sqrt{m}$ and $\mathbb{E}|V| \leq \sqrt{n}$ gives

$$\begin{aligned}
\mathbb{E}|W - W^{*(2)}| &\leq \frac{1}{\sqrt{m}}(1 + \mathbb{E}|X^3|) + \frac{1}{\sqrt{n}}(1 + \mathbb{E}|Y^3|) + \frac{1}{\sqrt{mn}}(1 + \mathbb{E}|X^3|)(1 + \mathbb{E}|Y^3|) \\
&\leq \frac{3\mathbb{E}|X^3|}{2\sqrt{m}} + \frac{3\mathbb{E}|Y^3|}{2\sqrt{n}} + \frac{9\mathbb{E}|X^3|\mathbb{E}|Y^3|}{4\sqrt{mn}},
\end{aligned}$$

where the final inequality follows as $1 \leq \mathbb{E}|X^3|$. A similar calculation verifies that

$$\mathbb{E}|W(W - W^{*(2)})| \leq \frac{3\mathbb{E}|X^3|}{2\sqrt{m}} + \frac{3\mathbb{E}|Y^3|}{2\sqrt{n}} + \frac{9\mathbb{E}|X^3|\mathbb{E}|Y^3|}{4\sqrt{mn}}.$$

Combining these two bounds and using that $A_1^2 = A_2^2 = 1$ gives the result. $\square$

When $\mathbb{E}X^3 = \mathbb{E}Y^3 = 0$ we can use the bound (5.24) of Theorem 5.19 to obtain a bound on the rate of convergence of W to its limiting Variance-Gamma distribution that is of order $m^{-1}+n^{-1}$ for smooth test functions. This approach is similar to the one used by Goldstein and Reinert [29], who used a zero bias coupling approach to obtain a bound of order n^{-1}, for smooth test functions, for normal approximation under the assumption that $\mathbb{E}X^3 = 0$.

Corollary 5.21. *Let the X_i, Y_j and W be defined as in Corollary 5.20, but with the extra condition that $\mathbb{E}X^3 = \mathbb{E}Y^3 = 0$, and $\mathbb{E}|X^5|$, $\mathbb{E}|Y^5| < \infty$. Then for $h \in C_b^3(\mathbb{R})$, we have*

$$|\mathbb{E}h(W) - \mathrm{VG}_{1,0}^{1,0}h| \leq \left(\frac{1}{m} + \frac{1}{n}\right)\left\{2N_1^2(h) + \frac{85}{2}\mathbb{E}X^4\mathbb{E}Y^4N_1^3(h) + 107\mathbb{E}|X|^5\mathbb{E}|Y|^5N_1^4(h)\right\},$$

where the $N_1^k(h)$ are defined as in Theorem 3.22.

Proof. We make use of the second bound in Theorem 5.19. It suffices to bound $\sqrt{\mathbb{E}\{\mathbb{E}(W - W^{*(2)})|U,V)^2\}}$, $\sqrt{\mathbb{E}\{\mathbb{E}(W^{*(2)}(W - W^{*(2)})|U,V)^2\}}$, $\mathbb{E}|W^{*(2)}(W - W^{*(2)})^2|$ and $\mathbb{E}(W - W^{*(2)})^2$. We have

$$\begin{aligned}
\mathbb{E}(W - W^{*(2)}|U,V) &= \frac{1}{\sqrt{mn}}\mathbb{E}(UV - U^*V^*|U,V) \\
&= \frac{1}{\sqrt{mn}}\mathbb{E}((X_I - X_I^*)V + (Y_J - Y_J^*)U - (X_I - X_I^*)(Y_J - Y_J^*)|U,V) \\
&= \frac{1}{\sqrt{mn}}(V\mathbb{E}((X_I|U) + U\mathbb{E}(Y_J|V) - \mathbb{E}(X_I|U)\mathbb{E}(Y_J|V)) \\
&= \frac{1}{\sqrt{mn}}\left(\frac{1}{m} + \frac{1}{n} - \frac{1}{mn}\right)UV,
\end{aligned}$$

where we used that X_I^* and S are independent and from part (ii) of Lemma 5.13 that $\mathbb{E}X_I^* = \frac{1}{2}\mathbb{E}X^3 = 0$ to obtain the third equality, and that $\mathbb{E}(X_I|U) = \frac{1}{m}U$ to obtain the final equality. We therefore have, as $\mathbb{E}U^2 = m$ and $\mathbb{E}V^2 = n$,

$$\sqrt{\mathbb{E}\{\mathbb{E}(W - W^{*(2)})|U,V)^2\}} = \sqrt{\frac{1}{mn}\left(\frac{1}{m} + \frac{1}{n} - \frac{1}{mn}\right)^2\mathbb{E}U^2V^2} = \frac{1}{m} + \frac{1}{n} - \frac{1}{mn} < \frac{1}{m} + \frac{1}{n}.$$

We now bound the second term

$$\begin{aligned}
&\mathbb{E}(W^{*(2)}(W - W^{*(2)})|U,V) \\
&= \frac{1}{mn}\mathbb{E}(U^*V^*(UV - U^*V^*)|U,V)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{mn}\mathbb{E}((U - X_I + X_I^*)(V - Y_j + Y_J^*)((X_I - X_I^*)V \\
&\quad + (Y_J - Y_J^*)U - (X_I - X_I^*)(Y_J - Y_J^*))|U, V) \\
&= \frac{1}{mn}\Big[UV^2\mathbb{E}(X_I - X_I^*|U) + U^2V\mathbb{E}(Y_J - Y_J^*|V) - \\
&\quad - 3UV\mathbb{E}(X_I - X_I^*|U)\mathbb{E}(Y_J - Y_J^*|V) - V^2\mathbb{E}((X_I - X_I^*)^2|U) - U^2\mathbb{E}((Y_J - Y_J^*)^2|V) \\
&\quad + 2V\mathbb{E}((X_I - X_I^*)^2|U)\mathbb{E}(Y_J - Y_J^*|V) + 2U\mathbb{E}(X_I - X_I^*|U)\mathbb{E}((Y_J - Y_J^*)^2|V) \\
&\quad - \mathbb{E}((X_I - X_I^*)^2|U)\mathbb{E}((Y_J - Y_J^*)^2|V)\Big].
\end{aligned} \tag{5.26}
$$

Since X_I^* and U are independent, $\mathbb{E}(X_I|U) = \frac{1}{m}U$ and by part (ii) of Lemma 5.13 $\mathbb{E}X_I^* = 0$ and $\mathbb{E}(X_I^*)^2 = \frac{1}{3}\mathbb{E}X^4$, we have

$$
\begin{aligned}
\mathbb{E}(X_I - X_I^*|U) &= \frac{1}{m}U, \\
\mathbb{E}((X_I - X_I^*)^2|U) &= \frac{1}{m}\sum_{i=1}^m \mathbb{E}((X_i - X_i^*)^2|U) \\
&= \frac{1}{m}\sum_{i=1}^m (\mathbb{E}(X_i^2|U) + \mathbb{E}(X_i^*)^2) \\
&= \frac{1}{3}\mathbb{E}X^4 + \frac{1}{m}\sum_{i=1}^m \mathbb{E}(X_i^2|U).
\end{aligned}
$$

Substituting these formulas into (5.26) gives

$$
\begin{aligned}
&\mathbb{E}(W^{*(2)}(W - W^{*(2)})|U, V) \\
&= \frac{1}{mn}\Bigg[\bigg(\frac{1}{m} + \frac{1}{n} - \frac{3}{mn}\bigg)U^2V^2 - \bigg(1 - \frac{2}{n}\bigg)\bigg(\frac{\mathbb{E}X^4}{3} + \frac{1}{m}\sum_{i=1}^m \mathbb{E}(X_i^2|U)\bigg)V^2 \\
&\quad - \bigg(1 - \frac{2}{m}\bigg)\bigg(\frac{\mathbb{E}Y^4}{3} + \frac{1}{n}\sum_{j=1}^n \mathbb{E}(Y_j^2|V)\bigg)U^2 \\
&\quad - \bigg(\frac{\mathbb{E}X^4}{3} + \frac{1}{m}\sum_{i=1}^m \mathbb{E}(X_i^2|U)\bigg)\bigg(\frac{\mathbb{E}Y^4}{3} + \frac{1}{n}\sum_{j=1}^n \mathbb{E}(Y_j^2|V)\bigg)\Bigg].
\end{aligned}
$$

Hence, on application of the inequalities $(a+b+c+d)^2 \leq 4(a^2+b^2+c^2+d^2)$, $(a+b)^2 \leq 2(a^2+b^2)$,

$|1-\frac{2}{m}|\leq 1$ and $|1-\frac{2}{n}|\leq 1$, we have

$$\begin{aligned}
&\sqrt{\mathbb{E}\{\mathbb{E}(W^{*(2)}(W-W^{*(2)})|U,V)^2\}} \\
&= 2\bigg\{\frac{1}{(mn)^2}\bigg[\bigg(\frac{1}{m}+\frac{1}{n}-\frac{3}{mn}\bigg)^2\mathbb{E}U^4\mathbb{E}V^4 \\
&\quad + 2\mathbb{E}V^4\bigg(\frac{\mathbb{E}(X^4)^2}{9}+\frac{1}{m^2}\bigg(\sum_{i=1}^m\mathbb{E}[(\mathbb{E}(X_i^2|U))^2]+\sum_{i\neq j}\mathbb{E}(\mathbb{E}X_i^2|U))\mathbb{E}(\mathbb{E}X_j^2|U))\bigg)\bigg) \\
&\quad + 2\mathbb{E}U^4\bigg(\frac{\mathbb{E}(Y^4)^2}{9}+\frac{1}{n^2}\bigg(\sum_{j=1}^n\mathbb{E}[(\mathbb{E}(Y_j^2|V))^2]+\sum_{i\neq j}\mathbb{E}(\mathbb{E}Y_i^2|V))\mathbb{E}(\mathbb{E}Y_j^2|V))\bigg)\bigg) \\
&\quad + 4\bigg(\frac{\mathbb{E}(X^4)^2}{9}+\frac{1}{m^2}\bigg(\sum_{i=1}^m\mathbb{E}[(\mathbb{E}(X_i^2|U))^2]+\sum_{i\neq j}\mathbb{E}(\mathbb{E}X_i^2|U))\mathbb{E}(\mathbb{E}X_j^2|U))\bigg)\bigg) \\
&\quad \times\bigg(\frac{\mathbb{E}(Y^4)^2}{9}+\frac{1}{n^2}\bigg(\sum_{j=1}^n\mathbb{E}[(\mathbb{E}(Y_j^2|V))^2]+\sum_{i\neq j}\mathbb{E}(\mathbb{E}Y_i^2|V))\mathbb{E}(\mathbb{E}Y_j^2|V))\bigg)\bigg)\bigg]\bigg\}^{1/2} \\
&\leq 2\bigg\{\frac{1}{(mn)^2}\bigg[\bigg(\frac{1}{m}+\frac{1}{n}-\frac{3}{mn}\bigg)^2\mathbb{E}U^4\mathbb{E}V^4+2\mathbb{E}V^4\bigg(\frac{\mathbb{E}(X^4)^2}{9}+\frac{1}{m^2}\Big(m\mathbb{E}X^4+m(m-1)\Big)\bigg) \\
&\quad + 2\mathbb{E}U^4\bigg(\frac{\mathbb{E}(Y^4)^2}{9}+\frac{1}{n^2}\Big(n\mathbb{E}Y^4+n(n-1)\Big)\bigg)+4\bigg(\frac{\mathbb{E}(X^4)^2}{9}+\frac{1}{m^2}\Big(m\mathbb{E}X^4+m(m-1)\Big)\bigg) \\
&\quad \times\bigg(\frac{\mathbb{E}(Y^4)^2}{9}+\frac{1}{n^2}\Big(n\mathbb{E}Y^4+n(n-1)\Big)\bigg)\bigg]\bigg\}^{1/2}.
\end{aligned}$$

Here to obtain the inequality we used that $\mathbb{E}(\mathbb{E}(X_i^2|U))=\mathbb{E}X_i^2=1$ and that, by Jensen's inequality,

$$\mathbb{E}[(\mathbb{E}(X_i^2|U))^2]\leq\mathbb{E}(\mathbb{E}(X_i^4|U))=\mathbb{E}X^4.$$

Since $EU^4<3m^2+m\mathbb{E}X^4$ and $\mathbb{E}V^4<3n^2+n\mathbb{E}Y^4$ (see Lemma 4.6), we have

$$\begin{aligned}
&\sqrt{\mathbb{E}\{\mathbb{E}(W^{*(2)}(W-W^{*(2)})|U,V)^2\}} \\
&< 2\sqrt{2}\bigg\{\bigg(\frac{1}{m^2}+\frac{1}{n^2}\bigg)\bigg(3+\frac{\mathbb{E}X^4}{m}\bigg)\bigg(3+\frac{\mathbb{E}Y^4}{n}\bigg)+\frac{1}{m^2}\bigg(3+\frac{\mathbb{E}Y^4}{n}\bigg)\bigg(\frac{(\mathbb{E}X^4)^2}{9}+1+\frac{\mathbb{E}X^4}{m}\bigg) \\
&\quad + \frac{1}{n^2}\bigg(3+\frac{\mathbb{E}X^4}{m}\bigg)\bigg(\frac{(\mathbb{E}Y^4)^2}{9}+1+\frac{\mathbb{E}Y^4}{n}\bigg) \\
&\quad + \frac{2}{(mn)^2}\bigg(\frac{(\mathbb{E}X^4)^2}{9}+1+\frac{\mathbb{E}X^4}{m}\bigg)\bigg(\frac{(\mathbb{E}Y^4)^2}{9}+1+\frac{\mathbb{E}Y^4}{n}\bigg)\bigg\}^{1/2}.
\end{aligned}$$

We can simplify this bound by using that $\mathbb{E}X^4\geq 1$ and $\mathbb{E}Y^4\geq 1$, as well as the simple inequality

$\frac{2}{(mn)^2} \leq \frac{1}{m^2} + \frac{1}{n^2} \leq (\frac{1}{m} + \frac{1}{n})^2$, for $m, n \geq 1$. This gives

$$\begin{aligned}\sqrt{\mathbb{E}\{\mathbb{E}(W_2^*(W-W_2^*)|U,V)^2\}} &< 2\sqrt{2}\sqrt{16+44+121}\bigg(\frac{1}{m}+\frac{1}{n}\bigg)\mathbb{E}X^4\mathbb{E}Y^4 \\ &< 39\bigg(\frac{1}{m}+\frac{1}{n}\bigg)\mathbb{E}X^4\mathbb{E}Y^4.\end{aligned}$$

We now bound the third term

$$\begin{aligned}&\mathbb{E}|W^{*(2)}(W-W^{*(2)})^2| \\ &= \frac{1}{(mn)^{3/2}}\mathbb{E}|U^*V^*(UV-U^*V^*)^2| \\ &= \frac{1}{(mn)^{3/2}}\mathbb{E}|(U-X_I+X_I^*)(V-Y_J+Y_J^*)((X_I-X_I^*)V \\ &\quad+(Y_J-Y_J^*)U-(X_I-X_I^*)(Y_J-Y_J^*))^2| \\ &\leq \frac{3}{(mn)^{3/2}}\mathbb{E}|(U-X_I+X_I^*)(V-Y_J+Y_J^*)((X_I-X_I^*)^2V^2 \\ &\quad+(Y_J-Y_J^*)^2U^2+(X_I-X_I^*)^2(Y_J-Y_J^*)^2)| \\ &\leq \frac{3}{(mn)^{3/2}}\Big\{\mathbb{E}|(X_I-X_I^*)^2U|\mathbb{E}|V^3|+\mathbb{E}|(Y_J-Y_J^*)^2V|\mathbb{E}|U^3|+\mathbb{E}|X_I-X_I^*|^3\mathbb{E}|V^3| \\ &\quad+\mathbb{E}|Y_J-Y_J^*|^3\mathbb{E}|U^3|+\mathbb{E}|(X_I-X_I^*)^2U|\mathbb{E}|(Y_J-Y_J^*)V^2|+\mathbb{E}|(X_I-X_I^*)U^2|\mathbb{E}|(Y_J-Y_J^*)^2V| \\ &\quad+\mathbb{E}|(X_I-X_I^*)^2U|\mathbb{E}|(Y_J-Y_J^*)^2V|+\mathbb{E}|X_I-X_I^*|^3\mathbb{E}|(Y_J-Y_J^*)V^2| \\ &\quad+\mathbb{E}|(X_I-X_I^*)U^2|\mathbb{E}|Y_J-Y_J^*|^3+\mathbb{E}|X_I-X_I^*|^3\mathbb{E}|(Y_J-Y_J^*)^2V| \\ &\quad+\mathbb{E}|(X_I-X_I^*)^2U|\mathbb{E}|Y_J-Y_J^*|^3+\mathbb{E}|X_I-X_I^*|^3\mathbb{E}|Y_J-Y_J^*|^3\Big\}.\end{aligned}$$

Using part (ii) of Lemma 5.13 gives

$$\mathbb{E}(X_I-X_I^*)^2 = \mathbb{E}X_I^2+\mathbb{E}(X_I^*)^2 = 1+\frac{1}{3}\mathbb{E}X^4 \leq \frac{4}{3}\mathbb{E}X^4, \tag{5.27}$$

$$\mathbb{E}|X_I-X_I^*|^3 \leq 4\mathbb{E}|X_I^3|+4\mathbb{E}|X_I^*|^3 = 4\mathbb{E}|X^3|+\mathbb{E}|X^5| \leq 5\mathbb{E}|X^5|. \tag{5.28}$$

By the Cauchy–Schwarz inequality and (5.27) we have

$$\mathbb{E}|(X_I-X_I^*)U^2| \leq \sqrt{\mathbb{E}(X_I-X_I^*)^2\mathbb{E}U^4} \leq m\sqrt{\frac{4}{3}\mathbb{E}X^4\bigg(3+\frac{\mathbb{E}X^4}{m}\bigg)} < \frac{4m}{\sqrt{3}}\mathbb{E}X^4, \tag{5.29}$$

and we again used that $m \geq 1$ and $\mathbb{E}X^4 \geq 1$. By Hölder's inequality and inequality (5.28) we have

$$\begin{aligned}\mathbb{E}|(X_I - X_I^*)^2U| &\leq \{\mathbb{E}|X_I - X_I^*|^3\}^{2/3}\{\mathbb{E}|U|^3\}^{1/3} \\ &\leq \{\mathbb{E}|X_I - X_I^*|^3\}^{2/3}\{\mathbb{E}U^4\}^{1/4} \\ &< \sqrt{m}(5\mathbb{E}|X|^5)^{2/3}\bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{1/4}.\end{aligned}$$

Using that $m \geq 1$ and $1 \leq \mathbb{E}X^4 \leq \mathbb{E}|X|^5$ we can simplify this bound as follows

$$\mathbb{E}|(X_I - X_I^*)^2U| \leq 5^{2/3} \cdot 4^{1/4} \cdot \sqrt{m}\{\mathbb{E}|X|^5\}^{2/3}\{\mathbb{E}|X|^5\}^{1/4} < 5\sqrt{m}\mathbb{E}|X|^5. \tag{5.30}$$

Applying the inequalities (5.28), (5.29) and (5.30) and using that $\mathbb{E}|U^3| \leq (\mathbb{E}U^4)^{3/4}$, by Hölder's inequality, then gives

$$\begin{aligned}&\mathbb{E}|W^{*(2)}(W - W^{*(2)})^2| \\ &< 3\bigg\{\frac{5}{m}\bigg(1 + \frac{1}{\sqrt{m}}\bigg)\mathbb{E}|X|^5\bigg(3 + \frac{\mathbb{E}Y^4}{n}\bigg)^{3/4} + \frac{5}{n}\bigg(1 + \frac{1}{\sqrt{n}}\bigg)\mathbb{E}|Y|^5\bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4} \\ &\quad + \frac{15}{m\sqrt{n}}\bigg(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}}\bigg)\mathbb{E}|X|^5\mathbb{E}Y^4 + \frac{15}{n\sqrt{m}}\bigg(\frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}}\bigg)\mathbb{E}X^4\mathbb{E}|Y|^5 \\ &\quad + \bigg(\frac{25}{mn} + \frac{25}{m^{3/2}n} + \frac{25}{mn^{3/2}} + \frac{9}{(mn)^{3/2}}\bigg)\mathbb{E}|X|^5\mathbb{E}|Y|^5\bigg\}.\end{aligned}$$

We can simplify this bound by using that $1 \leq \mathbb{E}X^4 \leq \mathbb{E}|X^5|$ and $1 \leq \mathbb{E}Y^4 \leq \mathbb{E}|Y^5|$, as well as the simple inequality $\frac{2}{mn} \leq \frac{1}{m} + \frac{1}{n}$, for $m, n \geq 1$. This gives

$$\begin{aligned}\mathbb{E}|W^{*(2)}(W - W^{*(2)})^2| &< 3\bigg\{5 \cdot 4^{3/4} + 15 + \frac{25 + 25 + 25 + 9}{2}\bigg\}\bigg(\frac{1}{m} + \frac{1}{n}\bigg)\mathbb{E}|X|^5\mathbb{E}|Y|^5 \\ &< 214\bigg(\frac{1}{m} + \frac{1}{n}\bigg)\mathbb{E}|X|^5\mathbb{E}|Y|^5.\end{aligned}$$

Finally, we bound the fourth term.

$$\begin{aligned}\mathbb{E}(W - W^{*(2)})^2 &= \frac{1}{mn}\mathbb{E}(UV - U^*V^*)^2 \\ &= \frac{1}{mn}\mathbb{E}((X_I - X_I^*)V + (Y_J - Y_J^*)U - (X_I - X_I^*)(Y_J - Y_J^*))^2 \\ &\leq \frac{3}{mn}[\mathbb{E}(X_I - X_I^*)^2\mathbb{E}V^2 + \mathbb{E}(Y_J - Y_J^*)^2\mathbb{E}U^2 \\ &\quad + \mathbb{E}(X_I - X_I^*)^2\mathbb{E}(Y_J - Y_J^*)^2],\end{aligned}$$

where we used that $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$ to obtain the inequality. Using that $\mathbb{E}U^2 = m$,

$\mathbb{E}V^2 = n$ and inequality (5.27) gives

$$\begin{aligned}
\mathbb{E}(W - W^{*(2)})^2 &= \frac{4\mathbb{E}X^4}{m} + \frac{4\mathbb{E}Y^4}{n} + \frac{16\mathbb{E}X^4\mathbb{E}Y^4}{3mn} \\
&\leq \left(4 + \frac{8}{3}\right)\left(\frac{1}{m} + \frac{1}{n}\right)\mathbb{E}X^4\mathbb{E}Y^4 \\
&< 7\left(\frac{1}{m} + \frac{1}{n}\right)\mathbb{E}X^4\mathbb{E}Y^4,
\end{aligned}$$

and we used our usual argument to obtain the first inequality. We have obtained bounds of order $m^{-1} + n^{-1}$ for all the terms, and so the proof is complete. □

We end this section with some remarks.

Remark 5.22. In Corollary 5.21 we were able to obtain an order $m^{-1} + n^{-1}$ bound on the convergence rate similar to that obtained in Theorem 4.13, but with a stronger assumption that $\mathbb{E}X^3 = \mathbb{E}Y^3 = 0$. The reason that we required that $\mathbb{E}X^3 = \mathbb{E}Y^3 = 0$ to achieve the faster convergence rate was because starting from Theorem 5.19 we were unable to use symmetry arguments since we had already taken the supremum norms of the derivatives of the solution to the VG$(1, 0, 1, 0)$ Stein equation (3.14).

Remark 5.23. If we could establish bounds on the required low order derivatives of the Normal Product Stein equation (5.8) for $r \geq 3$, then we could apply Theorem 5.19 and calculations similar to those used in the proof of Corollaries 5.20 and 5.21 to easily obtain bounds on the error in approximating the statistic $S_r = \prod_{k=1}^{r}(\frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k} X_{ik})$ by its limiting Product Normal distribution. This is in contrast to the local approach coupling, for which the calculations, involving Taylor expansions that would be used to bound the expectation (5.25), would get very tedious for large r.

5.4 Exchangeable Pair approach bounds for Symmetric Variance-Gamma approximation

The exchangeable pair coupling for use in normal approximation by Stein's method were introduced by Stein [73]. This coupling is most useful when we deal with global but weak dependence between random variables. The exchangeable pair approach has also proved useful in non–normal contexts, see Chatterjee et al. [16] and Röllin [65]. Recently, Chatterjee and Meckes [17] and Reinert and Röllin [61] have been able to extend the approach to multivariate normal approximation. In this section we illustrate how exchangeable pair couplings can be used to prove limit theorems for Symmetric Variance-Gamma distributions.

We begin with a definition.

Definition 5.24. *A pair (W, W') of random variables defined on the same probability space is called* exchangeable *if for all measurable sets B and B',*

$$\mathbb{P}(W \in B, W' \in B') = \mathbb{P}(W \in B', W' \in B).$$

Following Stein [73] we assume that $\mathbb{E}W = 0$, $\mathrm{Var}W = 1$, and that there is a $0 < \lambda < 1$ such that

$$\mathbb{E}(W|W') = (1-\lambda)W. \tag{5.31}$$

Notice that if the pair (W, W') has a bivariate normal distribution with correlation ρ, then (5.31) is satisfied with $\lambda = 1 - \rho$. Therefore, heuristically, (5.31) can be understood as a linear regression condition (see Reinert and Röllin [61] for more details). Under (5.31) it is easy to see that

$$\mathbb{E}Wf(W) = \frac{1}{2\lambda}\mathbb{E}(W - W')(f(W) - f(W')). \tag{5.32}$$

Reinert and Röllin [61] introduced a multivariate generalisation of condition (5.31). According to Reinert and Röllin we assume that $(\mathbf{W}, \mathbf{W}')$ is an exchangeable pair of $\mathbb{R}^d$–valued random vectors such that $\mathbb{E}\mathbf{W} = \mathbf{0}$ and $\mathbb{E}\mathbf{W}\mathbf{W}^T = \Sigma$, and that there is an invertible $d \times d$ matrix Λ such that

$$\mathbb{E}^{\mathbf{W}}(\mathbf{W}' - \mathbf{W}) = -\Lambda\mathbf{W}. \tag{5.33}$$

We shall make use of condition (5.33) in the proof of Theorem 6.24. However, for now we restrict our attention to the one-dimensional condition (5.31) and demonstrate how the exchangeable pair approach can be used to obtain limit theorems for Symmetric Variance-Gamma distributions.

Theorem 5.25. *Let (S, S') and (T, T') be exchangeable pairs satisfying (5.32) and assume $\mathbb{E}S = \mathbb{E}T = 0$, $\mathrm{Var}S = \mathrm{Var}T = 1$. Define $W = ST$. Then for $h \in C_b^2(\mathbb{R})$, we have*

$$\begin{aligned}|\mathbb{E}h(W) - \mathrm{VG}^{1,0}_{1,0}h| \leq{}& (N_1^1(h) + N_1^2(h))\sqrt{\mathbb{E}\left(1 - \frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\right)^2} \\ &+ \frac{N_1^2(h)}{4\lambda}\mathbb{E}|S-S'|^3\mathbb{E}|T^3| + \frac{N_1^2(h)}{4\lambda\mu}\mathbb{E}|S(S-S')^2|\mathbb{E}|T-T'|^3 \\ &+ \frac{N_1^3(h)}{8\lambda\mu}\mathbb{E}S^2(S-S')^2\mathbb{E}|T(T-T')^3|,\end{aligned} \tag{5.34}$$

where the $N_1^k(h)$ are defined as in Theorem 3.22.

Proof. Conditioning on T and using (5.32) gives

$$\mathbb{E}Wf(W) = \mathbb{E}STf(ST) = \mathbb{E}(\mathbb{E}(STf(ST)|T)) = \frac{1}{2\lambda}\mathbb{E}T(S-S')(f(ST)-f(S'T)).$$

Taylor expanding $f(S'T)$ about ST gives

$$\mathbb{E}Wf(W) = \frac{1}{2\lambda}\mathbb{E}(S-S')^2T^2f'(ST) + R_1,$$

where

$$|R_1| \leq \frac{\|f''\|}{4\lambda}\mathbb{E}|(S-S')^3T^3| = \frac{\|f''\|}{4\lambda}\mathbb{E}|S-S'|^3\mathbb{E}|T|^3.$$

We now condition on S and S', then use (5.32) to obtain

$$\begin{aligned}\mathbb{E}Wf(W) &= \frac{1}{2\lambda}\mathbb{E}(\mathbb{E}(S-S')^2T^2f'(ST)|S,S')) + R_1 \\ &= \frac{1}{2\lambda}\mathbb{E}(\mathbb{E}(S-S')^2T\cdot Tf'(ST)|S,S')) + R_1 \\ &= \frac{1}{4\lambda\mu}\mathbb{E}(S-S')^2(T-T')(Tf'(ST)-T'f(ST')) + R_1.\end{aligned}$$

Taylor expanding $f'(ST')$ about ST gives

$$\begin{aligned}\mathbb{E}Wf(W) &= \frac{1}{4\lambda\mu}\mathbb{E}(S-S')^2(T-T')[(T-T')f'(ST) + ST'(T-T')f''(ST)] + R_1 + R_2 \\ &= \frac{1}{4\lambda\mu}\mathbb{E}(S-S')^2(T-T')^2f'(W) + \frac{1}{4\lambda\mu}\mathbb{E}(S-S')^2(T-T')^2Wf''(W) \\ &\quad + R_1 + R_2 + R_3,\end{aligned}$$

where

$$\begin{aligned}|R_2| &\leq \frac{\|f^{(3)}\|}{8\lambda\mu}\mathbb{E}|(S-S')^2(T-T')^3S^2T'| = \frac{\|f^{(3)}\|}{8\lambda\mu}\mathbb{E}S^2(S-S')^2\mathbb{E}|T(T-T')^3|, \\ |R_3| &= \frac{1}{4\lambda\mu}|\mathbb{E}(S-S')^2(T-T')^3Sf''(W)| \leq \frac{\|f''\|}{4\lambda\mu}\mathbb{E}|S(S-S')^2|\mathbb{E}|T-T'|^3.\end{aligned}$$

Therefore

$$\begin{aligned}|\mathbb{E}\{Wf''(W) + f'(W) - Wf(W)\}| &\leq \left|\mathbb{E}f'(W)\left(1 - \frac{1}{4\lambda\mu}(S-S')^2(T-T')^2\right)\right| \\ &\quad + \left|\mathbb{E}Wf''(W)\left(1 - \frac{1}{4\lambda\mu}(S-S')^2(T-T')^2\right)\right| \\ &\quad + |R_1| + |R_2| + |R_3|. \qquad (5.35)\end{aligned}$$

Conditioning on S and T allows us to bound the first term of (5.35) as follows

$$\begin{aligned}
&\left|\mathbb{E}f'(W)\left(1-\frac{1}{4\lambda\mu}(S-S')^2(T-T')^2\right)\right| \\
&=\left|\mathbb{E}\left(\mathbb{E}\left(f'(W)\left(1-\frac{1}{4\lambda\mu}(S-S')^2(T-T')^2\right)\middle|S,T\right)\right)\right| \\
&=\left|\mathbb{E}\left[f'(W)\left(1-\frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\right)\right]\right| \\
&\leq \|f'\|\sqrt{\mathbb{E}\left(1-\frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\right)^2},
\end{aligned}$$

where we used the Cauchy-Schwarz inequality to obtain the final inequality. The second term of (5.35) can be bounded in a similar manner:

$$\begin{aligned}
&\left|\mathbb{E}Wf''(W)\left(1-\frac{1}{4\lambda\mu}(S-S')^2(T-T')^2\right)\right| \\
&=\left|\mathbb{E}\left(\mathbb{E}\left(Wf''(W)\left(1-\frac{1}{4\lambda\mu}(S-S')^2(T-T')^2\right)\middle|S,T\right)\right)\right| \\
&=\left|\mathbb{E}\left[Wf''(W)\left(1-\frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\right)\right]\right| \\
&\leq \|f''\|(\mathbb{E}W^2)^{1/2}\sqrt{\mathbb{E}\left(1-\frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\right)^2} \\
&= \|f''\|\sqrt{\mathbb{E}\left(1-\frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\right)^2},
\end{aligned}$$

where we used the Cauchy-Schwarz inequality to obtain the final inequality. Summing the remainder terms completes the proof. □

The following corollary provides a simple example of how Theorem 5.25 can be used to proved Symmetric Variance-Gamma limit theorems.

Corollary 5.26. *Let $X_1,\ldots,X_m$ be a sequence of i.i.d. random variables with zero mean and unit variance. Similarly, let $Y_1,\ldots,X_n$ be a sequence of i.i.d random variables with zero mean and unit variance. Suppose further that the σ-fields $\sigma\{X_i : i=1,\ldots,m\}$ and $\sigma\{X_i : i=1,\ldots,m\}$ are independent. Define $S=\frac{1}{\sqrt{m}}\sum_{i=1}^m X_i$, $T=\frac{1}{\sqrt{n}}\sum_{j=1}^n Y_j$ and set $W=ST$. Then* *for $h \in C_b^2(\mathbb{R})$, we have*

$$|\mathbb{E}h(W) - \mathrm{VG}_{1,0}^{1,0}h| \leq \frac{N_1^1(h)+N_1^2(h)}{4}\sqrt{\frac{4}{m}(\mathbb{E}X^4-1)+\frac{4}{n}(\mathbb{E}Y^4-1)+\frac{1}{mn}(\mathbb{E}X^4-1)(\mathbb{E}Y^4-1)}$$
$$+\frac{2N_1^2(h)}{\sqrt{m}}\mathbb{E}|X^3|\left(3+\frac{\mathbb{E}Y^4}{n}\right)^{3/4}+\frac{2N_1^2(h)}{\sqrt{n}}\mathbb{E}|Y^3|(3+\sqrt{\mathbb{E}X^4})$$
$$+\frac{N_1^3(h)}{2\sqrt{n}}(3\mathbb{E}|Y^3|+\mathbb{E}Y^4)\left(1+\sqrt{\mathbb{E}X^4\left(3+\frac{\mathbb{E}X^4}{m}\right)}\right),$$

where the $N_1^k(h)$ are defined as per Theorem 3.22.

Proof. Following Reinert [58], to construct S' such that (S,S') is an exchangeable pair, pick an index I uniformly from $\{1,\ldots,m\}$. If $I=i$, we replace $\frac{1}{\sqrt{m}}X_i$ by an independent copy $\frac{1}{\sqrt{m}}X_i^*$, and we put

$$S' = S - \frac{1}{\sqrt{m}}X_I + \frac{1}{\sqrt{m}}X_I^*.$$

Then (S,S') is exchangeable, and

$$\mathbb{E}(S'|S) = S - \frac{1}{m}S + \mathbb{E}X_I^* = \left(1-\frac{1}{m}\right)S,$$

so (5.31) is satisfied with $\lambda = \frac{1}{m}$. To construct T such that (T,T') is an exchangeable pair we proceed in a similar way; we pick and index J from $\{1,\ldots,n\}$ and put $T' = T - \frac{1}{\sqrt{n}}Y_J + \frac{1}{\sqrt{n}}Y_J^*$. We can also verify that (5.31) holds with $\mu = \frac{1}{n}$. We now set about bounding the expectations in the bound (5.34). For ease of notation we denote the four expectation terms of the bound (5.34) by $\tilde{R}_k$, for $k=1,2,3,4$. We begin by bounding the expectation of the first term ($\tilde{R_1}$) of (5.34). We now calculate $\frac{1}{\lambda}\mathbb{E}((S-S')^2|S)$. As $\lambda=\frac{1}{m}$, we have

$$\begin{aligned}\frac{1}{\lambda}\mathbb{E}((S-S')^2|S) &= \mathbb{E}((X_I-X_I^*)^2|S) \\ &= \frac{1}{m}\sum_{i=1}^m \mathbb{E}((X_i-X_i^*)^2|S) \\ &= \frac{1}{m}\sum_{i=1}^m (\mathbb{E}(X_i^2|S)+\mathbb{E}(X_i^*)^2) \\ &= 1+\frac{1}{m}\sum_{i=1}^m \mathbb{E}(X_i^2|S).\end{aligned}$$

With this we can write $\tilde{R}_1$ as

$$\begin{aligned}
\tilde{R}_1 &= \sqrt{\mathbb{E}\bigg(1 - \frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\bigg)^2} \\
&= \sqrt{\mathbb{E}\bigg(1 - \frac{1}{4}\bigg(1 + \frac{1}{m}\sum_{i=1}^m \mathbb{E}(X_i^2|S)\bigg)\bigg(1 + \frac{1}{n}\sum_{j=1}^n \mathbb{E}(Y_j^2|T)\bigg)\bigg)^2} \\
&= \frac{1}{4}\bigg\{\frac{4}{m}(\mathbb{E}[(\mathbb{E}(X_i^2|S))^2] - 1) + \frac{4}{n}(\mathbb{E}[(\mathbb{E}(Y_j^2|T))^2] - 1) \\
&\quad + \frac{1}{mn}(\mathbb{E}[(\mathbb{E}(X_i^2|S))^2] - 1)(\mathbb{E}[(\mathbb{E}(Y_j^2|T))^2] - 1)\bigg\}^{1/2},
\end{aligned}$$

where the final equality follows from a straightforward direct calculation, in which we made use of the fact that the X_i are identically distributed and that $\mathbb{E}(\mathbb{E}(X_i^2|S)) = \mathbb{E}X_i^2 = 1$. By Jensen's inequality we have

$$\mathbb{E}[(\mathbb{E}(X_i^2|S))^2] \le \mathbb{E}(\mathbb{E}(X_i^4|S)) = \mathbb{E}X^4,$$

and so we may bound $\tilde{R}_1$ as follows

$$\tilde{R}_1 \le \sqrt{\frac{4}{m}(\mathbb{E}X^4 - 1) + \frac{4}{n}(\mathbb{E}Y^4 - 1) + \frac{1}{mn}(\mathbb{E}X^4 - 1)(\mathbb{E}Y^4 - 1)}.$$

We now bound $\tilde{R}_2$:

$$\begin{aligned}
\tilde{R}_2 &= \frac{1}{4\lambda}\mathbb{E}|S-S'|^3\mathbb{E}|T^3| \\
&= \frac{1}{4\sqrt{m}}\mathbb{E}|X_I - X_I^*|^3\mathbb{E}|T^3| \\
&\le \frac{1}{\sqrt{m}}(\mathbb{E}|X_I|^3 + \mathbb{E}|X_I^*|^3)(\mathbb{E}T^4)^{3/4} \\
&\le \frac{2\mathbb{E}|X^3|}{\sqrt{m}}\bigg(3 + \frac{\mathbb{E}Y^4}{n}\bigg)^{3/4}.
\end{aligned}$$

To simplify the calculation of the bound for $\tilde{R}_3$ we use that $\mathbb{E}|Y_J - Y_J^*|^3 \le 8\mathbb{E}|Y^3|$, as we did in obtaining our bound for $\tilde{R}_2$,

$$\begin{aligned}
\tilde{R}_3 &= \frac{1}{4\lambda\mu}\mathbb{E}|S(S-S')^2|\mathbb{E}|T-T'|^3 \\
&= \frac{1}{4\sqrt{n}}\mathbb{E}|S(X_I - X_I^*)^2|\mathbb{E}|Y_J - Y_J^*|^3 \\
&\le \frac{2\mathbb{E}|Y^3|}{\sqrt{n}}(\mathbb{E}|SX_I^2| + 2\mathbb{E}|X_I^*|\mathbb{E}|SX_I^*| + \mathbb{E}(X_I^*)^2\mathbb{E}|S|)
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{2\mathbb{E}|Y^3|}{\sqrt{n}}(\sqrt{\mathbb{E}S^2\mathbb{E}X_I^4} + 2\sqrt{\mathbb{E}S^2\mathbb{E}X_I^2} + \sqrt{\mathbb{E}S^2}) \\
&= \frac{2\mathbb{E}|Y^3|}{\sqrt{n}}(\sqrt{\mathbb{E}X^4} + 3).
\end{aligned}$$

We now bound $\tilde{R}_4$. Let $T_J = T - \frac{1}{\sqrt{n}}Y_J$. Then Y_J and T_J are independent. We bound $\tilde{R}_4$ as follows

$$\tilde{R}_4 = \frac{1}{8\lambda\mu}\mathbb{E}S^2(S-S')^2\mathbb{E}|T(T-T')^3| = \frac{1}{8\sqrt{n}}\mathbb{E}S^2(X_I - X_I^*)^2\mathbb{E}|T(Y_J - Y_J^*)^3|$$

Finally, we bound $\mathbb{E}S^2(X_I - X_I^*)^2$ and $\mathbb{E}|T(Y_J - Y_J^*)^3|$:

$$\begin{aligned}
\mathbb{E}S^2(X_I - X_I^*)^2 &= \mathbb{E}(S^2(X_I^2 - 2X_IX_I^* + (X_I^*)^2) \\
&= \mathbb{E}S^2X_I^2 + \mathbb{E}(X_I^*)^2\mathbb{E}S^2 \\
&\leq \sqrt{\mathbb{E}X_I^4\mathbb{E}S^4} + 1) \\
&= \sqrt{\mathbb{E}X^4\left(3 + \frac{\mathbb{E}X^4}{m}\right)} + 1.
\end{aligned}$$

We also have

$$\begin{aligned}
\mathbb{E}|T(Y_J - Y_J^*)^3| &\leq 4\mathbb{E}|T_J + Y_J|(|Y_J|^3 + |Y_J^*|^3) \\
&\leq 4(\mathbb{E}|T_J|\mathbb{E}|Y_J^3| + \mathbb{E}|T_J|\mathbb{E}|Y_J^*|^3 + \mathbb{E}Y_J^4 + \mathbb{E}|Y_J|\mathbb{E}|Y_J^*|^3) \\
&< 4(3\mathbb{E}|Y^3| + \mathbb{E}Y^4),
\end{aligned}$$

where we used that $\mathbb{E}|T_J| \leq \sqrt{\mathbb{E}T_J^2} = \sqrt{\frac{n-1}{n}} < 1$ to obtain the final inequality. Summing up the remainders gives us bound (5.34). □

Moreover, the exchangeable pair approach is not restricted to requiring that condition (5.31) is satisfied. Following Rinott and Rotar [63], assume that (W, W') is an exchangeable pair such that $\mathbb{E}W = 0$, $\mathrm{Var}W = 1$, and let $R = R(W)$ be such that

$$\mathbb{E}(W|W') = (1-\lambda)W + R \tag{5.36}$$

for some $0 < \lambda < 1$. Similarly for Theorem 5.25, we can show that under the above setting, for W real–valued:

Theorem 5.27. *Let (S,S') and (T,T') be exchangeable pairs satisfying (5.36) and assume* $\mathbb{E}S=\mathbb{E}T=0$, $\mathrm{Var}S=\mathrm{Var}T=1$. *Define $W=ST$. Then for $h\in C_b^2(\mathbb{R})$, we have*

$$\begin{aligned}|\mathbb{E}h(W)-\mathrm{VG}^{1,0}_{1,0}h| &\le (N_1^1(h)+N_1^2(h))\sqrt{\mathbb{E}\Big(1-\frac{1}{4\lambda\mu}\mathbb{E}((S-S')^2|S)\mathbb{E}((T-T')^2|T)\Big)^2} \\ &\quad+\frac{N_1^2(h)}{4\lambda}\mathbb{E}|S-S'|^3\mathbb{E}|T^3|+\frac{N_1^2(h)}{4\lambda\mu}\mathbb{E}|S(S-S')^2|\mathbb{E}|T-T'|^3 \\ &\quad+\frac{N_1^3(h)}{8\lambda\mu}\mathbb{E}S^2(S-S')^2\mathbb{E}|T(T-T')^3|+\frac{N_1^0(h)}{\lambda}\sqrt{\mathbb{E}R^2} \\ &\quad+\frac{N_1^1(h)}{2\lambda\mu}\sqrt{\mathbb{E}(S-S')^4}(\mathbb{E}T^4)^{1/4}(\mathbb{E}R^4)^{1/4}, \qquad (5.37)\end{aligned}$$

where the $N_1^k(h)$ are defined as per Theorem 3.22.

Proof. Under condition (5.36) it is easy to see that

$$\mathbb{E}Sf(S)=\frac{1}{2\lambda}\mathbb{E}(S-S')(f(S)-f(S'))+\frac{1}{\lambda}\mathbb{E}Rf(W), \qquad (5.38)$$

with a similar expression holding for T. We now proceed as in the proof of Theorem 5.25; Taylor expanding and making use of (5.38), rather than (5.32). Doing so introduces two extra remainder terms, which are

$$\begin{aligned}R_4 &= \frac{1}{\lambda}\mathbb{E}TRf(W), \\ R_5 &= \frac{1}{2\lambda\mu}\mathbb{E}(S-S')^2TRf'(W).\end{aligned}$$

The Cauchy-Schwarz inequality is used to bound R_4 and R_5, which yields the bound (5.37). □

Thus, if R is small the $\mathrm{VG}(1,0,1,0)$ approximation will still be good.

Chapter 6

Convergence to real-valued functions of the Multivariate Normal distribution

Let $(X_{ik})_{1\leq i\leq n_k}$, for $1 \leq k \leq d$, be collections of random variables with zero mean. Define $W_k = \frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k} X_{ik}$ and let $\mathbf{W}$ denote the vector $(W_1, \ldots, W_d)$. In the case that $d = 1$, we shall suppress the k subscripts, for instance we would write $W = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i$. Suppose that $\mathbf{W}$ converges in distribution to a multivariate normal random variable $\mathbf{Z}$ with mean $\mathbf{0}$ and $d \times d$ covariance matrix Σ. Suppose that $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is continuous, then a standard result is that $g(\mathbf{W})$ converges in distribution to $g(\mathbf{Z})$. In this chapter we see how Stein's method can be used to obtain a bound on the error, in a weak convergence setting, in approximating $g(\mathbf{W})$ by its limiting distribution. The Variance-Gamma and Product Normal limit theorems considered in Chapters 4 and 5 are special cases of this problem. For example, the statistic (4.2) is of the form $g(\mathbf{W})$, where $g(\mathbf{w}) = \sum_{k=1}^{r} w_k w_{r+k}$.

The approach we consider involves the multivariate normal Stein equation with test function $h(g(\mathbf{w}))$:

$$\nabla^T\Sigma\nabla f(\mathbf{w}) - \mathbf{w}^T\nabla f(\mathbf{w}) = h(g(\mathbf{w})) - \mathbb{E}h(g(\mathbf{Z})). \tag{6.1}$$

When Σ is positive-definite, we know from Lemma 2.14 that

$$f(\mathbf{w}) = -\int_0^{\infty} \mathbb{E}[h(g(\mathrm{e}^{-s}\mathbf{w} + \sqrt{1-\mathrm{e}^{-2s}}\mathbf{Z})) - \mathbb{E}h(g(\mathbf{Z}))]\,\mathrm{d}s \tag{6.2}$$

solves the Stein equation (6.1), provided that the second order partial derivatives of f exist. Evaluating both sides of (6.1) at the random variable $\mathbf{W}$ and taking expectations, we see that we can bound the quantity $\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z}))$ by bounding $\mathbb{E}\{\nabla^T\Sigma\nabla f(\mathbf{W}) - \mathbf{W}^T\nabla f(\mathbf{W})\}$. Therefore we can obtain bounds on the error in approximating $g(\mathbf{W})$ by $g(\mathbf{Z})$ in much the same way as we would for approximating $\mathbf{W}$ by $\mathbf{Z}$, although in general the derivatives of g will be not be bounded, and so the derivatives of the solution (6.2) of the Stein equation may also be unbounded even when h has bounded derivatives. Hence, we will be unable to bound the derivatives of the solution by using the bounds given in Lemma 2.14.

The approach is general and allows us to obtain explicit bounds on the error in approximating the statistic $g(\mathbf{W})$ by $g(\mathbf{Z})$, provided that g satisfies certain differentiability conditions and that its partial derivatives satisfy certain integrability conditions; the X_{ik} satisfy certain integrability conditions and the covariance matrix of $\mathbf{W}$ is positive-definite. An advantage of the multivariate normal approach to proving limit theorems for statistics of the form $g(\mathbf{W})$, where g is a continuous function, is that we can deduce a Stein equation and smoothness estimates for its solution from the multivariate normal Stein equation. This is particularly useful for distributions that have Stein equations for which it is difficult to establish smoothness estimates. Indeed, we are able to use the multivariate normal Stein equation to prove a limit theorem for a statistic that has an asymptotic Product Normal distribution – see Theorem 6.22.

To illustrate this approach we obtain general limit theorems for the univariate case ($d = 1$) with the assumptions that the X_i are i.i.d., g is twice differentiable and that these derivatives are of polynomial growth (we say that a function $f : \mathbb{R} :\rightarrow \mathbb{R}$ has polynomial growth if there exist constants $A, B, k \geq 0$ such that $|f(x)| \leq A + B|x|^k$ for all $x \in \mathbb{R}$). Our first limit theorem has bound on the convergence rate of $O(n^{-1/2})$ for smooth test functions, which is what would arise from the Berry-Esséen Theorem. However, with the additional assumption that g is an even function ($g(w) = g(-w)$), or that the X_i have vanishing third moments, we obtain a faster order $O(n^{-1})$ bound on the rate of convergence. For the case of even g, we achieve this better bound by using symmetry arguments similar to those used in Chapter 4.

The techniques used to obtain these bounds can be applied in much more general situations. For instance, we can weaken the assumptions on the growth rate of g, extend to $g : \mathbb{R}^d \rightarrow \mathbb{R}$, and consider statistics for which the X_{ik} are not independent. In Section 6.2 we illustrate how this can be done without giving explicit bounds. Moreover, we see that if $g : \mathbb{R}^d \rightarrow \mathbb{R}$ is an even function, then the bound on the convergence rate will be of order $n_1^{-1} + \cdots + n_d^{-1}$ provided that g satisfies certain differentiability and integrability conditions.

In Section 6.3 we illustrate our results with applications chosen to be of interest in their own right and also to demonstrate the techniques involved in the multivariate normal Stein equation approach that is developed in this chapter. We obtain bounds of order n^{-1} for χ^2 and Symmetric Variance-Gamma limit theorems. We also obtain limit theorems for a statistic that is asymptotically distributed as the product of r independent normal random variables; succeeding where we failed in Chapter 5. Our final application is Friedman's χ^2 statistic. This application demonstrates how the multivariate normal approach can be used to prove limit theorems for statistics with a complicated dependence structure, and it also provides an example of how we can deal with problems for which the covariance matrix of $\mathbf{W}$ is not positive-definite.

6.1 Limit Theorems for functions of the normal distribution

In this section we consider the convergence of the statistic $g(W)$ to its limiting distribution in the case that g and its first and second derivatives are of polynomial growth and the sequence of random variables $X, X_1, \ldots, X_n$ are independent and identically distributed. We present three theorems, with the first theorem giving a bound of order $n^{-1/2}$, and the second and third theorems providing bounds of order n^{-1}. To achieve these better bounds, we need the additional assumptions that g is an even function, or that $\mathbb{E}X^3 = 0$.

6.1.1 A limit theorem with a bound of order $n^{-1/2}$

Here is our first limit theorem. The proof is straightforward and illustrates how the standard normal Stein equation can be used to prove limit theorems for statistics of the form $g(W)$, where g is continuous.

Theorem 6.1. *Suppose that $g : \mathbb{R} \to \mathbb{R}$ is absolutely continuous and that its first derivative, g', is of polynomial growth, with $|g'(w)| \leq A + B|w|^k$ almost everywhere, where A, B and k are non-negative constants. Let $X, X_1, X_2, \ldots, X_n$ be a collection of i.i.d. random variables with mean zero, unit variance and $\mathbb{E}|X|^{k+4} < \infty$. Let $W = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$. Then, for absolutely continuous $h : \mathbb{R} \to \mathbb{R}$, we have*

$$|\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))| \leq \frac{\alpha}{\sqrt{n}} \|h'\|, \tag{6.3}$$

where

$$\begin{aligned}
\alpha = A\bigg[&\bigg(1 + \frac{1}{\sqrt{2\pi}}\bigg)(2 + \mathbb{E}|X|^3) + \frac{2 + \mathbb{E}X^4}{\sqrt{2\pi n}}\bigg] \\
&+ 2^{k-1}B\bigg[(2 + \mathbb{E}|X|^3)\bigg(\frac{2^{k-1/2}}{\sqrt{\pi}}\mathbb{E}|W|^{k+1} + (2^k + 1)\mathbb{E}|W|^k + \frac{2^{k/2+1/2}\Gamma(\frac{k}{2} + 1)}{\sqrt{\pi}} \\
&+ \frac{2^{k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}}\bigg) + \frac{2^{k/2+1/2}\Gamma(\frac{k}{2} + 1)}{\sqrt{\pi n}}(2 + \mathbb{E}X^4) + \frac{2^k + 2}{n^{k/2}}(\mathbb{E}|X|^{k+3} + 2\mathbb{E}|X|^{k+1}) \\
&+ \frac{2^{k-1/2}}{\sqrt{\pi}n^{(k+1)/2}}(2\mathbb{E}|X|^{k+2} + \mathbb{E}|X|^{k+4})\bigg].
\end{aligned}$$

Remark 6.2. Recall from Lemma 4.6 that providing that $\mathbb{E}|X|^k$ exists, $\mathbb{E}|W|^k = O(1)$. Hence, the bound in Theorem 6.1 is of order $n^{-1/2}$.

Let us now prove Theorem 6.1.

Proof Part I: Expansions and Bounding

Consider the standard normal Stein equation with test function $h(g(w))$:

$$f''(w) - wf'(w) = h(g(w)) - \mathbb{E}h(g(Z)). \tag{6.4}$$

Let $W_i = W - \frac{1}{\sqrt{n}}X_i$, so that W_i and X_i are independent. Then evaluating both sides of (6.4) at the random variable W, taking expectations and then Taylor expanding and using the local approach coupling, as we did in Example 2.8, gives

$$\begin{aligned} |\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))| &\leq \frac{1}{n^{3/2}}\sum_{i=1}^n \left\{\frac{1}{2}\sup_\theta \mathbb{E}|X_i^3 f^{(3)}(W_\theta^{(i)})| + \mathbb{E}X_i^2 \sup_\theta \mathbb{E}|X_i f^{(3)}(W_\theta^{(i)})|\right\} \\ &= \frac{1}{2\sqrt{n}}\left\{\sup_\theta \mathbb{E}|X_1^3 f^{(3)}(W_\theta^{(1)})| + 2\sup_\theta \mathbb{E}|X_1 f^{(3)}(W_\theta^{(1)})|\right\}, \end{aligned} \tag{6.5}$$

where $W_\theta^{(i)} = W_i + \frac{\theta}{\sqrt{n}}X_i$ for some $\theta \in (0,1)$. Substituting $f^{(3)}(w) = wf''(w) + f'(w) + g'(w)h'(g(w))$ into (6.5) and then applying the triangle inequality gives

$$\begin{aligned} |\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))| \leq \frac{1}{2\sqrt{n}}\bigg\{ &\sup_\theta \mathbb{E}|X_1^3 W_\theta^{(1)} f''(W_\theta^{(1)})| + 2\sup_\theta \mathbb{E}|X_1 W_\theta^{(1)} f''(W_\theta^{(1)})| \\ &+ \sup_\theta \mathbb{E}|X_1^3 f'(W_\theta^{(1)})| + 2\sup_\theta \mathbb{E}|X_1 f'(W_\theta^{(1)})| \\ &+ \|h'\|\bigg(\sup_\theta \mathbb{E}|X_1^3 g'(W_\theta^{(1)})| + 2\sup_\theta \mathbb{E}|X_1 g'(W_\theta^{(1)})|\bigg)\bigg\}. \end{aligned} \tag{6.6}$$

By substituting the expression $f^{(3)}(w) = wf''(w) + f'(w) + g'(w)h'(g(w))$ into (6.5), which results in a bound for $\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))$ involving $f'(w)$ and $f''(w)$ rather than $f^{(3)}(w)$, we shall obtain a final bound that only involves the first derivative of g, as opposed to its first and second derivatives, which would have been the case had we not made the substitution.

Proof Part II: Smoothness estimates for the solution of the Stein equation

We now require bounds on the derivatives of f. Since we cannot assume that $g'(w)$ is bounded as $|w| \to \infty$ (this is only the case for $k = 0$), we cannot use the bounds of Lemma 2.14 to bound the derivatives of f. We bound the derivatives of f through a series of lemmas. We obtain bounds for $f'(w)$, $f''(w)$, and also a bound for $f^{(3)}(w)$ which we use in the proof of Theorem 6.6.

Lemma 6.3. *Let f be the solution of the Stein equation (6.4). Suppose $h:\mathbb{R}\rightarrow\mathbb{R}$ and $g:\mathbb{R}\rightarrow\mathbb{R}$ are absolutely continuous then, provided the following integrals exist, we have*

$$|f'(w)| \leq \|h'\|\int_0^\infty\int_{-\infty}^\infty \mathrm{e}^{-s}|g'(z_s^x)|\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \tag{6.7}$$

$$|f''(w)| \leq \|h'\|\int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-2s}}{\sqrt{1-\mathrm{e}^{-2s}}}|g'(z_s^x)||x|\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \tag{6.8}$$

where

$$z_s^x = \mathrm{e}^{-s}w + x\sqrt{1-\mathrm{e}^{-2s}}$$

and $\phi(x)$ is the p.d.f. of the standard normal distribution.

Suppose now that $h\in C_b^2(\mathbb{R})$ and $g\in C^2(\mathbb{R})$ then, provided the following integral exists, we have

$$|f^{(3)}(w)| \leq \int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\Big[\|h''\|g'(z_s^x)^2 + \|h'\||g''(z_s^x)|\Big]|x|\phi(x)\,\mathrm{d}x\,\mathrm{d}s. \tag{6.9}$$

Proof. The solution of the Stein equation (6.4) is given by

$$f(w) = -\int_0^\infty\int_{-\infty}^\infty [h(g(z_s^x)) - \mathbb{E}h(g(Z))]\phi(x)\,\mathrm{d}x\,\mathrm{d}s,$$

where $Z\sim N(0,1)$. From (2.19) and (2.20), we have

$$\begin{aligned} f'(w) &= -\int_0^\infty\int_{-\infty}^\infty \mathrm{e}^{-s}g'(z_s^x)h'(g(z_s^x))\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \\ f''(w) &= -\int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-2s}}{\sqrt{1-\mathrm{e}^{-2s}}}g'(z_s^x)h'(g(z_s^x))x\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \\ f^{(3)}(w) &= -\int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\Big[g'(z_s^x)^2h''(g(z_s^x)) + g''(z_s^x)h'(g(z_s^x))\Big]x\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \end{aligned}$$

and the result now easily follows. □

We now specialise to the case that the derivatives of g are of polynomial growth.

Lemma 6.4. *Suppose $|g'(w)|\leq A_1+B_1|w|^{k_1}$ and $|g''(w)|\leq A_2+B_2|w|^{k_2}$, where A_1, A_2, B_1, B_2, k_1 and k_2 are non-negative constants, then for all $w\in\mathbb{R}$,*

$$|f'(w)| \leq \|h'\|\bigg[A_1 + 2^{k_1}B_1\bigg(|w|^{k_1} + \frac{2^{k_1/2}\Gamma(\frac{k_1+1}{2})}{\sqrt{\pi}}\bigg)\bigg], \tag{6.10}$$

$$|f''(w)| \leq \sqrt{\frac{2}{\pi}}\|h'\|\bigg[A_1 + 2^{k_1}B_1\bigg(|w|^{k_1} + 2^{k_1/2}\Gamma\bigg(\frac{k_1}{2}+1\bigg)\bigg)\bigg], \tag{6.11}$$

$$\begin{aligned} |f^{(3)}(w)| &\leq \frac{\sqrt{2\pi}}{4}\bigg\{2\|h''\|\bigg[A_1^2 + 2^{2k_1}B_1^2(|w|^{2k_1} + 2^{k_1}\Gamma(k_1+1))\bigg] \\ &\quad + \|h'\|\bigg[A_2 + 2^{k_2}B_2\bigg(|w|^{k_2} + 2^{k_2/2}\Gamma\bigg(\frac{k_2}{2}+1\bigg)\bigg)\bigg]\bigg\}. \end{aligned} \tag{6.12}$$

Proof. We begin by obtaining simple inequalities for $|g^{(i)}(z_s^x)|^p$, for $i = 1, 2$ and $p = 1, 2$,

$$\begin{aligned} |g^{(i)}(z_s^x)|^p &= (A_i + B_i|z_s^x|^{k_i})^p \\ &\leq 2^{p-1}(A_i^p + B_i^p|z_s^x|^{pk_i}) \\ &\leq 2^{p-1}(A_i^p + B_i^p(\mathrm{e}^{-pk_is}|w|^{pk_i} + (1-\mathrm{e}^{-2s})^{pk_i/2}|x|^{pk_i})) \\ &\leq 2^{p-1}(A_i^p + 2^{pk_i}B_i^p(|w|^{pk_i} + |x|^{pk_i})), \end{aligned} \tag{6.13}$$

where the first inequality follows from the inequality $|a+b|^q \leq 2^{q-1}(a^q + b^q)$ for $q \geq 1$, and to obtain the second inequality we used the inequality $|a+b|^r \leq 2^r(|a|^r + |b|^r)$ for $r \geq 0$. We may now obtain a bound on $|f'(w)|$, $|f''(w)|$ and $|f^{(3)}(w)|$ by using inequalities (6.7), (6.8) and (6.9) combined with the inequality (6.13). We have

$$\begin{aligned} |f'(w)| &\leq \|h'\|\int_0^\infty\int_{-\infty}^\infty \mathrm{e}^{-s}[A_1 + 2^{k_1}B_1(|w|^{k_1} + |x|^{k_1})]\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \\ |f''(w)| &\leq \|h'\|\int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-2s}}{\sqrt{1-\mathrm{e}^{-2s}}}[A_1 + 2^{k_1}B_1(|w|^{k_1} + |x|^{k_1})]|x|\phi(x)\,\mathrm{d}x\,\mathrm{d}s, \\ |f^{(3)}(w)| &\leq \int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\Big\{2\|h''\|[A_1^2 + 2^{2k_1}B_1^2(|w|^{2k_1} + |x|^{2k_1})] \\ &\quad + \|h'\|[A_2 + 2^{k_2}B_2(|w|^{k_2} + |x|^{k_2})]\Big\}|x|\phi(x)\,\mathrm{d}x\,\mathrm{d}s. \end{aligned}$$

We then use the formulas $\int_{-\infty}^\infty |x|^k\phi(x)\,\mathrm{d}x = \mathbb{E}|Z|^k = \frac{2^{k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}}$ (this is formula 17 of Winkelbauer [75]) and $\int_0^\infty \frac{\mathrm{e}^{-ks}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s = \frac{\sqrt{\pi}\Gamma(k/2)}{2\Gamma((k+1)/2)}$ (see (2.21)) to obtain inequalities (6.10), (6.11) and (6.12). □

With these bounds for the derivatives of the solution of the Stein equation, we are able to bound the expectations on the right-hand side of (6.6).

Lemma 6.5. *Suppose $\theta \in (0,1)$, $k \geq 0$, $p \geq 0$ and that $\mathbb{E}|X|^{k+p+1} < \infty$, then*

$$\mathbb{E}|X_1^p g'(W_\theta^{(1)})| \leq A_1\mathbb{E}|X|^p + 2^{k_1}B_1\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_1} + \frac{\mathbb{E}|X|^{k_1+p}}{n^{k_1/2}}\bigg), \tag{6.14}$$

$$\begin{aligned}\mathbb{E}|X_1^pW_\theta^{(1)}f''(W_\theta^{(1)})| &\leq \sqrt{\frac{2}{\pi}}\|h'\|\bigg[A_1\bigg(\mathbb{E}|X|^p+\frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg)+2^{k_1}B_1\bigg(2^{k_1-1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_1+1} \\ &\quad+\frac{\mathbb{E}|X|^{p+k_1+1}}{n^{(k_1+1)/2}}\bigg)+2^{k_1/2}\Gamma\bigg(\frac{k_1}{2}+1\bigg)\bigg(\mathbb{E}|X|^p+\frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg)\bigg)\bigg], \qquad (6.15)\\ \mathbb{E}|X_1^pf'(W_\theta^{(1)})| &\leq \|h'\|\bigg[A_1\mathbb{E}|X|^p+2^{k_1}B_1\bigg(2^{k_1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_1}+\frac{\mathbb{E}|X|^{k_1+p}}{n^{k_1/2}}\bigg) \\ &\quad+\frac{2^{k_1/2}\Gamma(\frac{k_1+1}{2})}{\sqrt{\pi}}\mathbb{E}|X|^p\bigg)\bigg]. \qquad (6.16)\end{aligned}$$

Proof. We begin by noting two simple inequalities for $|W_\theta^{(i)}|^r$:

$$|W_\theta^{(i)}|^r \leq 2^r\bigg(|W_i|^r+\theta^r\frac{|X_i|^r}{n^{r/2}}\bigg)\leq 2^r\bigg(|W_i|^r+\frac{|X_i|^r}{n^{r/2}}\bigg), \qquad r\geq 0 \qquad (6.17)$$

and

$$|W_\theta^{(i)}|^r \leq 2^{r-1}\bigg(|W_i|^r+\theta^r\frac{|X_i|^r}{n^{r/2}}\bigg)\leq 2^{r-1}\bigg(|W_i|^r+\frac{|X_i|^r}{n^{r/2}}\bigg), \qquad r\geq 1 \qquad (6.18)$$

where we used that $|a+b|^r\leq 2^r(|a|^r+|b|^r)$ for all $r\geq 0$ and $|a+b|^r\leq 2^{r-1}(|a|^r+|b|^r)$ for $r\geq 1$.

Using inequality (6.17) we are able to establish inequality (6.14):

$$\mathbb{E}|X_1^pg'(W_\theta^{(1)})|\leq \mathbb{E}|X_1^p(A_1+B_1|W_\theta^{(1)}|^{k_1}|\leq A_1\mathbb{E}|X|^p+2^{k_1}B_1\bigg(\mathbb{E}|X|^p\mathbb{E}|W_1|^{k_1}+\frac{\mathbb{E}|X|^{k_1+p}}{n^{k_1/2}}\bigg).$$

Inequality (6.14) now follows on using the inequality $\mathbb{E}|W_1|^k\leq\mathbb{E}|W|^k$ for $k\geq 0$ (see Lemma 4.7).

We now prove that inequality (6.15) holds. Applying inequality (6.11) gives

$$\mathbb{E}|X_1^pW_\theta^{(1)}f''(W_\theta^{(1)})|\leq\sqrt{\frac{2}{\pi}}\|h'\|\mathbb{E}\bigg|X_i^pW_\theta^{(1)}\bigg[A_1+2^{k_1}B_1\bigg(|W_\theta^{(i)}|^{k_1}+2^{k_1/2}\Gamma\bigg(\frac{k_1}{2}+1\bigg)\bigg)\bigg]\bigg|.$$

We then use inequalities (6.17) and (6.18) to obtain

$$\begin{aligned}\mathbb{E}|X_1^pW_\theta^{(1)}f''(W_\theta^{(1)})|&\leq\sqrt{\frac{2}{\pi}}\|h'\|\bigg[A_1\bigg(\mathbb{E}|X|^p+\frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg)+2^{k_1}B_1\bigg(2^{k_1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W^{(1)}|^{k_1+1}\\ &\quad+\frac{\mathbb{E}|X|^{p+k_1+1}}{n^{(k_1+1)/2}}\bigg)+2^{k_1/2}\Gamma\bigg(\frac{k_1}{2}+1\bigg)\bigg(\mathbb{E}|X|^p+\frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg)\bigg)\bigg],\end{aligned}$$

and the result then follows from again using the inequality $\mathbb{E}|W_1|^k\leq\mathbb{E}|W|^k$ for $k\geq 0$.

The verification of inequality (6.16) follows from a similar, but simpler, calculation to the one that was used to prove inequality (6.15). □

Substituting the bounds of Lemma 6.5 into inequality (6.6) completes the proof of Theorem 6.1.

6.1.2 Limit theorems with bounds of order n^{-1}

We now present our limit theorems with bounds of order n^{-1}.

Theorem 6.6. *Suppose that $g:\mathbb{R}\rightarrow\mathbb{R}$ is an even function, and that its first and second derivatives are of polynomial growth, with $|g'(w)|\leq A_1+B_2|w|^{k_1}$ and $|g''(w)|\leq A_2+B_2|w|^{k_2}$ almost everywhere, where A_j, B_j and k_j, for $j=1,2$, are non-negative constants. Let $X,X_1,X_2,\ldots,X_n$ be a collection of i.i.d. random variables with mean zero, unit variance, $\mathbb{E}|X|^{2k_1+5}<\infty$ and $\mathbb{E}|X|^{k_2+5}<\infty$. Let $W=\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$. Then, for $h\in C_b^2(\mathbb{R})$, we have*

$$|\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))|\leq\frac{1}{12n}\Big\{[A_2(\alpha_1+\beta_1)+B_2(\alpha_2+\beta_2)+A_1(\alpha_3+\beta_3)+B_1(\alpha_4+\beta_4)]\|h'\| \\ [A_1^2(\alpha_5+\beta_5)+B_1^2(\alpha_6+\beta_6)]\|h''\|\Big\},$$

where

$$\begin{aligned}
\alpha_1 &= 18+\frac{9\sqrt{2\pi}}{2}+\frac{9\sqrt{2\pi}\mathbb{E}|X|^3}{2\sqrt{n}}+\left(2+\frac{\sqrt{2\pi}}{2}\right)\mathbb{E}X^4+\frac{\sqrt{2\pi}\mathbb{E}|X|^5}{2\sqrt{n}},\\
\beta_1 &= |\mathbb{E}X^3|\left[15+\frac{15\sqrt{2\pi}}{4}\left(1+\frac{1}{\sqrt{n}}\right)+\left(6+\frac{3\sqrt{2\pi}}{2}+\frac{4}{n}\right)\mathbb{E}|X|^3+\frac{3\sqrt{2\pi}\mathbb{E}X^4}{2\sqrt{n}}+\frac{2\mathbb{E}|X|^5}{n}\right],\\
\alpha_2 &= 2^{2k_2-1/2}\sqrt{\pi}(9+\mathbb{E}X^4)\mathbb{E}|W|^{k_2+1}+2^{k_2+1}(9+\mathbb{E}X^4)\mathbb{E}|W|^{k_2}\\
&\quad+2^{3k_2/2-1/2}\sqrt{\pi}\Gamma\left(\frac{k_2}{2}+1\right)\left[9+\frac{9\mathbb{E}|X|^3}{\sqrt{n}}+\mathbb{E}X^4+\frac{\mathbb{E}|X|^5}{\sqrt{n}}\right]\\
&\quad+\frac{2^{k_2+1}}{n^{k_2/2}}(9\mathbb{E}|X|^{k_2+2}+\mathbb{E}|X|^{k_2+4})+\frac{2^{2k_2-1/2}\sqrt{\pi}}{n^{(k_2+1)/2}}(9\mathbb{E}|X|^{k_2+3}+\mathbb{E}|X|^{k_2+5}),\\
\beta_2 &= 3\mathbb{E}|X^3|\Big\{2^{k_2+1}3^{k_2-1}(2+\mathbb{E}|X|^3)\mathbb{E}|W|^{k_2+2}+2^{2k_2-3/2}\sqrt{\pi}(2\mathbb{E}|X|^3+5)\mathbb{E}|W|^{k_2+1}\\
&\quad+2^{k_2}(3+\mathbb{E}|X|^3)\mathbb{E}|W|^{k_2}+2^{k_2}3^{k_2-1}(2+\mathbb{E}|X|^3)\mathbb{E}|W|^{k_2}\\
&\quad+2^{3k_2/2-3/2}\sqrt{\pi}\Gamma\left(\frac{k_2}{2}+1\right)\left[5+\frac{5}{\sqrt{n}}+2\mathbb{E}|X|^3+\frac{2\mathbb{E}X^4}{n}\right]\\
&\quad+2^{k_2/2+1}3^{k_2-1}\Gamma\left(\frac{k_2}{2}+1\right)\left[6+\left(3+\frac{4}{n}\right)\mathbb{E}|X|^3+\frac{2\mathbb{E}|X|^5}{n}\right]\\
&\quad+\frac{2^{k_2}}{n^{k_2/2}}(3\mathbb{E}|X|^{k_2+1}+\mathbb{E}|X|^{k_2+3})+\frac{2^{k_2}3^{k_2-1}}{n^{k_2/2}}(2\mathbb{E}|X|^{k_2+1}+\mathbb{E}|X|^{k_2+3})\\
&\quad+\frac{2^{3k_2/2-3/2}\sqrt{\pi}}{n^{(k_2+1)/2}}(5\mathbb{E}|X|^{k_2+2}+2\mathbb{E}|X|^{k_2+4})+\frac{2^{k_2+1}3^{k_2-1}}{n^{k_2/2+1}}(2\mathbb{E}|X|^{k_2+3}+\mathbb{E}|X|^{k_2+5})\Big\},\\
\alpha_3 &= \frac{4\sqrt{2}}{\sqrt{\pi}}(9+\mathbb{E}X^4),
\end{aligned}$$

$$
\begin{aligned}
\beta_3 &= \frac{6\sqrt{2}}{\sqrt{\pi}}|\mathbb{E}X^3|(3+\mathbb{E}|X|^3), \\
\alpha_4 &= \frac{2^{3k_1/2+5/2}}{\sqrt{\pi}}\bigg[(9+\mathbb{E}X^4)\bigg(2^{k_1/2}\mathbb{E}|W|^{k_1}+\Gamma\bigg(\frac{k_1}{2}+1\bigg)\bigg)+\frac{2^{k_1/2}}{n^{k_1/2}}(9\mathbb{E}|X|^{k_1+2}+\mathbb{E}|X|^{k_1+4})\bigg], \\
\beta_4 &= \frac{3\cdot 2^{3k_1/2+3/2}}{\sqrt{\pi}}|\mathbb{E}X^3|\bigg\{(3+\mathbb{E}|X|^3)\bigg(2^{k_1/2}\mathbb{E}|W|^{k_1}+\Gamma\bigg(\frac{k_1}{2}+1\bigg)\bigg) \\
&\quad +\frac{2^{k_1/2}}{n^{k_1/2}}(3\mathbb{E}|X|^{k_1+1}+\mathbb{E}|X|^{k_1+3})\bigg\}, \\
\alpha_5 &= 36+9\sqrt{2\pi}+\frac{9\sqrt{2\pi}\mathbb{E}|X|^3}{\sqrt{n}}+(4+\sqrt{2\pi})\mathbb{E}X^4+\frac{\sqrt{2\pi}\mathbb{E}|X|^5}{\sqrt{n}}, \\
\beta_5 &= |\mathbb{E}X^3|\bigg[30+\frac{15\sqrt{2\pi}}{2}\bigg(1+\frac{1}{\sqrt{n}}\bigg)+\bigg(12+3\sqrt{2\pi}+\frac{8}{n}\bigg)\mathbb{E}|X|^3+\frac{3\sqrt{2\pi}\mathbb{E}X^4}{\sqrt{n}}+\frac{4\mathbb{E}|X|^5}{n}\bigg], \\
\alpha_6 &= 2\bigg\{2^{4k_1-1/2}\sqrt{\pi}(9+\mathbb{E}X^4)\mathbb{E}|W|^{2k_1+1}+2^{2k_1+1}(9+\mathbb{E}X^4)\mathbb{E}|W|^{2k_1} \\
&\quad +2^{3k_1-1/2}\sqrt{\pi}\Gamma(k_1+1)\bigg[9+\frac{9\mathbb{E}|X|^3}{\sqrt{n}}+\mathbb{E}X^4+\frac{\mathbb{E}|X|^5}{\sqrt{n}}\bigg] \\
&\quad +\frac{2^{2k_1+1}}{n^{k_1}}(9\mathbb{E}|X|^{2k_1+2}+\mathbb{E}|X|^{2k_1+4})+\frac{2^{4k_1-1/2}\sqrt{\pi}}{n^{k_1+1/2}}(9\mathbb{E}|X|^{2k_1+3}+\mathbb{E}|X|^{2k_1+5})\bigg\}, \\
\beta_6 &= 6|\mathbb{E}X^3|\bigg\{2^{2k_1+1}3^{2k_1-1}(2+\mathbb{E}|X|^3)\mathbb{E}|W|^{2k_1+2}+2^{4k_1-3/2}\sqrt{\pi}(2\mathbb{E}|X|^3+5)\mathbb{E}|W|^{2k_1+1} \\
&\quad +2^{2k_1}(3+\mathbb{E}|X|^3)\mathbb{E}|W|^{2k_1}+2^{2k_1}3^{2k_1-1}(2+\mathbb{E}|X|^3)\mathbb{E}|W|^{2k_1} \\
&\quad +2^{3k_1-3/2}\sqrt{\pi}\Gamma(k_1+1)\bigg[5+\frac{5}{\sqrt{n}}+2\mathbb{E}|X|^3+\frac{2\mathbb{E}X^4}{n}\bigg] \\
&\quad +2^{k_1+1}3^{2k_1-1}\Gamma(k_1+1)\bigg[6+\bigg(3+\frac{4}{n}\bigg)\mathbb{E}|X|^3+\frac{2\mathbb{E}|X|^5}{n}\bigg] \\
&\quad +\frac{2^{2k_1}}{n^{k_1}}(3\mathbb{E}|X|^{2k_1+1}+\mathbb{E}|X|^{2k_1+3})+\frac{2^{2k_1}3^{2k_1-1}}{n^{k_1}}(2\mathbb{E}|X|^{2k_1+1}+\mathbb{E}|X|^{2k_1+3}) \\
&\quad +\frac{2^{3k_1-3/2}\sqrt{\pi}}{n^{k_1+1/2}}(5\mathbb{E}|X|^{2k_1+2}+2\mathbb{E}|X|^{2k_1+4})+\frac{2^{2k_1+1}3^{2k_1-1}}{n^{k_1+1}}(2\mathbb{E}|X|^{2k_1+3}+\mathbb{E}|X|^{2k_1+5})\bigg\}.
\end{aligned}
$$

If we drop the assumption that g is symmetric and replace it with the condition that $\mathbb{E}X^3 = 0$ then we can also get a bound of order n^{-1}:

Theorem 6.7. *Let the X_i and $g : \mathbb{R} \to \mathbb{R}$ be defined as in Theorem 6.6, but let us drop the assumption that g is an even function and replace it with the condition that $\mathbb{E}X^3 = 0$. Then, for $h \in C^2(\mathbb{R})$, we have*

$$
|\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))| \le \frac{1}{12n}\Big[(A_2\alpha_1 + B_2\alpha_2 + A_1\alpha_3 + B_1\alpha_4)\|h'\| + (A_1^2\alpha_5 + B_1^2\alpha_6)\|h''\|\Big],
$$

where the α_i and β_i are as in Theorem 6.6.

Firstly, we prove Theorem 6.6.

Proof of Theorem 6.6 Part I: Symmetry Argument for Optimal Rate

The following key lemma, which is for the more general multivariate setting and applies even to functions that have a growth rate that is faster than polynomial, is crucial in enabling us to obtain a bound of order n^{-1}.

Lemma 6.8. *Suppose that Σ is positive-definite $d \times d$ matrix and let $f(\mathbf{w})$ be the solution (6.2) of the Stein equation (6.1). Suppose that $g : \mathbb{R}^d \to \mathbb{R}$ is an even function: $g(\mathbf{u}) = g(-\mathbf{u})$ for all $\mathbf{u} \in \mathbb{R}^d$. Then, the solution f of the Stein equation is an even function. Moreover, for $k \geq 1$ and odd, provided that $\frac{\partial^k f(\mathbf{w})}{\prod_{j=1}^k \partial w_{i_j}}$ exists, we have*

$$\mathbb{E}\bigg[\frac{\partial^k f(\mathbf{Z})}{\prod_{j=1}^k \partial w_{i_j}}\bigg] = 0.$$

Proof. As Σ is positive-definite, from (6.2) we have the following formula for the solution of the Stein equation (6.1):

$$f(-\mathbf{w}) = -\int_0^\infty \int_{\mathbb{R}^d} \Big[h(g(-\mathrm{e}^{-s}\mathbf{w} + \sqrt{1-\mathrm{e}^{-2s}}\mathbf{x})) - \mathbb{E}h(g(\mathbf{Z}))\Big] p(\mathbf{x})\,\mathrm{d}\mathbf{x}\,\mathrm{d}s,$$

where

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}\sqrt{\det(\Sigma)}} \exp\bigg(-\frac{1}{2}\mathbf{x}^T\Sigma^{-1}\mathbf{x}\bigg).$$

Substituting $\mathbf{x} = -\mathbf{y}$ gives

$$f(-\mathbf{w}) = -\int_0^\infty \int_{\mathbb{R}^d} \Big[h(g(-\mathrm{e}^{-s}\mathbf{w} - \sqrt{1-\mathrm{e}^{-2s}}\mathbf{y})) - \mathbb{E}h(g(\mathbf{Z}))\Big] p(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s.$$

Since g is an even function we have

$$f(-\mathbf{w}) = -\int_0^\infty \int_{\mathbb{R}^d} \Big[h(g(\mathrm{e}^{-s}\mathbf{w} + \sqrt{1-\mathrm{e}^{-2s}}\mathbf{y})) - \mathbb{E}h(g(\mathbf{Z}))\Big] p(\mathbf{y})\,\mathrm{d}\mathbf{y}\,\mathrm{d}s = f(\mathbf{w}).$$

Hence, we have shown that the solution (6.2) of the Stein equation (6.1) is an even function.

Since f is an even function, the partial derivatives of order k, where $k \geq 1$ is an odd number, are odd functions, provided that they exist. Therefore, since $\mathbf{Z} \stackrel{\mathcal{D}}{=} -\mathbf{Z}$, we have

$$\mathbb{E}\bigg[\frac{\partial^k f(\mathbf{Z})}{\prod_{j=1}^k \partial w_{i_j}}\bigg] = -\mathbb{E}\bigg[\frac{\partial^k f(-\mathbf{Z})}{\prod_{j=1}^k \partial w_{i_j}}\bigg] = -\mathbb{E}\bigg[\frac{\partial^k f(\mathbf{Z})}{\prod_{j=1}^k \partial w_{i_j}}\bigg],$$

and so

$$\mathbb{E}\bigg[\frac{\partial^k f(\mathbf{Z})}{\prod_{j=1}^k \partial w_{i_j}}\bigg]=0,$$

as required ☐

Proof of Theorem 6.6 Part II: Expansions and Bounding

As we did in the proof of Theorem 6.1, we consider the standard normal Stein equation with test function $h(g(x))$:

$$f''(x)-xf'(x)=h(g(x))-\mathbb{E}h(g(Z)).$$

Evaluating both sides at W and taking expectations gives

$$\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))=\mathbb{E}f''(W)-\mathbb{E}Wf'(W).$$

We may therefore obtain a bound on $\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))$ by bounding $\mathbb{E}f''(W)-\mathbb{E}Wf'(W)$. To obtain our bounds, we will use the local approach coupling $W_i=W-\frac{X_i}{\sqrt{n}}$. We will make repeated use of Taylor expansions and we write $W_{\theta_p}^{(i)}=W_i+\theta_p\frac{X_i}{\sqrt{n}}$, where $\theta_p\in(0,1)$ for $p=1,2,\ldots,6$.

Taylor expanding $f'(W)$ about W_i, and using independence and that the X_i have mean zero and unit variance, gives

$$\begin{aligned}\mathbb{E}Wf'(W)&=\frac{1}{\sqrt{n}}\sum_{i=1}^n\mathbb{E}X_if'(W)\\&=\frac{1}{\sqrt{n}}\sum_{i=1}^n\mathbb{E}X_if'(W_i)+\frac{1}{n}\sum_{i=1}^n\mathbb{E}X_i^2f''(W_i)+\frac{1}{2n^{3/2}}\sum_{i=1}^n\mathbb{E}X_i^3f^{(3)}(W_i)+R_1\\&=\frac{1}{n}\sum_{i=1}^n\mathbb{E}f''(W_i)+\frac{\mathbb{E}X^3}{2n^{3/2}}\sum_{i=1}^n\mathbb{E}f^{(3)}(W_i)+R_1,\end{aligned}$$

where

$$R_1=\frac{1}{6n^2}\sum_{i=1}^n\mathbb{E}X_i^4f^{(4)}(W_{\theta_1}^{(i)}).$$

We now Taylor expand $f''(W_i)$ and $f^{(3)}(W_i)$ about W to obtain

$$\begin{aligned}\mathbb{E}Wf'(W)&=\frac{1}{n}\sum_{i=1}^n\mathbb{E}f''(W)-\frac{1}{n^{3/2}}\sum_{i=1}^n\mathbb{E}X_if^{(3)}(W)+\frac{\mathbb{E}X^3}{2n^{3/2}}\sum_{i=1}^n\mathbb{E}f^{(3)}(W)+R_1+R_2+R_3\\&=\mathbb{E}f''(W)-\frac{1}{n^{3/2}}\sum_{i=1}^n\mathbb{E}X_if^{(3)}(W)+\frac{\mathbb{E}X^3}{2\sqrt{n}}\mathbb{E}f^{(3)}(W)+R_1+R_2+R_3,\end{aligned}$$

where

$$\begin{aligned} R_2 &= -\frac{1}{2n^2}\sum_{i=1}^n \mathbb{E}X_i^2 f^{(4)}(W_{\theta_2}^{(i)}), \\ R_3 &= -\frac{\mathbb{E}X^3}{2n^2}\sum_{i=1}^n \mathbb{E}X_i f^{(4)}(W_{\theta_3}^{(i)}). \end{aligned}$$

We now Taylor expand $f^{(3)}(W_i)$ about W, and use independence and that the X_i have mean zero, to obtain

$$\begin{aligned} \mathbb{E}Wf'(W) &= \mathbb{E}f''(W) - \frac{1}{n^{3/2}}\sum_{i=1}^n \mathbb{E}X_i f^{(3)}(W_i) + \frac{\mathbb{E}X^3}{2\sqrt{n}}\mathbb{E}f^{(3)}(W) + R_1 + R_2 + R_3 + R_4 \\ &= \mathbb{E}f''(W) + \frac{\mathbb{E}X^3}{2\sqrt{n}}\mathbb{E}f^{(3)}(W) + R_1 + R_2 + R_3 + R_4, \qquad (6.19) \end{aligned}$$

where

$$R_4 = -\frac{1}{n^2}\sum_{i=1}^n \mathbb{E}X_i^2 f^{(4)}(W_{\theta_5}^{(i)}).$$

To achieve the desired $O(n^{-1})$ bound we need to show that $\mathbb{E}f^{(3)}(W)$ is of order $n^{-1/2}$, since in general $\mathbb{E}X^3 \neq 0$. As was the case in the proof of Theorem 4.4, we consider the standard normal Stein equation with test function $f^{(3)}(w)$:

$$\psi''(w) - w\psi'(w) = f^{(3)}(w) - \mathbb{E}f^{(3)}(Z). \qquad (6.20)$$

By Lemma 6.8, we have that $\mathbb{E}f^{(3)}(Z) = 0$, and therefore

$$\mathbb{E}f^{(3)}(W) = \mathbb{E}f^{(3)}(W) - \mathbb{E}f^{(3)}(Z) = \mathbb{E}\{\psi''(W) - W\psi'(W)\}.$$

We can bound $\mathbb{E}\{\psi''(W) - W\psi'(W)\}$ using a similar approach to that used in Example 2.6 to obtain

$$\mathbb{E}f^{(3)}(W) = R_5 + R_6$$

where

$$\begin{aligned} R_5 &= -\frac{1}{2n^{3/2}}\sum_{i=1}^n \mathbb{E}X_i^3 \psi^{(3)}(W_{\theta_5}^{(i)}), \\ R_6 &= \frac{1}{n^{3/2}}\sum_{i=1}^n \mathbb{E}X_i \psi^{(3)}(W_{\theta_6}^{(i)}). \end{aligned}$$

We have therefore shown that

$$\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z)) = \mathbb{E}\{f''(W) - Wf'(W)\} = R_1 + R_2 + R_3 + R_4 + \frac{\mathbb{E}X^3}{2\sqrt{n}}(R_5 + R_6),$$

and on using the triangle inequality and that the X_i are identically distributed this expression becomes

$$\begin{aligned} |\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))| \leq \frac{1}{n}\bigg\{ & \frac{1}{6}\sup_\theta \mathbb{E}|X_1^4 f^{(4)}(W_\theta^{(i)})| + \frac{3}{2}\sup_\theta \mathbb{E}|X_1^2 f^{(4)}(W_\theta^{(i)})| \\ & + \frac{|\mathbb{E}X^3|}{2}\bigg(\sup_\theta \mathbb{E}|X_1 f^{(4)}(W_\theta^{(i)})| + \frac{1}{2}\sup_\theta \mathbb{E}|X_1^3 \psi^{(3)}(W_\theta^{(i)})| \\ & + \sup_\theta \mathbb{E}|X_1 \psi^{(3)}(W_\theta^{(i)})|\bigg)\bigg\}. \end{aligned} \tag{6.21}$$

We can express $f^{(4)}(w)$ and $\psi^{(3)}(w)$ as follows

$$f^{(4)}(w) = wf^{(3)}(w) + 2f''(w) + g'(w)^2 h''(g(w)) + g''(w)h'(g(w)), \tag{6.22}$$

$$\psi^{(3)}(w) = (1 + w^2)\psi'(w) + f^{(3)}(w), \tag{6.23}$$

and substituting these expressions into (6.21) gives

$$\begin{aligned} & |\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))| \\ & \leq \frac{1}{12n}\bigg\{2\|h''\|\Big(\sup_\theta \mathbb{E}X_1^4 g'(W_\theta^{(1)})^2 + 9\sup_\theta \mathbb{E}X_1^2 g'(W_\theta^{(1)})^2\Big) + 2\|h'\|\Big(\sup_\theta \mathbb{E}|X_1^4 g''(W_\theta^{(1)})| \\ & \quad + 9\sup_\theta \mathbb{E}|X_1^2 g''(W_\theta^{(1)})|\Big) + 4\sup_\theta \mathbb{E}|X_1^4 f''(W_\theta^{(1)})| + 36\sup_\theta \mathbb{E}|X_1^2 f''(W_\theta^{(1)})| \\ & \quad + 2\sup_\theta \mathbb{E}|X_1^4 W_\theta^{(1)} f^{(3)}(W_\theta^{(1)})| + 18\sup_\theta \mathbb{E}|X_1^2 W_\theta^{(1)} f^{(3)}(W_\theta^{(1)})| \\ & \quad + 3|\mathbb{E}X^3|\Big[\|h''\|\Big(\sup_\theta \mathbb{E}|X_1^3 g'(W_\theta^{(1)})^2| + 3\sup_\theta \mathbb{E}|X_1 g'(W_\theta^{(1)})^2|\Big) + \|h'\|\Big(\sup_\theta \mathbb{E}|X_1^3 g''(W_\theta^{(1)})| \\ & \quad + 3\sup_\theta \mathbb{E}|X_1 g''(W_\theta^{(1)})|\Big) + 2\sup_\theta \mathbb{E}|X_1^3 f''(W_\theta^{(1)})| + 6\sup_\theta \mathbb{E}|X_1 f''(W_\theta^{(1)})| \\ & \quad + 2\sup_\theta \mathbb{E}|X_1^3 W_\theta^{(1)} f^{(3)}(W_\theta^{(1)})| + 5\sup_\theta \mathbb{E}|X_1 W_\theta^{(1)} f^{(3)}(W_\theta^{(1)})| \\ & \quad + \sup_\theta \mathbb{E}|X_1^3(1 + (W_\theta^{(1)})^2)\psi'(W_\theta^{(1)})| + 2\sup_\theta \mathbb{E}|X_1(1 + (W_\theta^{(1)})^2)\psi'(W_\theta^{(1)})|\Big]\bigg\}. \end{aligned} \tag{6.24}$$

As was the case in the proof of Theorem 6.1, making use of the substitutions (6.22) and (6.23) means the we have a final bound for $\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))$ that involves lower order derivatives of g. In this case, we improve from a bound involving the first four derivatives of g to one involving just the first two derivatives of g.

Proof Part III: Smoothness estimates for the solution of the Stein equation

To apply bound (6.24) we require bounds for $f'(w)$, $f''(w)$, $f^{(3)}(w)$ and $\psi'(w)$. We have already bounded $f'(w)$, $f''(w)$ and $f^{(3)}(w)$ in Lemma 6.4 and we now set about bounding $\psi'(w)$.

Lemma 6.9. *Let ψ be the solution of the Stein equation (6.20). Suppose $h:\mathbb{R}\rightarrow\mathbb{R}$ and $g:\mathbb{R}\rightarrow\mathbb{R}$ are twice differentiable then, provided the following integral exists, we have*

$$\begin{aligned}|\psi'(w)|\leq\int_0^{\infty}\int_0^{\infty}\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-4t}}{\sqrt{1-\mathrm{e}^{-2t}}}\Big[\|h''\|g'(z_{s,t}^{x,y})^2+\|h'\||g''(z_{s,t}^{x,y})|\Big]\\ \times|x|\phi(x)|y|\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}t,\end{aligned}\tag{6.25}$$

where

$$z_{s,t}^{x,y}=\mathrm{e}^{-s-t}w+y\mathrm{e}^{-s}\sqrt{1-\mathrm{e}^{-2t}}+x\sqrt{1-\mathrm{e}^{-2s}}.$$

Proof. From (2.20), we have

$$\psi'(w)=-\int_0^{\infty}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-t}}{\sqrt{1-\mathrm{e}^{-2t}}}f^{(3)}(\mathrm{e}^{-t}w+y\sqrt{1-\mathrm{e}^{-2t}})y\phi(y)\,\mathrm{d}y\,\mathrm{d}t.\tag{6.26}$$

We can use (2.20) to write $f^{(3)}(z_t^y)$ as

$$f^{(3)}(z_t^y)=-\int_0^{\infty}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-3(s+t)}}{\sqrt{1-\mathrm{e}^{-2s}}}[(g'(z_{s,t}^{x,y}))^2h''(g(z_{s,t}^{x,y})+g''(z_{s,t}^{x,y})h'(g(z_{s,t}^{x,y}))]x\phi(x)\,\mathrm{d}x\,\mathrm{d}s,$$

where

$$z_{s,t}^{x,y}=\mathrm{e}^{-s}(\mathrm{e}^{-t}w+y\sqrt{1-\mathrm{e}^{-2t}})+x\sqrt{1-\mathrm{e}^{-2s}}=\mathrm{e}^{-s-t}w+y\mathrm{e}^{-s}\sqrt{1-\mathrm{e}^{-2t}}+x\sqrt{1-\mathrm{e}^{-2s}}.$$

Substituting our expression for $f^{(3)}(z_t^y)$ into (6.26) gives

$$\begin{aligned}\psi'(w)=\int_0^{\infty}\int_0^{\infty}\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-4t}}{\sqrt{1-\mathrm{e}^{-2t}}}[(g'(z_{s,t}^{x,y}))^2h''(g(z_{s,t}^{x,y}))\\ +g''(z_{s,t}^{x,y})h'(g(z_{s,t}^{x,y}))]x\phi(x)y\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}t,\end{aligned}$$

from which we can easily deduce (6.25). □

We now specialise to the case that g' and g'' are of polynomial growth. Applying Lemma 6.9 with $|g'(w)|\leq A_1+B_1|w|^{k_1}$ and $|g''(w)|\leq A_2+B_2|w|^{k_2}$ gives:

Lemma 6.10. *Suppose* $|g'(w)| \le A_1 + B_1|w|^{k_1}$ *and* $|g''(w)| \le A_2 + B_2|w|^{k_2}$*, where* A_1*,* A_2*,* B_1*,* B_2*,* k_1 *and* k_2 *are non-negative constants, then for all* $w \in \mathbb{R}$*,*

$$\begin{aligned}|\psi'(w)| \le \frac{2\|h''\|}{3}\bigg[A_1^2 + 3^{2k_1}B_1^2(|w|^{2k_1} + 2^{k_1+1}\Gamma(k_1+1))\bigg] \\ + \frac{\|h'\|}{3}\bigg[A_2 + 3^{k_2}B_2\bigg(|w|^{k_2} + 2^{k_2/2+1}\Gamma\bigg(\frac{k_2}{2}+1\bigg)\bigg)\bigg].\end{aligned} \tag{6.27}$$

Proof. To prove (6.27) we make use of the following simple inequality

$$|g^{(i)}(z_{s,t}^{x,y})|^p \le 2^{p-1}(A_i^p + 3^{pk_i}B_i^p(|w|^{pk_i} + |x|^{pk_i} + |y|^{pk_i})), \qquad \text{for } i = 1,2 \text{ and } p \ge 1, \tag{6.28}$$

which can be verified by a similar calculation that was used to prove inequality (6.13). Using inequalities (6.26) and (6.28) we have

$$\begin{aligned}|\psi'(w)| \le \int_0^\infty\int_{-\infty}^\infty \frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-4t}}{\sqrt{1-\mathrm{e}^{-2t}}}\bigg\{2\|h''\|[A_1^2 + 3^{2k_1}B_1^2(|w|^{2k_1} + |x|^{2k_1} + |y|^{2k_1})] \\ + \|h'\|[A_2 + 3^{k_2}B_2(|w|^{k_2} + |x|^{k_2} + |y|^{k_2})]\bigg\}|x|\phi(x)\,\mathrm{d}x\,\mathrm{d}s.\end{aligned}$$

We then use the formulas $\int_{-\infty}^\infty |x|^k\phi(x)\,\mathrm{d}x = \mathbb{E}|Z|^k = \frac{2^{k/2}\Gamma(\frac{k+1}{2})}{\sqrt{\pi}}$ and $\int_0^\infty \frac{\mathrm{e}^{-ks}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s = \frac{\sqrt{\pi}\Gamma(k/2)}{2\Gamma((k+1)/2)}$ (see (2.21)) to obtain inequality (6.27). □

Finally, we present our bounds for the expectations on the right-hand side of (6.24). For $\theta \in (0,1)$, $k_1 \ge 0$, $k_2 \ge 0$, $p \ge 0$, and $\mathbb{E}|X|^{2k_1+p+2} < \infty$ and $\mathbb{E}|X|^{k_2+p+2} < \infty$, we have

$$\begin{aligned}\mathbb{E}|X_i^p g'(W_\theta^{(i)})^2| &\le 2A_1^2\mathbb{E}|X|^p + 2^{2k_1+1}B_1^2\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{2k_1} + \frac{\mathbb{E}|X|^{2k_1+p}}{n^{k_1}}\bigg), \\ \mathbb{E}|X_i^p g''(W_\theta^{(i)})| &\le A_2\mathbb{E}|X|^p + 2^{k_2}B_2\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_2} + \frac{\mathbb{E}|X|^{k_2+p}}{n^{k_2/2}}\bigg), \\ \mathbb{E}|X_i^p f''(W_\theta^{(i)})| &\le \sqrt{\frac{2}{\pi}}\|h'\|\bigg[A_1\mathbb{E}|X|^p + 2^{k_1}B_1\bigg(2^{k_1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_1} + \frac{\mathbb{E}|X|^{k_1+p}}{n^{k_1/2}}\bigg) \\ &\quad +2^{k_1/2}\Gamma\bigg(\frac{k_1}{2}+1\bigg)\mathbb{E}|X|^p\bigg)\bigg],\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}|X_i^p W_\theta^{(i)} f^{(3)}(W_\theta^{(i)})| \\
&\leq \frac{\sqrt{2\pi}}{4}\bigg\{\|h''\|\bigg[2A_1^2\bigg(\mathbb{E}|X|^p + \frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg) + 2^{3k_1+1}B_1^2\bigg(2^{k_1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{2k_1+1} + \frac{\mathbb{E}|X|^{2k_1+p+1}}{n^{k_1/2+1}}\bigg) \\
&\quad + \Gamma(k_1+1)\bigg(\mathbb{E}|X|^p + \frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg)\bigg)\bigg] + \|h'\|\bigg[A_2\bigg(\mathbb{E}|X|^p + \frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg) \\
&\quad + 2^{3k_2/2}B_2\bigg(2^{k_2/2}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_2+1} + \frac{\mathbb{E}|X|^{k_2+p+1}}{n^{(k_2+1)/2}}\bigg) \\
&\quad + \Gamma\bigg(\frac{k_2}{2}+1\bigg)\bigg(\mathbb{E}|X|^p + \frac{\mathbb{E}|X|^{p+1}}{\sqrt{n}}\bigg)\bigg)\bigg]\bigg\}, \\
&\mathbb{E}|X_i^p(1+(W_\theta^{(i)})^2)\psi'(W_\theta^{(i)})| \\
&\leq \frac{2\|h''\|}{3}\bigg[A_1^2\bigg(3\mathbb{E}|X|^p + \frac{2\mathbb{E}|X|^{p+2}}{n}\bigg) + 3^{2k_1}B_1^2\bigg(2^{2k_1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{2k_1} + \frac{\mathbb{E}|X|^{2k_1+p}}{n^{k_1}}\bigg) \\
&\quad 2^{2k_1+1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W^{(i)}|^{2k_1+2} + \frac{\mathbb{E}|X|^{2k_1+p+2}}{n^{k_1+1}}\bigg) + 2^{k_1+1}\Gamma(k_1+1)\bigg(3\mathbb{E}|X|^p + \frac{2\mathbb{E}|X|^{p+2}}{n}\bigg)\bigg)\bigg] \\
&\quad + \frac{\|h'\|}{3}\bigg[A_2\bigg(3\mathbb{E}|X|^p + \frac{2\mathbb{E}|X|^{p+2}}{n}\bigg) + 3^{k_2}B_2\bigg(2^{k_2}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_2} + \frac{\mathbb{E}|X|^{k_2+p}}{n^{k_2/2}}\bigg) \\
&\quad + 2^{k_2+1}\bigg(\mathbb{E}|X|^p\mathbb{E}|W|^{k_2+2} + \frac{\mathbb{E}|X|^{k_2+p+2}}{n^{k_2/2+1}}\bigg) + 2^{k_2/2+1}\Gamma\bigg(\frac{k_2}{2}+1\bigg)\bigg(3\mathbb{E}|X|^p + \frac{2\mathbb{E}|X|^{p+2}}{n}\bigg)\bigg)\bigg].
\end{aligned}
$$

These inequalities can be verified using similar calculations to those used to prove Lemma 6.5. Substituting these bounds into inequality (6.24) completes the proof of Theorem 6.6.

Proof of Theorem 6.7

Recall expression (6.19), which we obtained in Part II of the proof of Theorem 6.6:

$$\mathbb{E}Wf'(W) = \mathbb{E}f''(W) + \frac{\mathbb{E}X^3}{2\sqrt{n}}\mathbb{E}f^{(3)}(W) + R_1 + R_2 + R_3 + R_4.$$

Up to this point in the proof we had not made use of the fact that g was an even function. On applying the symmetry arguments we had that

$$\mathbb{E}Wf'(W) = \mathbb{E}f''(W) + \frac{\mathbb{E}X^3}{2\sqrt{n}}\mathbb{E}f^{(3)}(W) + R_1 + R_2 + R_3 + R_4 + \frac{\mathbb{E}X^3}{2\sqrt{n}}(R_5 + R_6).$$

We now note that the additional remainder terms R_5 and R_6 that arise from our symmetry conditions are pre multiplied by $\mathbb{E}X^3$. Hence, provided that the first and second derivatives of $g(w)$ are of polynomial growth, Theorem 6.7 follows on taking $\mathbb{E}X^3 = 0$ in the bound of Theorem 6.6.

6.2 Extensions to more general limit theorems

In this section we consider how the limit theorems of the previous section can be generalised. We consider three possible generalisations: weaker assumptions on the growth rate of the derivatives of g, an extension to $g:\mathbb{R}^d\rightarrow\mathbb{R}$, and removing the assumption that the X_{ik} are independent and identically distributed.

6.2.1 Weaker assumptions on the growth rate of the derivatives

We can use the multivariate normal Stein equation to obtain limit theorems, similar to those of Section 6.1, involving statistics of the form $g(W)$ for which the first and second derivatives of $g:\mathbb{R}\rightarrow\mathbb{R}$ have growth rate that is faster than polynomial. To illustrate how this would be done we consider the case of an extension of Theorem 6.1; the extension of Theorems 6.6 and 6.7 are similar. We would make use of bound (6.6) for $\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))$ and then use Lemma 6.3 to bound the first and second derivatives of the solution of the Stein equation. From (6.7) and (6.8) we can see that in principle this could be done provided that $|g'(w)|\leq A\mathrm{e}^{tw^2}$, for some $A>0$ and $t<1/2$. However, in general bounding the derivatives of the solution of the Stein equation would be more difficult than was the case we considered in Section 6.1, and the bounds on the derivatives would more complicated than those that we obtained in Lemma 6.4. Moreover, weakening on the assumptions on the growth rate would mean that we would have to impose stricter conditions on the expectations of the X_i.

6.2.2 Generalisation to real-valued functions of more than one variable

We now illustrate how we can obtain limit theorems for the statistic $g(\mathbf{W})$, where $g:\mathbb{R}^d\rightarrow\mathbb{R}$ is continuous. For simplicity, we consider the case that the X_{ik} are i.i.d. with zero mean and unit variance. We will therefore be considering the multivariate normal Stein equation (2.11) with test function $h(g(\mathbf{w}))$:

$$\sum_{k=1}^d\left(\frac{\partial^2 f}{\partial w_k^2}(\mathbf{w})+w_k\frac{\partial f}{\partial w_k}(\mathbf{w})\right)=h(g(\mathbf{w}))-\mathbb{E}h(g(\mathbf{Z})), \tag{6.29}$$

which has solution

$$f(\mathbf{w})=\int_0^\infty\int_{\mathbb{R}^d}[h(g(\mathrm{e}^{-s}\mathbf{w}+\sqrt{1-\mathrm{e}^{-2s}}\mathbf{x}))-\mathbb{E}h(g(\mathbf{Z}))]\prod_{l=1}^d\phi(x_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}s. \tag{6.30}$$

Since we are considering the case of i.i.d. random variables, we will use the local approach coupling $\mathbf{W}_k^{(i)} = \mathbf{W} - \frac{1}{\sqrt{n_k}}\mathbf{X}_{ik}$, where $\mathbf{X}_{ik}$ is a vector which has X_{ik} as its k-th entry and the other $d-1$ entries are given by zero.

Suppose we wish to obtain a $O(n_1^{-1/2} + \cdots + n_d^{-1/2})$ bound for the convergence of $g(\mathbf{W})$ to its limiting distribution. Then instead of using the univariate bound bound (6.6), which we obtained in the proof of Theorem 6.1, we would use:

Lemma 6.11. *Let $(X_{ik})_{1\leq i\leq n_k}$, for $1 \leq k \leq d$, be collections of i.i.d. random variables with zero mean and unit variance. Suppose that the third order partial derivatives of the solution (6.30) of the Stein equation (6.29) exist, then providing that the expectations on the right-hand side of (6.31) exist, we have*

$$|\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z}))| \leq \sum_{k=1}^{d} \frac{1}{\sqrt{n_k}}\bigg\{\frac{1}{2}\sup_{\theta}\mathbb{E}\bigg|X_{ik}^3\frac{\partial^3 f}{\partial w_k^3}(\mathbf{W}_{k,\theta}^{(i)})\bigg| + \sup_{\theta}\mathbb{E}\bigg|X_{ik}\frac{\partial^3 f}{\partial w_k^3}(\mathbf{W}_{k,\theta}^{(i)})\bigg|\bigg\}, \quad (6.31)$$

where $\mathbf{W}_{k,\theta}^{(i)} = \mathbf{W}_k^{(i)} + \frac{\theta}{\sqrt{n_k}}\mathbf{X}_{ik}$ for some $\theta \in (0,1)$.

Proof. Evaluating both side of (6.29) at $\mathbf{W}$ and taking expectations gives

$$\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z})) = \mathbb{E}\bigg\{\sum_{k=1}^{d}\bigg(\frac{\partial^2 f}{\partial w_k^2}(\mathbf{W}) + W_k\frac{\partial f}{\partial w_k}(\mathbf{W})\bigg)\bigg\}. \quad (6.32)$$

We use the coupling $\mathbf{W}_k^{(i)} = \mathbf{W} - \frac{1}{\sqrt{n_k}}\mathbf{X}_{ik}$ for $1 \leq k \leq d$. Since $\mathbf{W}_k^{(i)}$ is independent of $\mathbf{X}_{ik}$, we can perform a similar calculation to the one used to prove (6.6) to bound the right-hand side of (6.32). Doing this gives (6.31), as required. □

The third order partial derivatives of f that appear in (6.31) could then be bounded using Lemma 6.13 (below).

With the additional assumption that g is even, and that its fourth order partial derivatives exist, we can obtain bounds of order $n_1^{-1} + \cdots + n_d^{-1}$ by applying the following lemma.

Lemma 6.12. *Let $(X_{ik})_{1\leq i\leq n_k}$, for $1 \leq k \leq d$, be collections of i.i.d. random variables with zero mean and unit variance. Suppose that the fourth order partial derivatives of the solution f of the Stein equation (6.29) exist and that the third order partial derivatives of the solution ψ_k of the partial differential equation*

$$\sum_{j=1}^{d}\bigg(\frac{\partial^2 \psi_k}{\partial w_j^2}(\mathbf{w}) - w_j\frac{\partial \psi_k}{\partial w_j}(\mathbf{w})\bigg) = \frac{\partial^3 f}{\partial w_k^3}(\mathbf{w}) \quad (6.33)$$

exist for $1 \leq k \leq d$. Then, providing that the expectations on the right-hand side of (6.34) exist, we have

$$\begin{aligned} |\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z}))| \leq \sum_{k=1}^{d} \bigg\{ \frac{1}{6n_k} \bigg[\sup_\theta \mathbb{E}\bigg| X_{ik}^4 \frac{\partial^4 f}{\partial w_k^4}(\mathbf{W}_{k,\theta}^{(i)}) \bigg| + 9 \sup_\theta \mathbb{E}\bigg| X_{ik}^2 \frac{\partial^4 f}{\partial w_k^4}(\mathbf{W}_{k,\theta}^{(i)}) \bigg| \\ + 3|\mathbb{E}X_k^3| \sup_\theta \mathbb{E}\bigg| X_{ik} \frac{\partial^4 f}{\partial w_k^4}(\mathbf{W}_{k,\theta}^{(i)}) \bigg| \bigg] \\ + \frac{|\mathbb{E}X_k^3|}{4\sqrt{n_k}} \sum_{j=1}^{d} \frac{1}{\sqrt{n_j}} \bigg[\sup_\theta \mathbb{E}\bigg| X_{ij}^3 \frac{\partial^3 \psi_k}{\partial w_j^3}(\mathbf{W}_{j,\theta}^{(i)}) \bigg| \\ + 2 \sup_\theta \mathbb{E}\bigg| X_{ij} \frac{\partial^3 \psi_k}{\partial w_j^3}(\mathbf{W}_{j,\theta}^{(i)}) \bigg| \bigg] \bigg\}, \end{aligned} \tag{6.34}$$

where $\mathbf{W}_{k,\theta}^{(i)} = \mathbf{W}_k^{(i)} + \frac{\theta}{\sqrt{n_k}} \mathbf{X}_{ik}$ for some $\theta \in (0,1)$.

Proof. We require a bound on

$$\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z})) = \mathbb{E}\bigg\{ \sum_{k=1}^{d} \bigg(\frac{\partial^2 f}{\partial w_k^2}(\mathbf{W}) + W_k \frac{\partial f}{\partial w_k}(\mathbf{W}) \bigg) \bigg\}.$$

We use the coupling $\mathbf{W}_k^{(i)} = \mathbf{W} - \frac{1}{\sqrt{n_k}} \mathbf{X}_{ik}$ for $1 \leq k \leq d$, and similar calculation to the one used to prove (6.19), to obtain

$$\begin{aligned} |\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z}))| \leq \frac{1}{n_k} \sum_{k=1}^{d} \bigg\{ \frac{1}{6} \bigg(\sup_\theta \mathbb{E}\bigg| X_{1k}^4 \frac{\partial^4 f}{\partial w_k^4}(\mathbf{W}_{k,\theta}^{(1)}) \bigg| + \frac{3}{2} \sup_\theta \mathbb{E}\bigg| X_{1k}^2 \frac{\partial^4 f}{\partial w_k^4}(\mathbf{W}_{k,\theta}^{(1)}) \bigg| \\ + \frac{1}{2} |\mathbb{E}X_k^3| \sup_\theta \mathbb{E}\bigg| X_{1k} \frac{\partial^4 f}{\partial w_k^4}(\mathbf{W}_{k,\theta}^{(1)}) \bigg| \bigg) + \frac{|\mathbb{E}X_k^3|}{2\sqrt{n_k}} \bigg| \mathbb{E} \frac{\partial^3 f}{\partial w_k^3}(\mathbf{W}) \bigg| \bigg\}. \end{aligned} \tag{6.35}$$

To obtain the desired $O(n_1^{-1} + \cdots + n_d^{-1})$ convergence rate, we use a similar argument to the one used in the proof of Theorem 6.6. We consider the multivariate normal Stein equation

$$\sum_{j=1}^{d} \bigg(\frac{\partial^2 \psi_k}{\partial w_j^2}(\mathbf{w}) - w_j \frac{\partial \psi_k}{\partial w_j}(\mathbf{w}) \bigg) = \frac{\partial^3 f}{\partial w_k^3}(\mathbf{w}) - \mathbb{E}\bigg[\frac{\partial^3 f}{\partial w_k^3}(\mathbf{Z}) \bigg].$$

By Lemma 6.8 we have that $\mathbb{E}\frac{\partial^3 f}{\partial w_k^3}(\mathbf{Z}) = 0$, and therefore

$$\mathbb{E}\bigg[\frac{\partial^3 f}{\partial w_k^3}(\mathbf{W}) \bigg] = \mathbb{E}\bigg[\frac{\partial^3 f}{\partial w_k^3}(\mathbf{W}) \bigg] - \mathbb{E}\bigg[\frac{\partial^3 f}{\partial w_k^3}(\mathbf{Z}) \bigg] = \sum_{j=1}^{d} \mathbb{E}\bigg(\frac{\partial^2 \psi_j}{\partial w_j^2}(\mathbf{W}) - W_j \frac{\partial \psi_k}{\partial w_j}(\mathbf{W}) \bigg). \tag{6.36}$$

We can use Lemma 6.11 to bound the right-hand side of (6.36), and substituting this bound into (6.35) gives (6.34), as required. □

The partial derivatives of f and ψ can be bounded using the following lemma.

Lemma 6.13. *Suppose $h:\mathbb{R}\rightarrow\mathbb{R}$ and $g:\mathbb{R}\rightarrow\mathbb{R}$ are four times differentiable then, provided the following integrals exist, we have*

$$\begin{aligned}
\left|\frac{\partial^3 f}{\partial w_k^3}(\mathbf{w})\right| &\leq \int_0^\infty\int_{\mathbb{R}^d}\frac{\mathrm{e}^{-3s}}{\sqrt{1-\mathrm{e}^{-2s}}}\left[\|h''\|\left(\frac{\partial g}{\partial w_k}(\mathbf{z}_s^x)\right)^2+\|h'\|\left|\frac{\partial^2 g}{\partial w_k^2}(\mathbf{z}_s^x)\right|\right]|x_k|\prod_{l=1}^d\phi(x_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}s,\\
\left|\frac{\partial^4 f}{\partial w_k^4}(\mathbf{w})\right| &\leq \int_0^\infty\int_{\mathbb{R}^d}\frac{\mathrm{e}^{-4s}}{\sqrt{1-\mathrm{e}^{-2s}}}\left[\|h^{(3)}\|\left|\frac{\partial g}{\partial w_k}(\mathbf{z}_s^x)\right|^3+3\|h''\|\left|\frac{\partial g}{\partial w_k}(\mathbf{z}_s^x)\frac{\partial^2 g}{\partial w_k^2}(\mathbf{z}_s^x)\right|\right.\\
&\quad\left.+\|h'\|\left|\frac{\partial^3 g}{\partial w_k^3}(\mathbf{z}_s^x)\right|\right]|x_k|\prod_{l=1}^d\phi(x_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}s, \qquad (6.37)\\
\left|\frac{\partial^3 \psi_k}{\partial w_j^3}(\mathbf{w})\right| &\leq \int_0^\infty\int_0^\infty\int_{\mathbb{R}^{2d}}\frac{\mathrm{e}^{-5s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-6t}}{\sqrt{1-\mathrm{e}^{-2t}}}\left\{\|h^{(4)}\|\left(\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\right)^2\left(\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\right)^2\right.\\
&\quad+\|h^{(3)}\|\left|\left(\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\right)^2\frac{\partial^2 g}{\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})+\left(\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\right)^2\frac{\partial^2 g}{\partial w_k^2}(\mathbf{z}_{s,t}^{x,y})\right.\\
&\quad\left.+4\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^2 g}{\partial w_k\partial w_j}(\mathbf{z}_{s,t}^{x,y})\right|+\|h''\|\left|2\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^3 g}{\partial w_k\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})\right.\\
&\quad\left.+2\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^3 g}{\partial w_k^2\partial w_j}(\mathbf{z}_{s,t}^{x,y})+2\left(\frac{\partial^2 g}{\partial w_k\partial w_j}(\mathbf{z}_{s,t}^{x,y})\right)^2+\frac{\partial^2 g}{\partial w_k^2}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^2 g}{\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})\right|\\
&\quad\left.+\|h'\|\left|\frac{\partial^4 g}{\partial w_k^2\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})\right|\right\}|x_k||y_j|\prod_{l=1}^d\phi(x_l)\phi(y_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}\,\mathrm{d}s\,\mathrm{d}t, \qquad (6.38)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{z}_s^x &= \mathrm{e}^{-s}\mathbf{w}+\mathbf{x}\sqrt{1-\mathrm{e}^{-2s}},\\
\mathbf{z}_{s,t}^{x,y} &= \mathrm{e}^{-s-t}\mathbf{w}+\mathbf{y}\mathrm{e}^{-s}\sqrt{1-\mathrm{e}^{-2t}}+\mathbf{x}\sqrt{1-\mathrm{e}^{-2s}}.
\end{aligned}$$

Proof. The inequalities for $\frac{\partial^3 f}{\partial w_k^3}(\mathbf{w})$ and $\frac{\partial^4 f}{\partial w_k^4}(\mathbf{w})$ can be verified using a straightforward generalisation of the calculations used in the proof of Lemma 6.3.

We now verify that inequality (6.38) holds. From (2.20), we have

$$\frac{\partial\psi_k}{\partial w_j}(\mathbf{w})=-\int_0^\infty\int_{\mathbb{R}^d}\frac{\mathrm{e}^{-t}}{\sqrt{1-\mathrm{e}^{-2t}}}\frac{\partial^3 f}{\partial w_k^3}(\mathbf{z}_t^y)y_j\prod_{l=1}^d\phi(y_l)\,\mathrm{d}\mathbf{y}\,\mathrm{d}t.$$

Partial differentiating with respect to w_j gives

$$\frac{\partial^3\psi_k}{\partial w_j^3}(\mathbf{w}) = -\int_0^\infty\int_{\mathbb{R}^d}\frac{\mathrm{e}^{-t}}{\sqrt{1-\mathrm{e}^{-2t}}}\frac{\partial^2}{\partial w_j^2}\bigg(\frac{\partial^3 f}{\partial w_k^3}(\mathbf{z}_t^y)\bigg)y_j\prod_{l=1}^d\phi(y_l)\,\mathrm{d}\mathbf{y}\,\mathrm{d}t. \tag{6.39}$$

We can use (2.20) to write $\frac{\partial^5 f}{\partial w_j^3\partial w_k^2}(\mathbf{z}_t^y)$ as follows

$$\frac{\partial^5 f}{\partial w_j^3\partial w_k^2}(\mathbf{z}_t^y) = -\int_0^\infty\int_{\mathbb{R}^d}\frac{\mathrm{e}^{-5(s+t)}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\partial^4(h\circ g)}{\partial w_j^2\partial w_k^2}(\mathbf{z}_{s,t}^{x,y})x_k\prod_{l=1}^d\phi(x_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}s,$$

where

$$\mathbf{z}_{s,t}^{x,y} = \mathrm{e}^{-s}(\mathrm{e}^{-t}\mathbf{w}+\mathbf{y}\sqrt{1-\mathrm{e}^{-2t}})+\mathbf{x}\sqrt{1-\mathrm{e}^{-2s}} = \mathrm{e}^{-s-t}\mathbf{w}+\mathbf{y}\mathrm{e}^{-s}\sqrt{1-\mathrm{e}^{-2t}}+\mathbf{x}\sqrt{1-\mathrm{e}^{-2s}}.$$

Substituting our expression for $\frac{\partial^5 f}{\partial w_j^2\partial w_k^3}(\mathbf{z}_s^x)$ into (6.39) gives

$$\begin{aligned}\frac{\partial^3\psi_k}{\partial w_j^3}(\mathbf{w}) &= \int_0^\infty\int_0^\infty\int_{\mathbb{R}^{2d}}\frac{\mathrm{e}^{-5s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-6t}}{\sqrt{1-\mathrm{e}^{-2t}}}\frac{\partial^4(h\circ g)}{\partial w_j^2\partial w_k^2}(\mathbf{z}_{s,t}^{x,y})x_ky_j\prod_{l=1}^d\phi(x_l)\phi(y_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}\,\mathrm{d}s\,\mathrm{d}t\\
&= \int_0^\infty\int_0^\infty\int_{\mathbb{R}^{2d}}\frac{\mathrm{e}^{-5s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-6t}}{\sqrt{1-\mathrm{e}^{-2t}}}\bigg\{h^{(4)}(g(\mathbf{z}_{s,t}^{x,y}))\bigg(\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\bigg)^2\bigg(\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\bigg)^2\\
&+h^{(3)}(g(\mathbf{z}_{s,t}^{x,y}))\bigg[\bigg(\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\bigg)^2\frac{\partial^2 g}{\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})+\bigg(\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\bigg)^2\frac{\partial^2 g}{\partial w_k^2}(\mathbf{z}_{s,t}^{x,y})\\
&+4\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^2 g}{\partial w_k\partial w_j}(\mathbf{z}_{s,t}^{x,y})\bigg]+h''(g(\mathbf{z}_{s,t}^{x,y}))\bigg[2\frac{\partial g}{\partial w_k}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^3 g}{\partial w_k\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})\\
&+2\frac{\partial g}{\partial w_j}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^3 g}{\partial w_k^2\partial w_j}(\mathbf{z}_{s,t}^{x,y})+2\bigg(\frac{\partial^2 g}{\partial w_k\partial w_j}(\mathbf{z}_{s,t}^{x,y})\bigg)^2+\frac{\partial^2 g}{\partial w_k^2}(\mathbf{z}_{s,t}^{x,y})\frac{\partial^2 g}{\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})\bigg]\\
&+h'(g(\mathbf{z}_{s,t}^{x,y}))\frac{\partial^4 g}{\partial w_k^2\partial w_j^2}(\mathbf{z}_{s,t}^{x,y})\bigg\}x_ky_j\prod_{l=1}^d\phi(x_l)\phi(y_l)\,\mathrm{d}\mathbf{x}\,\mathrm{d}\mathbf{y}\,\mathrm{d}s\,\mathrm{d}t,\end{aligned}$$

from which we can easily deduce (6.38). □

The bounds for the derivatives of f and ψ in Lemma 6.13 are more complicated than those used for the univariate case. We will therefore not provide any general bounds. However, in Theorem 6.20 we use Lemmas 6.12 and 6.13 with $g(x,y)=xy$ to prove a VG$(1,0,1,0)$ limit theorem.

We now make some remarks regarding the above lemmas.

Remark 6.14. As in the univariate case, we can achieve a bound of order $n_1^{-1}+\cdots+n_d^{-1}$ in the case that $\mathbb{E}X_k^3=0$, for all $1\le k\le d$, without needing to use symmetry considerations. This is because all terms involving ψ_k disappear when $\mathbb{E}X_k^3=0$.

Remark 6.15. A limit theorem with a bound of order $n_1^{-1/2}+\cdots+n_d^{-1/2}$ resulting from an application of Lemma 6.11 would require that $h\in C_b^2(\mathbb{R})$ and $g\in C^2(\mathbb{R}^d)$, as well as appropriate growth rates on g' and g'', which are slightly stronger conditions than in the univariate Theorem 6.1. Similarly, a bound of order $n_1^{-1}+\cdots+n_d^{-1}$, in the case of symmetric g, that is obtained by an application of Lemma 6.12 would require the stronger conditions that $h\in C_b^4(\mathbb{R})$ and $g\in C^4(\mathbb{R}^d)$, as well as appropriate growth rates on the first four derivatives of g.

These stronger conditions arise as an artefact of the calculations we used in proving Lemmas (6.11) and (6.12), and it may be possible to impose weaker differentiability conditions on h and g and still arrive at bounds of order $n_1^{-1}+\cdots+n_d^{-1}$. We were able to weaken the conditions on h and g in the univariate case by exploiting the the fact that the standard normal Stein equation is a first order linear differential equation, but this is not the case for the multivariate normal Stein equation for the case $d\geq 2$.

6.2.3 Locally dependent random variables

We now illustrate how Theorem 6.6 can be generalised to locally dependent random variables; the extension of Theorem 6.1 is straightforward and is omitted. In Lemma 6.16 we generalise the bound (6.21) for $|\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))|$ to the case that the X_i are locally dependent. The proof of Lemma 6.16 is a straightforward generalisation of the calculation used to prove (6.21), but leads to a more complicated bound.

We use a similar dependence structure to the one used in Section 4.5. We suppose that for each $i=1,\ldots,n$, there exist index sets $\{i\}\subseteq A_i^{(1)}\subseteq A_i^{(2)}\subseteq A_i^{(3)}\subseteq\{1,\ldots,n\}$ such that $X_i \perp\!\!\!\perp \sigma\{X_j\,:\,j\in A_i^{(1)}\}$, if $X_p\in\sigma\{X_j\,:\,j\in A_i^{(1)}\}$ then $X_p\perp\!\!\!\perp\sigma\{X_j\,:\,j\in A_i^{(2)}\}$, and if $X_q\in\sigma\{X_j\,:\,j\in A_i^{(2)}\}$ then $X_q\perp\!\!\!\perp\sigma\{X_j\,:\,j\in A_i^{(3)}\}$, where $\sigma\{X_j\}$ denotes the σ–algebra generated by X_j. Recalling the notation from Section 4.5, we write $X_i^{(k)}=\sum_{u\in A_i^{(k)}}X_u$ for $k=1,2,3$.

Lemma 6.16. *Let $X_1,\ldots,X_n$ be a sequence of mean zero random variables, that have a dependence structure as outlined above. Let $W=\sum_{i=1}^n X_i$ and suppose that the X_i are normalised so that $\mathbb{E}W^2=1$. Suppose that $g:\mathbb{R}\rightarrow\mathbb{R}$ is an even function. Then, provided that the fourth derivative of the solution f of the Stein equation (6.4) exists and that the M_k, which are defined below, exist, we have*

$$|\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))|\leq\sum_{k=1}^8 M_k+\left|\sum_{j=1}^n\left\{\mathbb{E}X_jX_j^{(1)}X_j^{(2)}-\frac{1}{2}\mathbb{E}X_j(X_j^{(1)})^2\right\}\right|(M_9+M_{10}+M_{11}),$$

where

$$\begin{aligned}
M_1 &= \frac{1}{6}\sum_{i=1}^n \sup_\theta \left|\mathbb{E}X_i(X_i^{(1)})^3 f^{(4)}(W_{1,\theta}^{(i)})\right|, \\
M_2 &= \frac{1}{2}\sum_{i=1}^n \sup_\theta \left|\mathbb{E}X_i X_i^{(1)}(X_i^{(2)}-X_i^{(1)})^2 f^{(4)}(W_{1,\theta}^{(i)})\right|, \\
M_3 &= \frac{1}{2}\sum_{i=1}^n \left|\mathbb{E}X_i X_i^{(1)}\right| \sup_\theta \left|\mathbb{E}((X_i^{(2)})^2 f^{(4)}(W_{1,\theta}^{(i)})\right|, \\
M_4 &= \frac{1}{2}\sum_{i=1}^n \sup_\theta \left|\mathbb{E}X_i(X_i^{(1)})^2(X_i^{(2)}-X_i^{(1)}) f^{(4)}(W_{2,\theta}^{(i)})\right|, \\
M_5 &= \frac{1}{2}\sum_{i=1}^n \left|\mathbb{E}X_i(X_i^{(1)})^2\right| \sup_\theta \left|\mathbb{E}X_i^{(2)} f^{(4)}(W_{3,\theta}^{(i)})\right|, \\
M_6 &= \sum_{i=1}^n |\mathbb{E}X_i X_i^{(1)}| \sup_\theta \left|\mathbb{E}X_i^{(2)}X_i^{(3)} f^{(4)}(W_{1,\theta}^{(i)})\right|, \\
M_7 &= \sum_{i=1}^n \sup_\theta \left|\mathbb{E}X_i X_i^{(1)}(X_i^{(2)}-X_i^{(1)})(X_i^{(3)}-X_i^{(2)}) f^{(4)}(W_{5,\theta}^{(i)})\right|, \\
M_8 &= \sum_{i=1}^n \left|\mathbb{E}X_i X_i^{(1)}(X_i^{(2)}-X_i^{(1)})\right| \sup_\theta \left|\mathbb{E}X_i^{(3)} f^{(4)}(W_{4,\theta}^{(i)})\right|, \\
M_9 &= \frac{1}{2}\sum_{i=1}^n \sup_\theta \left|\mathbb{E}X_i(X_i^{(1)})^2 \psi^{(3)}(W_{1,\theta}^{(i)})\right|, \\
M_{10} &= \sum_{i=1}^n \sup_\theta \left|\mathbb{E}X_i X_i^{(1)}(X_i^{(2)}-X_i^{(1)}) \psi^{(3)}(W_{2,\theta}^{(i)})\right|, \\
M_{11} &= \sum_{i=1}^n \left|\mathbb{E}X_i X_i^{(1)}\right| \sup_\theta \left|\mathbb{E}X_i^{(2)} \psi^{(3)}(W_{3,\theta}^{(i)})\right|,
\end{aligned}$$

and $W_{1,\theta}^{(i)} = W_i^{(1)} + \theta X_i^{(1)}$, $W_{2,\theta}^{(i)} = W_i^{(2)} + \theta(X_i^{(2)} - X_i^{(1)})$, $W_{3,\theta}^{(i)} = W_i^{(2)} + \theta X_i^{(2)}$, $W_{4,\theta}^{(i)} = W_i^{(3)} + \theta X_i^{(3)}$ *and* $W_{5,\theta}^{(i)} = W_i^{(3)} + \theta(X_i^{(3)} - X_i^{(2)})$ *for some* $\theta \in (0,1)$.

Proof. As has been the case throughout this chapter, we bound $|\mathbb{E}h(g(W)) - \mathbb{E}h(g(Z))|$ by bounding $\mathbb{E}f''(W) - \mathbb{E}Wf'(W)$. Through repeated use of Taylor expansions, and using that $\mathbb{E}X_i = 0$ and $\sum_{i=1}^n \mathbb{E}X_i X_i^{(1)} = \mathbb{E}W^2 = 1$, we have

$$\begin{aligned}
\mathbb{E}Wf'(W) &= \sum_{i=1} \mathbb{E}X_i f'(W) \\
&= \sum_{i=1}^n \mathbb{E}X_i f'(W_i^{(1)}) + \sum_{i=1}^n \mathbb{E}X_i X_i^{(1)} f''(W_i^{(1)}) + N_1 + R_1
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}f''(W_i^{(2)}) + N_1 + N_2 + R_1 + R_2 \\
&= \mathbb{E}f''(W) + N_1 + N_2 + N_3 + R_1 + R_2 + R_3,
\end{aligned}
$$

where

$$
\begin{aligned}
N_1 &= \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^2 f^{(3)}(W_i^{(1)}), \\
N_2 &= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)}) f^{(3)}(W_i^{(2)}), \\
N_3 &= -\sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}X_i^{(2)} f^{(3)}(W), \\
R_1 &= \frac{1}{6}\sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^3 f^{(4)}(W_{1,\theta_1}^{(i)}), \\
R_2 &= \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)})^2 f^{(4)}(W_{2,\theta_2}^{(i)}), \\
R_3 &= -\frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}(X_i^{(2)})^2 f^{(4)}(W_{3,\theta_3}^{(i)}),
\end{aligned}
$$

and $W_{1,\theta_1}^{(i)} = W_i^{(1)} + \theta_1 X_i^{(1)}$, $W_{2,\theta_2}^{(i)} = W_i^{(2)} + \theta_2 (X_i^{(2)} - X_i^{(1)})$, $W_{3,\theta_3}^{(i)} + W_i^{(2)} = \theta_3 X_i^{(2)}$, for some $\theta_k \in (0,1)$, $k = 1,2,3$.

We now deal with N_1, N_2 and N_3. We begin with N_1:

$$
N_1 = \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^2 \mathbb{E}f^{(3)}(W_i^{(2)}) + R_4 = \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^2 \mathbb{E}f^{(3)}(W) + R_4 + R_5,
$$

where

$$
\begin{aligned}
R_4 &= \frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^2 (X_i^{(2)} - X_i^{(1)}) f^{(4)}(W_{2,\theta_4}^{(i)}), \\
R_5 &= -\frac{1}{2}\sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^2 \mathbb{E}X_i^{(2)} f^{(4)}(W_{3,\theta_5}^{(i)}),
\end{aligned}
$$

and $W_{2,\theta_4}^{(i)} = W_i^{(2)} + \theta_4 (X_i^{(2)} - X_i^{(1)})$, $W_{3,\theta_5}^{(i)} + W_i^{(2)} = \theta_5 X_i^{(2)}$ for some $\theta_k \in (0,1)$, $k = 4,5$.

We defer bounding N_2 as this is somewhat more involved and now consider N_3:

$$
N_3 = -\sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}X_i^{(2)} \mathbb{E}f^{(3)}(W_i^{(3)}) + R_6 = R_6,
$$

where

$$R_6 = \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}X_i^{(2)} X_i^{(3)} f^{(4)}(W_{3,\theta_6}^{(i)}),$$

and $W_{3,\theta}^{(i)} = W_i^{(3)} + \theta X_i^{(3)}$ and $\theta_6 \in (0,1)$.

Finally, we consider N_2:

$$\begin{aligned} N_2 &= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)}) \mathbb{E}f^{(3)}(W_i^{(3)}) + R_7 \\ &= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)}) \mathbb{E}f^{(3)}(W) + R_7 + R_8, \end{aligned}$$

where

$$\begin{aligned} R_7 &= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)})(X_i^{(3)} - X_i^{(2)}) f^{(4)}(W_{5,\theta_7}^{(i)}), \\ R_8 &= -\sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)}) \mathbb{E}X_i^{(3)} f^{(4)}(W_{4,\theta_8}^{(i)}), \end{aligned}$$

and $W_{4,\theta}^{(i)} = W_i^{(3)} + \theta X_i^{(3)}$, $W_{5,\theta}^{(i)} = W_i^{(3)} + \theta(X_i^{(3)} - X_i^{(2)})$ and $\theta_k \in (0,1)$ for $k = 4, 5$.

As in the proof of Theorem 6.6, we use a symmetry argument to bound $\mathbb{E}f^{(3)}(W)$. To achieve such a bound we proceed as we did in the i.i.d. case. Since g is even, we have by Lemma 6.8 that $\mathbb{E}f^{(3)}(Z) = 0$, and therefore

$$\mathbb{E}f^{(3)}(W) = \mathbb{E}f^{(3)}(W) - \mathbb{E}f^{(3)}(Z) = \mathbb{E}\{\psi''(W) - W\psi'(W)\} = R_9 + R_{10} + R_{11},$$

where

$$\begin{aligned} R_9 &= \frac{1}{2} \sum_{i=1}^{n} \mathbb{E}X_i (X_i^{(1)})^2 \psi^{(3)}(W_{1,\theta_9}^{(i)}), \\ R_{10} &= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} (X_i^{(2)} - X_i^{(1)}) \psi^{(3)}(W_{2,\theta_{10}}^{(i)}), \\ R_{11} &= \sum_{i=1}^{n} \mathbb{E}X_i X_i^{(1)} \mathbb{E}X_i^{(2)} \psi^{(3)}(W_{3,\theta_{11}}^{(i)}), \end{aligned}$$

and $W_{1,\theta}^{(i)} = W_i^{(1)} + \theta X_i^{(1)}$, $W_{2,\theta}^{(i)} = W_i^{(2)} + \theta(X_i^{(2)} - X_i^{(1)})$, $W_{3,\theta}^{(i)} + W_i^{(2)} = \theta X_i^{(2)}$, with $\theta_k \in (0,1)$ for $k = 9, 10, 11$. The bound (6.16) now follows from bounding the remainder terms $R_1 - R_{11}$ by the triangle inequality. □

We now consider how Lemma 6.16 can be used to generalise Theorem 6.6 to locally dependent random variables. We could use the expressions (6.22) and (6.23) for $f^{(4)}(w)$ and $\psi^{(3)}(w)$, respectively, to obtain a bound for $\mathbb{E}h(g(W))-\mathbb{E}h(g(Z))$ in terms of $f'(w)$, $f''(w)$, $f^{(3)}(w)$ and $\psi'(w)$, thereby enabling us to impose the same conditions on g as those of Theorem 6.6. For g with polynomial growth we would use the bounds for $f'(w)$, $f''(w)$, $f^{(3)}(w)$ and $\psi'(w)$ that are given in Lemmas 6.4 and 6.10. We would bound the relevant expectations by performing similar calculations to those used in the proof of Lemma 6.5. However, these bounds, would be even more complicated than in the i.i.d. case.

6.3 Applications

In this section we apply the results of the previous sections to obtain some limit theorems for statistics that are asymptotically Chi, χ^2, Variance-Gamma and Product Normal distributed. Our final application is an explicit bound on the error in approximating Friedman's statistic by its limiting distribution.

Before stating our limit theorems, we note that the results of Sections 6.1 and 6.2 have helped shed light on the $O(n^{-1})$ bounds for χ^2 approximation of Pickett [55], as well as our $O(m^{-1}+n^{-1})$ bounds for Variance-Gamma approximation. For instance, the χ^2 statistic $\sum_{k=1}^d W_k^2$ that was studied by Pickett [55], and the asymptotically Variance-Gamma distributed statistic (4.2) are both of the form $g(\mathbf{W})$, where g is an even function with derivatives of polynomial growth. Hence, in light of Theorem 6.6 and Lemmas 6.8 and Lemma 6.12, these faster bounds on the convergence rates were no accident and, in fact, are to be expected.

In Theorems 6.18 and 6.20 we obtain $O(n^{-1})$ and $O(m^{-1}+n^{-1})$ bounds on the convergence rate for these asymptotically χ^2 and Variance-Gamma distributed statistics. These bounds follows from applying Theorem 6.6 and Lemma 6.12, respectively. We also obtain a bound of order $n_1^{-1}+\cdots+n_r^{-1}$ for the error in approximating the statistic (5.1), which is given by $\prod_{k=1}^r(\frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k}X_{ik})$,by its limiting distribution, for the case the r is even. Again, this to be expected, as this statistic is of the form $g(\mathbf{W})$, where g is an even function with derivatives of polynomial growth

6.3.1 A Chi distribution limit theorem

Suppose $Z\sim N(0,1)$, then $|Z|$ has a Chi distribution with p.d.f. $\sqrt{\frac{2}{\pi}}\mathrm{e}^{-x^2/2}$, for $x\geq 0$. Let $X,X_1,X_2,\ldots$ be a collection of i.i.d. random variables with zero mean and unit variance. Let

$W = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$. Then, since the function $g(x) = |x|$ is continuous, it follows that $|W| \xrightarrow{\mathcal{D}} |Z|$. Now $g(x) = |x|$ is an even function with polynomial growth rate, but is not twice differentiable, so we cannot apply Theorem 6.6 to obtain a $O(n^{-1})$ bound on the convergence. But g is absolutely continuous and has a derivative that is bounded almost everywhere, and we can therefore apply Theorem 6.1 to obtain a bound of order $n^{-1/2}$.

Proposition 6.17. *Let the $X, X_1, X_2\ldots$ be a collection of random variables as defined above, with $\mathbb{E}X^4 < \infty$. Then, for absolutely continuous h, we have*

$$|\mathbb{E}h(|W|) - \mathbb{E}h(|Z|)| \leq \frac{\|h'\|}{\sqrt{n}}\bigg\{\bigg(1 + \frac{1}{\sqrt{2\pi}}\bigg)(2 + \mathbb{E}|X|^3) + \frac{2 + \mathbb{E}X^4}{\sqrt{2\pi n}}\bigg\}.$$

Proof. We have that $|g'(w)| = 1$ almost everywhere, and so the result follows on applying bound (6.3) of Lemma 6.1 with $A = 1$ and $B = 0$. ☐

6.3.2 A χ^2 limit theorem

We now present a $\chi^2_{(d)}$ limit theorem. A limit theorem for the $\chi^2_{(1)}$ distribution follow immediately from Theorem 6.6 and we extend to the the case of general d be using a Lindeberg scheme. Alternatively, we could have obtained $\chi^2_{(d)}$ limit theorems by taking $g(\mathbf{w}) = \sum_{k=1}^d w_k^2$ and then used Lemma 6.12 to arrive at a $O(dn^{-1})$ bound (we would take $n_1 = \cdots = n_d = n$). However, this would have meant that we would have to impose stronger conditions on the the derivatives of the test function h.

Theorem 6.18. *Let $X, X_{1k}, X_{2k}, \ldots, X_{nk}$, for $k \leq d$, be a collection of i.i.d. random variables with mean zero, unit variance and $\mathbb{E}|X|^7 < \infty$. Define $W_k = \frac{1}{\sqrt{n}}\sum_{i=1}^n X_{ik}$ and let $S_d = \sum_{k=1}^d W_k^2$. Suppose that $h \in C^2(\mathbb{R})$, then*

$$|\mathbb{E}h(S_d) - \chi^2_{(d)}h| \leq \frac{d}{12n}(\alpha_1\|h'\| + \alpha_2\|h''\|), \tag{6.40}$$

where $\chi^2_{(d)}h$ denotes the expectation of $h(Z)$, for $Z \sim \chi^2_{(d)}$, and

$$\begin{aligned}\alpha_1 = 180 + 9\sqrt{2\pi} + \frac{288\sqrt{2}}{\sqrt{\pi}} + \frac{9\sqrt{2}}{\sqrt{n}}\bigg(1 + \frac{32}{\sqrt{\pi}}\bigg)\mathbb{E}|X|^3 + \bigg(20 + \sqrt{2\pi} + \frac{32\sqrt{2}}{\sqrt{\pi}}\bigg)\mathbb{E}X^4 \\ + \bigg(\sqrt{2\pi} + \frac{32\sqrt{2}}{\sqrt{\pi}}\bigg)\frac{\mathbb{E}|X|^5}{\sqrt{n}} + |\mathbb{E}X^3|\bigg\{102 + \frac{15\sqrt{2\pi}}{2}\bigg(1 + \frac{1}{\sqrt{n}}\bigg) + \frac{144}{\sqrt{2}}\sqrt{\pi}\bigg(1 + \frac{1}{\sqrt{n}}\bigg) \\ + \bigg(36 + 3\sqrt{2\pi} + \frac{48\sqrt{2}}{\sqrt{\pi}} + \frac{8}{n}\bigg)\mathbb{E}|X|^3 + \bigg(3\sqrt{2\pi} + \frac{48\sqrt{2}}{\sqrt{\pi n}}\bigg)\mathbb{E}X^4 + \frac{4\mathbb{E}|X|^5}{n}\bigg\},\end{aligned}$$

$$\alpha_2 = 8\bigg\{2\sqrt{2\pi}(9+\mathbb{E}X^4)\bigg(3+\frac{\mathbb{E}X^4}{n}\bigg)^{3/4}+18+9\sqrt{2\pi}+\frac{9\sqrt{2\pi}\mathbb{E}|X|^3}{\sqrt{n}}+\bigg(2+\sqrt{2\pi}+\frac{18}{n}\bigg)\mathbb{E}X^4$$
$$+\frac{\sqrt{2\pi}}{\sqrt{n}}\bigg(1+\frac{18}{n}\bigg)\mathbb{E}|X|^7+\frac{2\mathbb{E}X^6}{n}+\frac{2\sqrt{2\pi}\mathbb{E}|X|^7}{n\sqrt{n}}\bigg\}+12|\mathbb{E}X^3|\bigg\{12(2+\mathbb{E}|X|^3)\bigg(3+\frac{\mathbb{E}X^4}{n}\bigg)$$
$$+2\sqrt{2\pi}(2\mathbb{E}|X|^3+5)\bigg(3+\frac{\mathbb{E}X^4}{n}\bigg)^{3/4}+54+5\sqrt{2\pi}\bigg(1+\frac{1}{\sqrt{n}}\bigg)+\bigg(26+2\sqrt{2\pi}+\frac{42}{n}\bigg)\mathbb{E}|X|^3$$
$$+\frac{\sqrt{2\pi}}{n}\bigg(2+\frac{5}{\sqrt{n}}\bigg)\mathbb{E}X^4+\frac{4}{n}\bigg(5+\frac{6}{n}\bigg)\mathbb{E}|X|^5+\frac{2\sqrt{2\pi}\mathbb{E}X^6}{n\sqrt{n}}+\frac{12\mathbb{E}|X|^7}{n^2}\bigg\}.$$

Proof. The bound, in the case $d=1$, follows by taking $g(x)=x^2$, and therefore $A_1=0$, $A_2=2$, $B_1=2$, $B_2=0$ and $k_1=1$, in Theorem 6.6. In obtaining this bound we used the inequality $\mathbb{E}|W|^3 \leq \{\mathbb{E}W^4\}^{3/4} < (3+\frac{\mathbb{E}X^4}{n})^{3/4}$ (see Lemma 4.6).

We now prove that the bound holds for general d by using a Lindeberg scheme. Let $Z_1,\ldots,Z_d$ be independent standard normal variables, so that $\sum_{i=1}^d Z_i^2 \sim \chi^2_{(d)}$. We define $U_0=\sum_{j=1}^d W_j^2$, $U_d=\sum_{i=1}^d Z_j^2$ and $U_k=\sum_{i=1}^k Z_i^2+\sum_{j=k+1}^d W_j^2$, for $k=1,\ldots,d-1$. With this notation, we have

$$\begin{aligned}|\mathbb{E}h(S_d)-\chi^2_{(d)}h| &= \bigg|\sum_{k=1}^d \mathbb{E}h(U_k)-\mathbb{E}h(U_{k-1})\bigg| \\ &\leq \sum_{k=1}^d |\mathbb{E}h(U_k)-\mathbb{E}h(U_{k-1})| \\ &= \sum_{k=1}^d |\mathbb{E}(\mathbb{E}(\{h(U_k)-h(U_{k-1})\}\mid Z_1,\ldots,Z_{k-1},W_{k+1},\ldots,W_d))| \\ &\leq \sum_{k=1}^d \mathbb{E}|\mathbb{E}(\{h(U_k)-h(U_{k-1})\}\mid Z_1,\ldots,Z_{k-1},W_{k+1},\ldots,W_d)|. \end{aligned} \tag{6.41}$$

Define $g_k(x)=h(x+\sum_{i=1}^{k-1}z_i+\sum_{j=k+1}^d w_j)$, where the z_i and w_j denote the variables that we have conditioned on.. Using the bound (6.40) for the case that $d=1$, we can bound the conditional expectation in expression (6.41) as follows

$$\begin{aligned}|\mathbb{E}h(S_d)-\chi^2_{(d)}h| &\leq \sum_{k=1}^d \mathbb{E}|\mathbb{E}((g_k(W_k)-g_k(Z_k))\mid Z_1,\ldots,Z_{k-1},W_{k+1},\ldots,W_d))| \\ &\leq \sum_{k=1}^d \mathbb{E}\bigg[\frac{1}{n}(\alpha_1\|g_k'\|+\alpha_2\|g_k''\|)\bigg] \\ &= \frac{d}{n}(\alpha_1\|h'\|+\alpha_2\|h''\|),\end{aligned}$$

where the final equality follows because $\|g_k'\|=\|h'\|$ and $\|g_k''\|=\|h''\|$ for all k. This completes the proof of the bound for general d. □

Remark 6.19. The bound (6.40) improves on the bound given in Theorem 4.7 of Pickett [55] in the sense that our bound requires the existence of h'', whereas Pickett required the third order derivative to exist. However, Pickett's bound performs better for large d; Pickett's bound is of order $d^{1/2}n^{-1}$, whereas ours is of order dn^{-1}.

6.3.3 Variance-Gamma and Product Normal distribution limit theorems

We now use Lemmas 6.12 and 6.13 with $g(x,y) = xy$ and a Lindeberg scheme to prove a Variance-Gamma limit theorem.

Theorem 6.20. *Let $X, X_{1k}, X_{2k}, \ldots, X_{mk}$, for $1 \leq k \leq r$, be a collection of i.i.d. random variables with mean zero, unit variance and $\mathbb{E}|X|^5 < \infty$. Similarly, let $Y, Y_{1k}, Y_{2k}, \ldots, Y_{nk}$ be a collection of i.i.d. random variable with mean zero, unit variance and $\mathbb{E}|Y|^5 < \infty$. Suppose further that the σ-fields $\sigma\{X_{ik} : i = 1, \ldots, m,\ k = 1, \ldots r\}$ and $\sigma\{Y_{jk} : j = 1, \ldots, n,\ k = 1, \ldots r\}$ are independent. Define $T_r = \frac{1}{\sqrt{mn}} \sum_{i,j,k=1}^{m,n,r} X_{ik}Y_{jk}$. Suppose that $h \in C_b^4(\mathbb{R})$, then*

$$|\mathbb{E}h(T_r) - \mathrm{VG}_{1,0}^{r,0}h| \leq \frac{r}{12}\Big(\beta_2(X,Y)\|h''\| + \beta_3(X,Y)\|h^{(3)}\| + \beta_4(X,Y)\|h^{(4)}\|\Big), \tag{6.42}$$

where $\mathrm{VG}_{1,0}^{r,0}h$ denotes the expectation of $h(Z)$ for $Z \sim \mathrm{VG}(r,0,1,0)$, and

$$\begin{aligned}
\beta_2(X,Y) &= \frac{1}{10\sqrt{mn}}\bigg[|\mathbb{E}X^3|(\mathbb{E}|Y|^3 + 2) + |\mathbb{E}Y^3|(\mathbb{E}|X|^3 + 2)\bigg], \\
\beta_3(X,Y) &= \frac{4\sqrt{2}}{9\sqrt{\pi}m}(\mathbb{E}X^4 + 3|\mathbb{E}X^3| + 9)\bigg(\frac{2\sqrt{2}}{\sqrt{\pi}} + \bigg(3 + \frac{\mathbb{E}Y^4}{n}\bigg)^{3/4}\bigg) \\
&\quad + \frac{4\sqrt{2}}{9\sqrt{\pi}n}(\mathbb{E}Y^4 + 3|\mathbb{E}Y^3| + 9)\bigg(\frac{2\sqrt{2}}{\sqrt{\pi}} + \bigg(3 + \frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg) \\
&\quad + \frac{1 + \sqrt{\frac{\pi}{2}} + \sqrt{\frac{2}{\pi}}}{5\sqrt{mn}}\bigg\{|\mathbb{E}X^3|\bigg[2\bigg(1 + \sqrt{\frac{\pi}{2}} + \sqrt{\frac{2}{\pi}} + \frac{1}{\sqrt{n}}\bigg) + \bigg(1 + \sqrt{\frac{\pi}{2}} + \sqrt{\frac{2}{\pi}}\bigg)\mathbb{E}|Y|^3 \\
&\quad + \frac{\mathbb{E}Y^4}{\sqrt{n}}\bigg] + |\mathbb{E}Y^3|\bigg[2\bigg(1 + \sqrt{\frac{\pi}{2}} + \sqrt{\frac{2}{\pi}} + \frac{1}{\sqrt{m}}\bigg) + \bigg(1 + \sqrt{\frac{\pi}{2}} + \sqrt{\frac{2}{\pi}}\bigg)\mathbb{E}|X|^3 + \frac{\mathbb{E}X^4}{\sqrt{m}}\bigg]\bigg\}, \\
\beta_4(X,Y) &= \frac{27|\mathbb{E}X^3|}{20m}\bigg(9 + \frac{\mathbb{E}Y^4}{n}\bigg) + \frac{27|\mathbb{E}Y^3|}{20n}\bigg(9 + \frac{\mathbb{E}X^4}{m}\bigg) + \frac{4}{5\sqrt{mn}}\bigg[|\mathbb{E}X^3|\bigg(10 + \bigg(5 + \frac{4}{n}\bigg)\mathbb{E}|Y|^3 \\
&\quad + \frac{2\mathbb{E}|Y|^5}{n}\bigg) + |\mathbb{E}Y^3|\bigg(10 + \bigg(5 + \frac{4}{m}\bigg)\mathbb{E}|X|^3 + \frac{2\mathbb{E}|X|^5}{m}\bigg)\bigg].
\end{aligned}$$

Proof. We prove the result for case $r = 1$; the result for general r follows by a Lindeberg scheme that is similar to one used to prove Theorem 6.18. We will prove the result for $r = 1$

by using Lemma 6.12 with $g(u,v)=uv$. Let $U=\frac{1}{\sqrt{m}}\sum_{i=1}^m X_i$ and $V=\frac{1}{\sqrt{n}}\sum_{j=1}^n Y_j$. Define $U_i=U-\frac{1}{\sqrt{m}}X_i$, $V_i=V-\frac{1}{\sqrt{n}}Y_j$, and $U_\theta^{(i)}=U_i+\frac{\theta}{\sqrt{m}}X_i$ and $V_\theta^{(j)}=V_i+\frac{\theta}{\sqrt{n}}Y_j$ for some $\theta\in(0,1)$. From the bound (6.34) of Lemma 6.12 we have the following bound on $|\mathbb{E}h(T_1)-\mathrm{VG}^{1,0}_{1,0}h|$:

$$\begin{aligned}
&|\mathbb{E}h(T_1)-\mathrm{VG}^{1,0}_{1,0}h|\\
&\le\frac{1}{6m}\bigg[\sup_\theta\mathbb{E}\bigg|X_i^4\frac{\partial^4 f}{\partial u^4}(U_\theta^{(i)},V)\bigg|+9\sup_\theta\mathbb{E}\bigg|X_i^2\frac{\partial^4 f}{\partial u^4}(U_\theta^{(i)},V)\bigg|+3|\mathbb{E}X^3|\sup_\theta\mathbb{E}\bigg|X_i\frac{\partial^4 f}{\partial u^4}(U_\theta^{(i)},V)\bigg|\bigg]\\
&\quad+\frac{1}{6n}\bigg[\sup_\theta\mathbb{E}\bigg|Y_j^4\frac{\partial^4 f}{\partial v^4}(U,V_\theta^{(j)})\bigg|+9\sup_\theta\mathbb{E}\bigg|Y_j^2\frac{\partial^4 f}{\partial v^4}(U,V_\theta^{(j)})\bigg|+3|\mathbb{E}Y^3|\sup_\theta\mathbb{E}\bigg|Y_j\frac{\partial^4 f}{\partial v^4}(U,V_\theta^{(j)})\bigg|\bigg]\\
&\quad+\frac{|\mathbb{E}X^3|}{4\sqrt{mn}}\bigg[\sup_\theta\mathbb{E}\bigg|X_i^3\frac{\partial^3\psi_u}{\partial u^3}(U_\theta^{(i)},V)\bigg|+2\sup_\theta\mathbb{E}\bigg|X_i\frac{\partial^3\psi_u}{\partial u^3}(U_\theta^{(i)},V)\bigg|+\sup_\theta\mathbb{E}\bigg|X_i^3\frac{\partial^3\psi_v}{\partial u^3}(U_\theta^{(i)},V)\bigg|\\
&\quad+2\sup_\theta\mathbb{E}\bigg|X_i\frac{\partial^3\psi_v}{\partial u^3}(U_\theta^{(i)},V)\bigg|\bigg]+\frac{|\mathbb{E}Y^3|}{4\sqrt{mn}}\bigg[\sup_\theta\mathbb{E}\bigg|Y_j^3\frac{\partial^3\psi_u}{\partial v^3}(U,V_\theta^{(j)})\bigg|+2\sup_\theta\mathbb{E}\bigg|Y_j\frac{\partial^3\psi_u}{\partial v^3}(U_\theta^{(j)},V)\bigg|\\
&\quad+\sup_\theta\mathbb{E}\bigg|Y_j^3\frac{\partial^3\psi_v}{\partial v^3}(U,V_\theta^{(j)})\bigg|+2\sup_\theta\mathbb{E}\bigg|Y_j\frac{\partial^3\psi_v}{\partial v^3}(U,V_\theta^{(j)})\bigg|\bigg]. \qquad (6.43)
\end{aligned}$$

We now obtain bounds for the expectations on the right-hand side of (6.43), and we begin by bounding the relevant partial derivatives of f and ψ. Using (6.37) we have

$$\bigg|\frac{\partial^4 f}{\partial u^4}(u,v)\bigg|\le\|h^{(3)}\|\int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-4s}}{\sqrt{1-\mathrm{e}^{-2s}}}|(\mathrm{e}^{-s}v+y\sqrt{1-\mathrm{e}^{-2s}})^3||x|\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s. \qquad (6.44)$$

We note that for $p\ge1$ we have

$$|\mathrm{e}^{-s}v+y\sqrt{1-\mathrm{e}^{-2s}}|^p\le2^{p-1}(\mathrm{e}^{-ps}|v|^p+(1-\mathrm{e}^{-2s})^{p/2}|y|^p)\le2^{p-1}(|v|^p+|y|^p).$$

Applying this inequality to the right-hand side of (6.44) gives

$$\bigg|\frac{\partial^4 f}{\partial u^4}(u,v)\bigg|\le\|h^{(3)}\|\int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-4s}}{\sqrt{1-\mathrm{e}^{-2s}}}\cdot4(|v|^3+|y|^3)|x|\phi(x)\phi(y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s.$$

We now evaluate the integral in the above expression using the formulas $\int_{-\infty}^\infty|x|\phi(x)\,\mathrm{d}x=\mathbb{E}|Z|=\sqrt{\frac{2}{\pi}}$, $\int_{-\infty}^\infty|x|^3\phi(x)\,\mathrm{d}x=\mathbb{E}|Z|^3=\frac{2\sqrt{2}}{\sqrt{\pi}}$ and $\int_0^\infty\frac{\mathrm{e}^{-4s}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s=\frac{\sqrt{\pi}\Gamma(2)}{2\Gamma(5/2)}=\frac{2}{3}$ (see (2.21)) to obtain

$$\bigg|\frac{\partial^4 f}{\partial u^4}(u,v)\bigg|\le\frac{8\sqrt{2}}{3\sqrt{\pi}}\|h^{(3)}\|\bigg(|v|^3+\frac{2\sqrt{2}}{\sqrt{\pi}}\bigg).$$

By symmetry we have

$$\bigg|\frac{\partial^4 f}{\partial v^4}(u,v)\bigg|\le\frac{8\sqrt{2}}{3\sqrt{\pi}}\|h^{(3)}\|\bigg(|u|^3+\frac{2\sqrt{2}}{\sqrt{\pi}}\bigg).$$

We now use (6.38) to bound the partial derivatives of ψ. We have

$$\begin{aligned}\left|\frac{\partial^3\psi_u}{\partial u^3}(u,v)\right| &\leq \int_0^\infty\int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-6s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-5t}}{\sqrt{1-\mathrm{e}^{-2t}}}\|h^{(4)}\|(\mathrm{e}^{-s-t}v+x_2\mathrm{e}^{-t}\sqrt{1-\mathrm{e}^{2s}} \\ &\quad + y_2\sqrt{1-\mathrm{e}^{-2t}})^4|x_1||y_1|\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\,\mathrm{d}t \\ &\leq 27\|h^{(4)}\|\int_0^\infty\int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-6s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-5t}}{\sqrt{1-\mathrm{e}^{-2t}}}(v^4+x_2^4+y_2^4) \\ &\quad \times|x_1||y_1|\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}y_1\,\mathrm{d}y_2\,\mathrm{d}s\,\mathrm{d}t,\end{aligned}$$

where to obtain the second inequality we used that, for $p\geq 1$, $|\mathrm{e}^{-s-t}v+x_2\mathrm{e}^{-t}\sqrt{1-\mathrm{e}^{2s}}+y_2\sqrt{1-\mathrm{e}^{-2t}}|^p\leq 3^{p-1}(v^4+x_2^4+y_2^4)$. Evaluating the integral using that $\mathbb{E}|Z|=\sqrt{\frac{2}{\pi}}$, $\mathbb{E}Z^4=3$ and the definite integral formulas $\int_0^\infty\frac{\mathrm{e}^{-6s}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s=\frac{\sqrt{\pi}\Gamma(3)}{2\Gamma(7/2)}=\frac{8}{15}$ and $\int_0^\infty\frac{\mathrm{e}^{-5s}}{\sqrt{1-\mathrm{e}^{-2s}}}\,\mathrm{d}s=\frac{\sqrt{\pi}\Gamma(5/2)}{2\Gamma(3)}=\frac{3\pi}{16}$ (see (2.21)) gives

$$\left|\frac{\partial^3\psi_u}{\partial u^3}(u,v)\right|\leq\frac{27}{5}\|h^{(4)}\|(v^4+6).$$

By symmetry, we have

$$\left|\frac{\partial^3\psi_v}{\partial v^3}(u,v)\right|\leq\frac{27}{5}\|h^{(4)}\|(u^4+6).$$

Using a similar calculation, we have

$$\begin{aligned}\left|\frac{\partial^3\psi_v}{\partial u^3}(u,v)\right| &\leq \int_0^\infty\int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-6s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-5t}}{\sqrt{1-\mathrm{e}^{-2t}}}[\|h^{(4)}\|z_u^2z_v^2+4\|h^{(3)}\||z_uz_v|+2\|h''\|] \\ &\quad\times|x_1||y_2|\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}t \\ &\leq \int_0^\infty\int_0^\infty\int_{\mathbb{R}^2}\frac{\mathrm{e}^{-6s}}{\sqrt{1-\mathrm{e}^{-2s}}}\frac{\mathrm{e}^{-5t}}{\sqrt{1-\mathrm{e}^{-2t}}}[\|9h^{(4)}\|(u^2+x_1^2+y_1^2)(v^2+x_2^2+y_2^2) \\ &\quad+4\|h^{(3)}\|(|u|+|x_1|+|y_1|)(|v|+|x_2|+|y_2|)+2\|h''\|] \\ &\quad\times|x_1||y_2|\phi(x_1)\phi(x_2)\phi(y_1)\phi(y_2)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}t \\ &=\frac{\pi}{10}\bigg[9\|h^{(4)}\|\bigg(\sqrt{\frac{2}{\pi}}u^2+\frac{2\sqrt{2}}{\sqrt{\pi}}+\sqrt{\frac{2}{\pi}}\bigg)\bigg(\sqrt{\frac{2}{\pi}}v^2+\frac{2\sqrt{2}}{\sqrt{\pi}}+\sqrt{\frac{2}{\pi}}\bigg) \\ &\quad+4\|h^{(3)}\|\bigg(\sqrt{\frac{2}{\pi}}|u|+1+\frac{2}{\pi}\bigg)\bigg(\sqrt{\frac{2}{\pi}}|v|+1+\frac{2}{\pi}\bigg)+\frac{4}{\pi}\|h''\|\bigg] \\ &=\frac{1}{5}\bigg[9\|h^{(4)}\|(u^2+3)(v^2+3) \\ &\quad+4\|h^{(3)}\|\bigg(u^2+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)\bigg(v^2+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)+2\|h''\|\bigg].\end{aligned}$$

Again, by symmetry we have

$$\left|\frac{\partial^3\psi_u}{\partial v^3}(u,v)\right|\leq\frac{1}{5}\bigg[9\|h^{(4)}\|(u^2+3)(v^2+3)+4\|h^{(3)}\|\bigg(|u|+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)\bigg(|v|+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)+2\|h''\|\bigg].$$

With these bounds on the partial derivatives of f and ψ we are able to obtain the following bounds, for $p \geq 0$,

$$\begin{aligned}
\mathbb{E}\bigg|X_i^p\frac{\partial^4 f}{\partial u^4}(U_\theta^{(i)},V)\bigg| &\leq \frac{8\sqrt{2}}{3\sqrt{\pi}}\|h^{(3)}\|\mathbb{E}|X|^p\bigg(\frac{2\sqrt{2}}{\sqrt{\pi}}+\bigg(3+\frac{\mathbb{E}Y^4}{n}\bigg)^{3/4}\bigg), \\
\mathbb{E}\bigg|Y_j^p\frac{\partial^4 f}{\partial v^4}(U,V_\theta^{(j)})\bigg| &\leq \frac{8\sqrt{2}}{3\sqrt{\pi}}\|h^{(3)}\|\mathbb{E}|Y|^p\bigg(\frac{2\sqrt{2}}{\sqrt{\pi}}+\bigg(3+\frac{\mathbb{E}X^4}{m}\bigg)^{3/4}\bigg), \\
\mathbb{E}\bigg|X_i^p\frac{\partial^3 \psi_u}{\partial u^3}(U_\theta^{(i)},V)\bigg| &\leq \frac{27}{5}\|h^{(4)}\|\mathbb{E}|X|^p\bigg(9+\frac{\mathbb{E}Y^4}{n}\bigg), \\
\mathbb{E}\bigg|Y_j^p\frac{\partial^3 \psi_v}{\partial v^3}(U,V_\theta^{(j)})\bigg| &\leq \frac{27}{5}\|h^{(4)}\|\mathbb{E}|Y|^p\bigg(9+\frac{\mathbb{E}X^4}{m}\bigg), \\
\mathbb{E}\bigg|X_i^p\frac{\partial^3 \psi_v}{\partial u^3}(U_\theta^{(i)},V)\bigg| &\leq \frac{2}{5}\bigg[36\|h^{(4)}\|\bigg(\mathbb{E}|X|^p+\frac{\mathbb{E}|X|^{p+2}}{m}\bigg)+2\bigg(1+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)\|h^{(3)}\| \\
&\quad\times\bigg(\bigg(1+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)\mathbb{E}|X|^p+\frac{\mathbb{E}|X|^{p+1}}{\sqrt{m}}\bigg)+\|h''\|\bigg], \\
\mathbb{E}\bigg|Y_j^p\frac{\partial^3 \psi_u}{\partial v^3}(U,V_\theta^{(j)})\bigg| &\leq \frac{2}{5}\bigg[36\|h^{(4)}\|\bigg(\mathbb{E}|Y|^p+\frac{\mathbb{E}|Y|^{p+2}}{n}\bigg)+2\bigg(1+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)\|h^{(3)}\| \\
&\quad\times\bigg(\bigg(1+\sqrt{\frac{\pi}{2}}+\sqrt{\frac{2}{\pi}}\bigg)\mathbb{E}|Y|^p+\frac{\mathbb{E}|Y|^{p+1}}{\sqrt{n}}\bigg)+\|h''\|\bigg].
\end{aligned}$$

These bounds follow from first using the bounds on the derivatives of f and ψ and then bounding the expectations using the same approach that was used to obtain the bounds of Lemma 6.5. Substituting these bounds into (6.43) completes the proof of the bound (6.42) for the case $r=1$. The generalisation to the case of general $r \geq 1$ follows from a similar Lindeberg scheme to the one used in the proof of Theorem 6.18. □

We now prove two Product Normal limit theorems. Again, to simplify the calculations, we use a Lindeberg scheme, rather than directly using Lemma 6.12.

Theorem 6.21. *Let $X_k, X_{1k}, X_{2k}, \ldots, X_{n_k k}$, for $1 \leq k \leq r$, be collections of i.i.d. random variables with mean zero, unit variance and $\mathbb{E}|X_k|^3 < \infty$. Let $W_k = \frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k} X_{ik}$ and define $W = \prod_{k=1}^r W_k$. Suppose that $h : \mathbb{R} \to \mathbb{R}$ is absolutely continuous, then*

$$|\mathbb{E}h(W) - \mathrm{PN}_r^1 h| \leq \sum_{k=1}^r \frac{\|h'\|}{\sqrt{n_k}}(2+\mathbb{E}|X_k|^3),$$

where $\mathrm{PN}_r^1 h$ denotes the expectation of $h(Z)$ for $Z \sim \mathrm{PN}(r,1)$.

Proof. The result in the case $r = 1$ was proved in Example 2.6. We extend this result to general r by using a Lindeberg scheme. We define $V_0 = \prod_{j=1}^r W_j$, $V_r = \prod_{i=1}^r Z_j$ and $V_k =$

$\prod_{i=1}^{k}\prod_{j=k+1}^{r} W_jZ_i$, for $k=1,\ldots,r-1$. With this notation we have

$$\begin{aligned}|\mathbb{E}h(W)-\mathrm{PN}_r^1h| &= |\mathbb{E}h(\textstyle\prod_{k=1}^r W_k)-\mathbb{E}h(\prod_{k=1}^r Z_k)|\\ &=\left|\sum_{k=1}^r \mathbb{E}h(V_k)-\mathbb{E}h(V_{k-1})\right|\\ &\leq \sum_{k=1}^r |\mathbb{E}h(V_k)-\mathbb{E}h(V_{k-1})|\\ &=\sum_{k=1}^r |\mathbb{E}(\mathbb{E}((h(V_k)-h(V_{k-1}))\mid Z_1,\ldots,Z_{k-1},W_{k+1},\ldots,W_r))|\\ &\leq \sum_{k=1}^r \mathbb{E}|\mathbb{E}((h(V_k)-h(V_{k-1}))\mid Z_1,\ldots,Z_{k-1},W_{k+1},\ldots,W_r)|.\end{aligned}$$

Define $g_k(x)=h((\prod_{i=1}^{k-1}\prod_{j=k+1}^r w_jz_i)x)$, where the w_j and z_i denote the variables we have conditioned on. Using the bound for the case that $r=1$, allows us to bound the conditional expectation, and we have

$$\begin{aligned}|\mathbb{E}h(W)-\mathrm{PN}_r^1h| &\leq \sum_{k=1}^r \mathbb{E}|\mathbb{E}((g_k(W_k)-g_k(Z_k))\mid Z_1,\ldots,Z_{k-1},W_{k+1},\ldots,W_r))|\\ &\leq \sum_{k=1}^r \mathbb{E}\left[\frac{|\prod_{i=1}^{k-1}\prod_{j=k+1}^r W_jZ_i|\|h'\|}{\sqrt{n_k}}(2+\mathbb{E}|X_k|^3)\right],\end{aligned}$$

where the final equality follows because $\|g_k'\|=\prod_{i=1}^{k-1}\prod_{j=k+1}^r|w_jz_i|\|h'\|$. We now use that $\mathbb{E}|Z|\leq\sqrt{\mathbb{E}Z^2}=1$ and $\mathbb{E}|W_k|\leq\sqrt{\mathbb{E}W_k^2}=1$, by the Cauchy-Schwarz inequality, to deduce that

$$|\mathbb{E}h(W)-\mathrm{PN}_r^1h|\leq\sum_{k=1}^r\frac{\|h'\|}{\sqrt{n_k}}(2+\mathbb{E}|X_k|^3),$$

as required. □

We arrive at the following theorem by using our bound from Theorem 6.20 and a Lindeberg scheme. This theorem demonstrates the power of the multivariate normal approach that we have used in this chapter. We have been unable to establish smoothness estimates for the solution of the $\mathrm{PN}(r,\sigma^2)$ Stein equation (5.8) for $r\geq 3$. Therefore an alternative approach is required and we are able to use the multivariate normal Stein equation to arrive at a $O(n_1^{-1}+\cdots+n_d^{-1})$ bound:

Theorem 6.22. *Let $X_k, X_{1k}, X_{2k},\ldots,X_{nk}$, for $1\leq k\leq 2r$, be a collection of i.i.d. random variables with mean zero, unit variance and $\mathbb{E}|X_k|^5<\infty$. Let $W_k=\frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k}X_{ik}$ and define*

$W = \prod_{k=1}^{2r} W_k$. *Suppose that* $h \in C_b^4(\mathbb{R})$, *then*

$$|\mathbb{E}h(W) - \mathrm{PN}_{2r}^1 h| \leq \gamma_2 \|h''\| + \gamma_3 \|h^{(3)}\| + \gamma_4 \|h^{(4)}\|,$$

where $\mathrm{PN}_{2r}^1 h$ *denotes the expectation of* $h(Z)$ *for* $Z \sim \mathrm{PN}(2r, 1)$, *and*

$$\begin{aligned} \gamma_2 &= \frac{1}{10} \sum_{k=1}^{r} \beta_2(X_{2k-1}, X_{2k}), \\ \gamma_3 &= \sum_{k=1}^{r} \bigg\{ \prod_{j \neq k}^{r} \bigg(3 + \frac{\mathbb{E}X_{2j-1}^4}{n_{2j-1}} \bigg)^{3/4} \bigg(3 + \frac{\mathbb{E}X_{2j}^4}{n_{2j}} \bigg)^{3/4} \beta_3(X_{2j-1}, X_{2j}) \bigg\}, \\ \gamma_4 &= \sum_{k=1}^{r} \bigg\{ \prod_{j \neq k}^{r} \bigg(3 + \frac{\mathbb{E}X_{2j-1}^4}{n_{2j-1}} \bigg) \bigg(3 + \frac{\mathbb{E}X_{2j}^4}{n_{2j}} \bigg) \beta_4(X_{2j-1}, X_{2j}) \bigg\}, \end{aligned}$$

and the β_i *are defined as in Theorem 6.20.*

Proof. The result in the case $r = 1$ was proved in Theorem 6.20. We extend this result to general $r \geq 1$ by using a Lindeberg scheme. With V_k and $g_k(x)$ defined as they were in the proof of Theorem 6.21, we can examine that proof to see that

$$\begin{aligned} &|\mathbb{E}h(W) - \mathrm{PN}_{2r}^1 h| \\ &= |\mathbb{E}h(\textstyle\prod_{k=1}^{2r} W_k) - \mathbb{E}h(\prod_{k=1}^{2r} Z_k)| \\ &\leq \sum_{k=1}^{r} \mathbb{E}|\mathbb{E}((g_{2k}(W_{2k-1}W_{2k}) - g_{2k}(Z_{2k-1}Z_{2k})) \mid Z_1, \ldots, Z_{k-1}, W_{k+1}, \ldots, W_r))|. \end{aligned}$$

We may use the bound from Theorem 6.20 to bound the conditional expectation, and so

$$\begin{aligned} \mathbb{E}h(W) - \mathrm{PN}_{2r}^1 h| \leq \sum_{k=1}^{r} \mathbb{E}\bigg[\frac{1}{12n} \bigg(&\beta_2(X_{2k-1}, X_{2k}) \prod_{i=1}^{k-1} \prod_{j=k+1}^{r} |Z_{2i-1}Z_{2i}W_{2j-1}W_{2j}|^2 \|h''\| \\ &+ \beta_3(X_{2k-1}, X_{2k}) \prod_{i=1}^{k-1} \prod_{j=k+1}^{r} |Z_{2i-1}Z_{2i}W_{2j-1}W_{2j}|^3 \|h^{(3)}\| \\ &+ \beta_4(X_{2k-1}, X_{2k}) \prod_{i=1}^{k-1} \prod_{j=k+1}^{r} |Z_{2i-1}Z_{2i}W_{2j-1}W_{2j}|^4 \|h^{(4)}\| \bigg) \bigg]. \end{aligned} \tag{6.45}$$

To bound the expectations on the right-hand side of (6.45) we proceed by using independence and that $\mathbb{E}Z^2 = 1$, $\mathbb{E}|Z|^3 = 2\sqrt{\frac{2}{\pi}} < 3^{3/4}$, $\mathbb{E}Z^4 = 3$, and $\mathbb{E}W_k^2 = 1$, $\mathbb{E}|W_k|^3 < (3 + \frac{\mathbb{E}X_k^4}{n_k})^{3/4}$,

$\mathbb{E}W_k^4 < 3 + \frac{\mathbb{E}X_k^4}{n_k}$. This gives

$$\begin{aligned}
\prod_{i=1}^{k-1}\prod_{j=k+1}^{r}\mathbb{E}|Z_{2i-1}Z_{2i}W_{2j-1}W_{2j}|^2 &= 1, \\
\prod_{i=1}^{k-1}\prod_{j=k+1}^{r}\mathbb{E}|Z_{2i-1}Z_{2i}W_{2j-1}W_{2j}|^3 &< (3^{3/4})^{2k-2}\prod_{j=k+1}^{r}\left(3+\frac{\mathbb{E}X_{2j-1}^4}{n_{2j-1}}\right)^{3/4}\left(3+\frac{\mathbb{E}X_{2j}^4}{n_{2j}}\right)^{3/4} \\
&\leq \prod_{j\neq k}^{r}\left(3+\frac{\mathbb{E}X_{2j-1}^4}{n_{2j-1}}\right)^{3/4}\left(3+\frac{\mathbb{E}X_{2j}^4}{n_{2j}}\right)^{3/4}, \\
\prod_{i=1}^{k-1}\prod_{j=k+1}^{r}\mathbb{E}|Z_{2i-1}Z_{2i}W_{2j-1}W_{2j}|^4 &< 3^{2k}\prod_{j=k+1}^{r}\left(3+\frac{\mathbb{E}X_{2j-1}^4}{n_{2j-1}}\right)\left(3+\frac{\mathbb{E}X_{2j}^4}{n_{2j}}\right) \\
&\leq \prod_{j\neq k}^{r}\left(3+\frac{\mathbb{E}X_{2j-1}^4}{n_{2j-1}}\right)\left(3+\frac{\mathbb{E}X_{2j}^4}{n_{2j}}\right).
\end{aligned}$$

Substituting these inequalities into (6.45) gives the desired bound. □

6.3.4 Friedman's Statistic

Our final application is to Friedman's χ^2 test, which was introduced by Friedman [28]. Friedman's test is non-parametric, and is used to detect differences in treatments across multiple test attempts. With Friedman's test we can test the null hypothesis that there is no treatment effect against the general alternative without needing to assume that the data are normally distributed.

Theorem 6.23 (**Friedman's χ^2 test**). *Consider J independent trials, with the j-th trial leading to a ranking $\pi_j(1),\dots,\pi_j(I)$, where $\pi_j(i)\in\{1,\dots,I\}$, over the I treatments. Under the null hypothesis, the rankings are independent permutations $\pi_1,\dots,\pi_J$, with each permutation being equally likely. Let*

$$Y_i = \frac{1}{J}\sum_{j=1}^{J}\pi_j(i) - \frac{I+1}{2}. \tag{6.46}$$

Then the Friedman χ^2 statistic, given by

$$W = \frac{12J}{I(I+1)}\sum_{i=1}^{I}Y_i^2, \tag{6.47}$$

is asymptotically $\chi^2_{(I-1)}$ distributed under the null hypothesis.

In this subsection, we demonstrate how the multivariate normal Stein equation can be used to obtain bounds on the distance between W and the χ^2_{I-1} distribution. We establish a bound of order $J^{-1/2}$, however, since the Y_i are asymptotically normally distributed and we can write W in the form $g(Y_1,\dots,Y_I)$ with $g(y_1,\dots,y_I)=g(-y_1,\dots,-y_I)$, we would expect that the true rate of convergence is $O(J^{-1})$. Such a bound would improve on the $O(J^{-\frac{I}{I+1}})$ convergence rate that Jensen [37] established. Pickett [55] used symmetry arguments to obtain a $O(n^{-1})$ bound for Pearson's χ^2 statistic, and we would that expect similar arguments, and some rather lengthy calculations, would verify this. We now state our result.

Theorem 6.24. *Suppose $I\geq 2$ and let W denote Friedman's statistic, defined as per equation (6.47). Then, for $h\in\mathcal{C}_\lambda\cap C_b^2(\mathbb{R})$, we have*

$$|\mathbb{E}h(W)-\chi^2_{(I-1)}h|\leq\left(\sqrt{\pi}+\frac{\mathrm{e}^{-1}}{\sqrt{2}}+\frac{1}{\sqrt{I}}\right)\sqrt{\frac{I}{J}}\left[\frac{1}{2\sqrt{5}}\|h\|+\left(\frac{1+\sqrt{3}}{12\sqrt{5}}+\frac{105}{2}\right)\|h'\|+\frac{2283}{4}\|h''\|\right],$$

where $\chi^2_{(I-1)}h$ denotes the expectation of $h(Z)$ for $Z\sim\chi^2_{(I-1)}$.

Before proving Theorem 6.24, we introduce some notation and state some preliminary lemmas. We let $\rho_j(i)=\pi_j(i)-\frac{I+1}{2}$ and write $S_i=\frac{\sqrt{12J}}{\sqrt{I(I+1)}}Y_i$. Recalling (6.46), we have

$$S_i=\frac{\sqrt{12}}{\sqrt{I(I+1)J}}\sum_{j=1}^J\rho_j(i)$$

and we can therefore write

$$W=\sum_{i=1}^I S_i^2.$$

Since $\rho_1(i),\dots,\rho_J(i)$ are i.i.d. with zero mean and variance $\frac{I^2-1}{12}$, it follows, by the central limit theorem, that $S_i\xrightarrow{\mathcal{D}}N(0,1)$ as $J\to\infty$. Whilst the sequence of random variables $\rho_1(i),\dots,\rho_J(i)$ are independent, the S_i are not independent, and we shall need the following simple lemma for the covariance matrix of $\mathbf{S}=(S_1,\dots,S_I)$.

Lemma 6.25. *The covariance matrix of $\mathbf{S}$, denoted by $\Sigma_{\mathbf{S}}=(\sigma_{ij})$, has entries*

$$\sigma_{ii}=\frac{I-1}{I}\quad\text{and}\quad\sigma_{ij}=-\frac{1}{I}\quad(i\neq j).$$

Proof. Since $\rho_1(i),\dots,\rho_J(i)$ are independent, we have

$$\sigma_{ii}=\mathrm{Var}S_i=\frac{12}{I(I+1)J}\sum_{j=1}^J\mathrm{Var}\rho_j(i)=\frac{12}{I(I+1)J}\times J\times\frac{I^2-1}{12}=\frac{I-1}{I}.$$

Suppose now that $i \neq j$. Since $\sum_{i=1}^{I} S_i = 0$, we have

$$0 = \mathbb{E}\bigg(S_j \sum_{k=1}^{I} S_k\bigg) = \mathbb{E}S_j^2 + \sum_{k \neq j} \mathbb{E}S_k S_j = \mathbb{E}S_j^2 + (I-1)\mathbb{E}S_1 S_j,$$

where we used that the S_j are identically distributed to obtain the equality. On rearranging, and using that $\mathbb{E}S_i^2 = \frac{I-1}{I}$, we have that $\mathbb{E}S_i S_j = -\frac{1}{I}$ for $i \neq j$. Since the $\mathbb{E}S_i = 0$, it follows that $\sigma_{ij} = \mathrm{Cov}(S_i, S_j) = \mathbb{E}S_i S_j = -\frac{1}{I}$, as required. $\square$

As we have done throughout this chapter, we shall use the multivariate normal Stein equation (2.12) (in this case with $\Sigma = \Sigma_{\mathbf{S}}$) to obtain bounds on the convergence of the statistic W. However, $\Sigma_{\mathbf{S}}$ is not positive-definite (it is clearly singular), and therefore the solution (2.16) of multivariate normal Stein equation is not well defined. We must therefore proceed in a different manner than usual, and do so by exploiting a connection between the Stein equations for $\chi^2_{(I-1)}$ and $\mathrm{MVN}(\mathbf{0}, \Sigma_{\mathbf{S}})$ distributions, which is given in Lemma 5.7 of Pickett [55].

Lemma 6.26 (**Multivariate normal characterisation**). *Let $\mathcal{A}_Z f$ denote the generator of the process with stationary distribution equal in law to Z, applied to a function f. Then, for any $f \in C^2(\mathbb{R})$,*

$$\mathcal{A}_X g(\mathbf{S}) = \mathcal{A}_Y f(W), \tag{6.48}$$

where $X \sim \mathrm{MVN}(\mathbf{0}, \Sigma_{\mathbf{S}})$, $Y \sim \chi^2_{(I-1)}$ and $g : \mathbb{R}^I \to \mathbb{R}$ is defined by $g(\mathbf{S}) = \frac{1}{4} f(W)$.

Remark 6.27. Pickett [55] established the connection between the $\chi^2_{(m-1)}$ and $\mathrm{MVN}(\mathbf{0}, \Sigma)$ Stein equations, where the covariance matrix Σ has entries

$$\sigma_{ii} = \frac{n(m-1)}{m^2} \quad \text{and} \quad \sigma_{ij} = -\frac{n}{m^2} \quad (i \neq j).$$

The covariance matrices Σ and $\Sigma_{\mathbf{S}}$ are equal upto a multiplicative constants and, on examining Pickett's proof, it is easy to see that the connection formula (6.48) holds for the covariance matrix $\Sigma_{\mathbf{S}}$.

As we shall see in the proof of Theorem 6.24, Lemma 6.26 allows us to bound the partial derivatives of f in terms of the derivatives of g, which are bounded using the smoothness estimates (2.35) for the solution of the $\chi^2_{(p)}$ Stein equation.

Due to the weak global dependence of the $\rho_j(i)$, we consider an exchangeable pair coupling to be appropriate. Before presenting our coupling, we state our final preliminary result, which appears as an intermediate inequality in the proof of Theorem 2.1 of Reinert and Röllin [61]. It is worth noting that up to this part of their proof Reinert and Röllin had not used that Σ is positive-definite.

Lemma 6.28. *Let* $\mathbf{W} = (W_1, \ldots, W_d) \in \mathbb{R}^d$*. Assume that* $(\mathbf{W}, \mathbf{W}')$ *is an exchangeable pair of* $\mathbb{R}^d$*–valued random vectors such that* $\mathbb{E}\mathbf{W} = \mathbf{0}, \mathbb{E}\mathbf{W}\mathbf{W}^T = \Sigma$*. Suppose further that* $\mathbb{E}^{\mathbf{W}}(\mathbf{W}' - \mathbf{W}) = -\Lambda \mathbf{W}$ *for an invertible* $d \times d$ *matrix* Λ*. Then provided that* $f \in C^3(\mathbb{R}^d)$*, we have*

$$\begin{aligned}
&|\mathbb{E}\{\nabla^T \Sigma \nabla f(\mathbf{W}) - \mathbf{W}^T \nabla f(\mathbf{W})\}| \\
&\leq \frac{1}{2} \sum_{m,i,j=1}^{d} |(\Lambda^{-1})_{m,i}| \sqrt{\mathrm{Var}\mathbb{E}^{\mathbf{W}}(W_i' - W_i)(W_j' - W_j)} \left| \mathbb{E}\left[\frac{\partial^2 f(\mathbf{W})}{\partial w_m \partial w_j}\right]\right| \\
&\quad + \frac{1}{4} \sum_{m,i,j,k=1}^{d} |(\Lambda^{-1})_{m,i}| \mathbb{E}\left|(W_i' - W_i)(W_j' - W_j)(W_k' - W_k) \frac{\partial^3 f(\mathbf{W}^*)}{\partial w_m \partial w_j \partial w_k}\right|,
\end{aligned}$$

where $\mathbf{W}^* = \theta \mathbf{W} + (1 - \theta)\mathbf{W}'$ *for some* $\theta \in (0, 1)$*.*

Let us now prove Theorem 6.24.

Proof of Theorem 6.24

We begin by presenting our exchangeable pair coupling. This coupling was introduced in a talk given by Gesine Reinert. Pick an index $M \in \{1, \ldots, J\}$ uniformly and indices $K, L \in \{1, \ldots, I\}$ uniformly and independently of M. If $M = m$, $K = k$, $L = l$, define the permutation π_m' by

$$\begin{aligned}
\pi_m'(k) &= \pi_m(l), \\
\pi_m'(l) &= \pi_m(k), \\
\pi_m'(i) &= \pi_m(i), \qquad i \neq k, l.
\end{aligned}$$

We then put

$$\begin{aligned}
Y_K' &= Y_K - \frac{1}{J}\rho_M(K) + \frac{1}{J}\rho_M(L), \\
Y_L' &= Y_L - \frac{1}{J}\rho_M(L) + \frac{1}{J}\rho_M(K), \\
Y_i' &= Y_i, \qquad i \neq K, L,
\end{aligned}$$

which can be written in terms of the S_i as

$$\begin{aligned}
S_K' &= S_K - \frac{\sqrt{12}}{\sqrt{I(I+1)J}}[\rho_M(K) - \rho_M(L)], \\
S_L' &= S_L - \frac{\sqrt{12}}{\sqrt{I(I+1)J}}[\rho_M(L) - \rho_M(K)],
\end{aligned}$$

and

$$S_i' = S_i, \qquad i \neq K, L.$$

It is clear that $(\mathbf{S}, \mathbf{S}')$ is an exchangeable pair, and we now verify that condition (5.33) holds.

$$\begin{aligned}
\mathbb{E}^{\mathbf{S}}(S_i' - S_i) &= \frac{1}{I^2 J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}(S_i' - S_i \mid K = k, L = l, M = m) \\
&= \frac{1}{I^2 J}\frac{\sqrt{12}}{\sqrt{I(I+1)J}}\left\{\sum_{k=1}^{I}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}(\rho_m(k) - \rho_m(i)) + \sum_{l=1}^{I}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}(\rho_m(l) - \rho_m(i))\right\} \\
&= -\frac{2}{IJ}\frac{\sqrt{12}}{\sqrt{I(I+1)J}}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}\rho_m(i) \\
&= -\frac{2}{IJ}S_i,
\end{aligned}$$

where we used that $\sum_{k=1}^{I}\rho_m(k) = 0$ to obtain the third equality. Therefore condition (5.33) holds with $\Lambda = \frac{2}{IJ}\mathrm{I}_I$, where I_I is the $I \times I$ identity matrix.

Having established an appropriate exchangeable pair coupling, we are finally in a position to bound the quantity $\mathbb{E}h(W) - \chi^2_{(I-1)}h$. From the χ^2 Stein equation (2.15) and the connection Lemma 6.26, we have

$$|\mathbb{E}h(W) - \chi^2_{(I-1)}h| = \left|\mathbb{E}\left\{f''(W) - \frac{1}{2}(I - 1 - W)f'(W)\right\}\right| = |\mathbb{E}\{\nabla^T\Sigma\nabla g(\mathbf{S}) - \mathbf{S}^T\nabla g(\mathbf{S})\}|,$$

where $g : \mathbb{R}^I \to \mathbb{R}$ is defined by $g(\mathbf{S}) = \frac{1}{4}f(W)$. Using Lemma 6.28, we have

$$\begin{aligned}
&|\mathbb{E}h(W) - \chi^2_{(I-1)}h| \\
&\leq \frac{1}{2}\sum_{m,i,j=1}^{I}|(\Lambda^{-1})_{m,i}|\sqrt{\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i' - S_i)(S_j' - S_j)}\mathbb{E}\left|\left[\frac{\partial^2 g(\mathbf{S})}{\partial s_m \partial s_j}\right]\right| \\
&\quad + \frac{1}{4}\sum_{m,i,j,k=1}^{I}|(\Lambda^{-1})_{m,i}|\mathbb{E}\left|(S_i' - S_i)(S_j' - S_j)(S_k' - S_k)\frac{\partial^3 g(\mathbf{S}^*)}{\partial s_m \partial s_j \partial s_k}\right| \\
&= \frac{IJ}{4}\sum_{i=1}^{I}\sqrt{\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i' - S_i)^2}\left|\mathbb{E}\left[\frac{\partial^2 g(\mathbf{S})}{\partial s_i^2}\right]\right| + \frac{IJ}{4}\sum_{i=1}^{I}\sum_{j\neq i}\sqrt{\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i' - S_i)(S_j' - S_j)}\left|\mathbb{E}\left[\frac{\partial^2 g(\mathbf{S})}{\partial s_i \partial s_j}\right]\right| \\
&\quad + \frac{IJ}{8}\sum_{i=1}^{I}\mathbb{E}\left|(S_i' - S_i)^3\frac{\partial^3 g(\mathbf{S}^*)}{\partial s_i^3}\right| + \frac{IJ}{4}\sum_{i=1}^{I}\sum_{j\neq i}\mathbb{E}\left|(S_i' - S_i)^2(S_j' - S_j)\frac{\partial^3 g(\mathbf{S}^*)}{\partial s_i^2 \partial s_j}\right|, \qquad (6.49)
\end{aligned}$$

where $\mathbf{S}^* = \theta\mathbf{S} + (1-\theta)\mathbf{S}$ for some $\theta \in (0, 1)$, and to obtain the second equality we used that, by construction, $(S_i' - S_i)(S_j' - S_j)(S_k' - S_k) = 0$, unless $i = j$, $j = k$ or $i = k$. Let us denote the first term of (6.49) by (A); the second term by (B); the third by (C) and the fourth by (D).

We now set about bounding these terms and we first deal with (A).

$$\begin{aligned}
\mathbb{E}^{\mathbf{S}}(S_i'-S_i)^2 &= \frac{1}{I^2J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}((S_i'-S_i)^2 \mid K=k, L=l, M=m)\\
&= \frac{1}{I^2J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J}\{\mathbf{1}(i=k)+\mathbf{1}(i=l)\}\frac{12}{I(I+1)J}\mathbb{E}^{\mathbf{Y}}(\rho_m(k)-\rho_m(l))^2\\
&= \frac{2\times 12}{I^3(I+1)J^2}\sum_{m=1}^{J}\sum_{k=1}^{I}\mathbb{E}^{\mathbf{Y}}(\rho_m(k)-\rho_m(i))^2\\
&= \frac{2\times 12}{I^3(I+1)J^2}\sum_{m=1}^{J}\sum_{k=1}^{I}\mathbb{E}^{\mathbf{Y}}(\rho_m^2(k)-2\rho_m(k)\rho_m(i)+\rho_m^2(i))\\
&= \frac{2(I-1)}{I^2J}+\frac{24}{I^2(I+1)J^2}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{Y}}\rho_m^2(i),
\end{aligned}$$

where to obtain the final equality we used that $\sum_{k=1}^{I}\rho_m(k)=0$ and that $\sum_{k=1}^{I}\mathbb{E}\rho_m^2(i) = I\mathrm{Var}\rho_m(1)=\frac{I(I^2-1)}{12}$. Hence,

$$\begin{aligned}
\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i'-S_i)^2 &= \frac{576}{I^4(I+1)^2J^4}\sum_{m=1}^{J}\mathrm{Var}\mathbb{E}^{\mathbf{Y}}\rho_m^2(i)\\
&\leq \frac{576}{I^4(I+1)^2J^4}\sum_{m=1}^{J}\mathrm{Var}\rho_m^2(i)\\
&= \frac{576}{I^4(I+1)^2J^3}(\mathbb{E}\rho_1^4(i)-(\mathbb{E}\rho_1^2(i))^2)\\
&= \frac{576}{I^4(I+1)^2J^5}\left(\frac{1}{240}(I^2-1)(3I^2-7)-\frac{(I^2-1)^2}{144}\right)\\
&= \frac{16(I-1)(I^2-4)}{5I^4(I+1)J^3}\\
&< \frac{16(I-1)^2}{5I^4J^3}. \qquad (6.50)
\end{aligned}$$

where the first inequality is due to the fact that, for any random variables X and Y, $\mathrm{Var}Y = \mathbb{E}[\mathrm{Var}(Y|X)]+\mathrm{Var}[\mathbb{E}(Y|X)]$. To obtain the third equality we used that $\mathbb{E}\rho_1^4(i)=\frac{1}{240}(I^2-1)(3I^2-7)$ (see Zwillinger [76]). Recalling that $g(\mathbf{S})=\frac{1}{4}f(W)$, we have

$$\begin{aligned}
\left|\mathbb{E}\left[\frac{\partial^2 g(\mathbf{S})}{\partial s_i^2}\right]\right| &= \left|\mathbb{E}\left\{S_i^2f''(W)+\frac{1}{2}f'(W)\right\}\right|\\
&\leq \|f''\|\mathbb{E}S_i^2+\frac{1}{2}\|f'\|\\
&= \frac{I-1}{12I}\|f''\|+\frac{1}{2}\|f'\|
\end{aligned}$$

$$\leq \frac{1}{12}\|f''\| + \frac{1}{2}\|f'\|. \tag{6.51}$$

Using bounds (6.50) and (6.51) gives

$$(A) \leq \frac{IJ}{4} \times I \times \sqrt{\frac{16}{5}}\frac{I-1}{I^2J^{3/2}} \times \left(\frac{\|f''\|}{12} + \frac{\|f'\|}{2}\right) = \frac{I-1}{12\sqrt{5}\sqrt{J}}(\|f''\| + 6\|f'\|).$$

We bound (B) using a similar approach to that used to bound (A):

$$\begin{aligned}
\mathbb{E}^{\mathbf{S}}(S_i'-S_i)(S_j'-S_j) &= \frac{1}{I^2J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}((S_i'-S_i)(S_j'-S_j) \mid K=k, L=l, M=m) \\
&= \frac{1}{I^2J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J}\{\mathbf{1}(i=k, j=l) + \mathbf{1}(i=l, j=k)\} \\
&\quad \times \mathbb{E}^{\mathbf{S}}((S_i'-S_i)(S_j'-S_j) \mid K=k, L=l, M=m) \\
&= \frac{2}{I^2J}\cdot\frac{12}{I(I+1)J}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}(\rho_m(i)-\rho_m(j))(\rho_m(j)-\rho_m(i)) \\
&= -\frac{24}{I^3(I+1)J^2}\sum_{m=1}^{J}\mathbb{E}^{\mathbf{S}}(\rho_m^2(i) - 2\rho_m(i)\rho(j) + \rho_m^2(j)).
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i'-S_i)(S_j-S_j') &= \frac{576}{I^6(I+1)^2J^4}\sum_{m=1}^{J}\mathrm{Var}\mathbb{E}^{\mathbf{S}}(\rho_m^2(i) - 2\rho_m(i)\rho(j) + \rho_m^2(j)) \\
&\leq \frac{576}{I^6(I+1)^2J^4}\sum_{m=1}^{J}\mathrm{Var}(\rho_m^2(i) - 2\rho_m(i)\rho(j) + \rho_m^2(j)).
\end{aligned}$$

We now note that for random variables X, Y, Z with finite second moment,

$$\begin{aligned}
\mathrm{Var}(X+Y+Z) &= \mathbb{E}(X+Y+Z-\mathbb{E}X-\mathbb{E}Y-\mathbb{E}Z)^2 \\
&\leq 3(\mathbb{E}(X-\mathbb{E}X)^2 + \mathbb{E}(Y-\mathbb{E}Y)^2 + \mathbb{E}(Z-\mathbb{E}Z)^2) \\
&= 3(\mathrm{Var}X + \mathrm{Var}Y + \mathrm{Var}Z).
\end{aligned}$$

Therefore

$$\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i'-S_i)(S_j-S_j') \leq \frac{1728}{I^6(I+1)^2J^3}\Big(2\mathrm{Var}\rho_1^2(i) + 4\mathrm{Var}\rho_1(i)\rho_2(j)\Big).$$

Now,

$$\mathrm{Var}\rho_1^2(i) = \mathbb{E}\rho_1^4(i) - (\mathbb{E}\rho_1^2(i))^2 = \frac{1}{240}(I^2-1)(I^2-4) - \frac{1}{144}I^2 - 1 = \frac{1}{180}(I^2-1)(I^2-4),$$

and we also have

$$\begin{aligned}\mathrm{Var}\rho_1(i)\rho_2(j) &= \mathbb{E}\rho_1^2(i)\rho_2^2(j) - (\mathbb{E}\rho_1(i)\rho_2(j))^2 \\ &\leq \mathbb{E}\rho_1^2(i)\rho_2^2(j) \\ &\leq \sqrt{\mathbb{E}\rho_1^4(i)\mathbb{E}\rho_2^4(j)} \\ &= \frac{1}{240}(I^2-1)(3I^2-7).\end{aligned}$$

Therefore

$$\begin{aligned}\mathrm{Var}\mathbb{E}^{\mathbf{S}}(S_i'-S_i)(S_j-S_j') &\leq \frac{1728}{I^6(I+1)^2J^3}\left(\frac{1}{90}(I^2-1)(I^2-4) + \frac{1}{60}(I^2-1)(3I^2-7)\right) \\ &= \frac{48(I-1)(11I^2-29)}{5I^6(I+1)J^3} \\ &< \frac{48(I-1)^2}{5I^6J^3}. \end{aligned} \tag{6.52}$$

Using the Cauchy-Schwarz inequality gives

$$\left|\mathbb{E}\left[\frac{\partial^2 g(\mathbf{S})}{\partial s_i \partial s_j}\right]\right| = |\mathbb{E}S_iS_jf''(W)| \leq \|f''\|\mathbb{E}|S_iS_j| \leq \|f''\|\sqrt{\mathbb{E}S_i^2\mathbb{E}S_j^2} = \frac{I-1}{12I}\|f''\| \leq \frac{\|f''\|}{12}. \tag{6.53}$$

From inequalities (6.52) and (6.53), we have

$$(B) \leq \frac{IJ}{4} \times I^2 \times \sqrt{\frac{48}{5}\frac{I-1}{I^3J^{3/2}}} \times \frac{1}{12}\|f''\| = \frac{\sqrt{3}(I-1)\|f''\|}{12\sqrt{5}\sqrt{J}}.$$

We now bound (C). We have

$$\frac{\partial^3 g(\mathbf{S})}{\partial s_i^3} = 2S_i^3 f^{(3)}(W) + 3S_i f''(W),$$

and using this formula gives

$$\begin{aligned}\mathbb{E}\left|(S_i'-S_i)^3\frac{\partial^3 g(\mathbf{S}^*)}{\partial s_i^3}\right| &\leq 2\|f^{(3)}\|\mathbb{E}|(S_i'-S_i)^3(\theta S_i + (1-\theta)S_i')^3| \\ &\quad + 3\|f''\|\mathbb{E}|(S_i'-S_i)^3(\theta S_i + (1-\theta)S_i')|\end{aligned}$$

$$\begin{aligned}
&\leq 8\|f^{(3)}\|\mathbb{E}|(S_i'-S_i)^3[\theta^3S_i^3+(1-\theta)^3(S_i')^3]| \\
&\quad+3\|f''\|\mathbb{E}|(S_i'-S_i)^3[\theta S_i+(1-\theta)S_i']| \\
&\leq 8\|f^{(3)}\|\{\theta^3\mathbb{E}|(S_i'-S_i)^3S_i^3|+(1-\theta)^3\mathbb{E}|(S_i'-S_i)^3(S_i')^3|\} \\
&\quad+3\|f''\|\{\theta\mathbb{E}|(S_i'-S_i)^3S_i|+(1-\theta)\mathbb{E}|(S_i'-S_i)^3S_i'|\} \\
&\leq 8\|f^{(3)}\|\mathbb{E}|(S_i'-S_i)^3S_i^3|+3\|f''\|\mathbb{E}|(S_i'-S_i)^3S_i|,
\end{aligned}$$

where to obtain the final inequality we used that S_i and S_i' have the same distribution, and that $\theta^3+(1-\theta)^3<1$, for $\theta\in(0,1)$. We have that

$$\begin{aligned}
\mathbb{E}|(S_i'-S_i)^3S_i^3| &= \frac{1}{I^2J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J}\mathbb{E}(|S_i'-S_i|^3|S_i|^3\mid K=k,L=l,M=m) \\
&= \frac{2}{I^2J^{5/2}}\cdot\frac{12^{3/2}}{I^{3/2}(I+1)^{3/2}}\sum_{k=1}^{I}\sum_{m=1}^{J}\mathbb{E}|(\rho_m(k)-\rho_m(i))^3S_i^3| \\
&\leq \frac{2\cdot 12^{3/2}}{I^{7/2}(I+1)^{3/2}J^{5/2}}\times IJ\times(I-1)^3\mathbb{E}|S_i|^3,
\end{aligned}$$

where we used that $|\rho_m(k)-\rho_m(i)|\leq I-1$ to obtain the inequality. By Hölder's inequality we have $\mathbb{E}|S_i|^3\leq\{\mathbb{E}S_i^4\}^{3/4}$, and so

$$\begin{aligned}
\mathbb{E}|(S_i'-S_i)^3S_i^3| &\leq \frac{2\cdot 12^{3/2}(I-1)^3}{I^{5/2}(I+1)^{3/2}J^{3/2}}\{\mathbb{E}S_i^4\}^{3/4} \\
&= \frac{2\cdot 12^{3/2}(I-1)^3}{I^{5/2}(I+1)^{3/2}J^{3/2}}\cdot\frac{12^{3/2}}{I^{3/2}(I+1)^{3/2}}\left(\frac{3(J-1)}{J}(\mathbb{E}\rho_1^2(i))^2+\frac{1}{J}\mathbb{E}\rho_1^4(i)\right)^{3/4} \\
&= \frac{3456(I-1)^3}{I^4(I+1)^3J^{3/2}}\left(\frac{J-1}{J}\cdot\frac{(I^2-1)^2}{48}+\frac{1}{J}\cdot\frac{1}{240}(I^2-1)(3I^2-7)\right)^{3/4} \\
&< \frac{3456(I-1)^3}{I^4(I+1)^3J^{3/2}}\left(\frac{1}{48}+\frac{1}{80}\right)^{3/4}I^3 \\
&= \frac{3456(I-1)^3}{30^{3/4}I(I+1)^3J^{3/2}} \\
&< \frac{270(I-1)}{I^2J^{3/2}}.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathbb{E}|(S_i'-S_i)^3S_i| &\leq \frac{2\cdot 12^{3/2}(I-1)^3}{I^{5/2}(I+1)^{3/2}J^{3/2}}\mathbb{E}|S_i| \\
&\leq \frac{2\cdot 12^{3/2}(I-1)^3}{I^{5/2}(I+1)^{3/2}J^{3/2}}\sqrt{\mathbb{E}S_i^2}
\end{aligned}$$

$$\begin{aligned}&= \frac{2\cdot 12^{3/2}(I-1)^{7/2}}{I^3(I+1)^{3/2}J^{3/2}} \\ &< \frac{84(I-1)}{I^2J^{3/2}}.\end{aligned}$$

Therefore

$$(C) \leq \frac{IJ}{8}\times I\times\bigg(8\|f^{(3)}\|\cdot\frac{270}{IJ^{3/2}}+3\|f''\|\cdot\frac{84}{IJ^{3/2}}\bigg)=\frac{I-1}{\sqrt{J}}\bigg[270\|f^{(3)}\|+\frac{63}{2}\|f''\|\bigg].$$

Finally, we bound (D). We have

$$\frac{\partial^3 g(\mathbf{S})}{\partial s_i^2\partial s_j}=2S_i^2S_jf^{(3)}(W)+S_jf''(W), \qquad i\neq j,$$

and using this formula gives

$$\begin{aligned}&\mathbb{E}\bigg|(S_i'-S_i)^2(S_j'-S_j)\frac{\partial^3 g(\mathbf{S}^*)}{\partial s_i^2\partial s_j}\bigg| \\ &\leq 2\|f^{(3)}\|\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)(S_i')^2S_j|+\|f''\|\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_j'| \\ &\leq 4\|f^{(3)}\|\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)(\theta^2S_i^2+(1-\theta)(S_i')^2)(\theta S_j+(1-\theta)S_j')| \\ &\quad+\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)(\theta S_j+(1-\theta)S_j')| \\ &= 4\|f^{(3)}\|\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)(\theta^3S_i^2S_j+\theta^2(1-\theta)S_i^2S_j'+\theta(1-\theta)(S_i')^2S_j \\ &\quad+(1-\theta)^3(S_i')^2S_j')|+\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)(\theta S_j+(1-\theta)S_j')| \\ &\leq 4\|f^{(3)}\|\{(\theta^3+(1-\theta)^3)\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_i^2S_j| \\ &\quad+(\theta^2(1-\theta)+\theta(1-\theta)^2)\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_i^2S_j'|\} \\ &\quad+\|f''\|\{\theta\mathbb{E}|(S_i'-S_j)^2(S_j'-S_j)S_j|+(1-\theta)\mathbb{E}|(S_i'-S_j)^2(S_j'-S_j)S_j'|\} \\ &\leq \|f^{(3)}\|\{4\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_i^2S_j|+\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_i^2S_j'|\} \\ &\quad+\|f''\|\mathbb{E}|(S_i'-S_j)^2(S_j'-S_j)S_j|,\end{aligned}$$

where to obtain the final inequality we used that S_i and S_i' have the same distribution, and that $\theta^2(1-\theta)+\theta(1-\theta)^2\leq\frac{1}{4}$, for $\theta\in(0,1)$. We now apply a similar calculation to one that was used to bound (C) to obtain

$$\begin{aligned}&\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_i^2S_j| \\ &=\frac{1}{I^2J}\sum_{k=1}^I\sum_{l=1}^I\sum_{m=1}^J\mathbb{E}(|(S_i'-S_i)^2(S_j'-S_j)S_i^2S_j|\mid K=k,L=l,M=m)\end{aligned}$$

$$
\begin{aligned}
&= \frac{2}{I^2 J^{5/2}} \cdot \frac{12^{3/2}}{I^{3/2}(I+1)^{3/2}} \sum_{m=1}^{J} \mathbb{E}|(\rho_m(i)-\rho_m(j))^2(\rho_m(j)-\rho_m(i)S_i^2 S_j| \\
&\le \frac{2\cdot 12^{3/2}}{I^{7/2}(I+1)^{3/2}J^{3/2}} \cdot (I-1)^3 \cdot \mathbb{E}|S_i^2 S_j| \\
&\le \frac{2\cdot 12^{3/2}(I-1)^3}{I^{7/2}(I+1)^{3/2}J^{3/2}} \sqrt{\mathbb{E}S_i^4 \mathbb{E}S_j^2} \\
&= \frac{2\cdot 12^{3/2}(I-1)^3}{I^{7/2}(I+1)^{3/2}J^{3/2}} \cdot \frac{12^{3/2}}{I^{3/2}(I+1)^{3/2}} \left(\frac{J-1}{J}\cdot\frac{(I^2-1)^2}{48} + \frac{1}{J}\cdot\frac{1}{240}(I^2-1)(3I^2-7)\right)^{1/2} \\
&\quad \times \left(\frac{I^2-1}{12}\right)^{1/2} \\
&< \frac{3456(I-1)^3}{I^5(I+1)^3 J^{3/2}} \left(\frac{1}{48}+\frac{1}{80}\right)^{1/2} \left(\frac{1}{12}\right)^{1/2} I^2\sqrt{I^2-1} \\
&= \frac{288\sqrt{2}(I-1)^{7/2}}{\sqrt{5}I^3(I+1)^{5/2}J^{3/2}} \\
&< \frac{183(I-1)}{I^3 J^{3/2}}.
\end{aligned}
\tag{6.54}
$$

Similarly,

$$
\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_j| \le \frac{2\cdot 12^{3/2}(I-1)^3}{I^{7/2}(I+1)^{3/2}J^{3/2}}\mathbb{E}|S_i| \le \frac{2\cdot 12^{3/2}(I-1)^{7/2}}{I^4(I+1)^{3/2}J^{3/2}} < \frac{84}{I^2 J^{3/2}}. \tag{6.55}
$$

We also have that

$$
\begin{aligned}
&\mathbb{E}|(S_i'-S_i)^2(S_j'-S_j)S_i^2 S_j'| \\
&= \frac{1}{I^2 J}\sum_{k=1}^{I}\sum_{l=1}^{I}\sum_{m=1}^{J} \mathbb{E}(|(S_i'-S_i)^2(S_j'-S_j)S_i^2 S_j'| \mid K=k, L=l, M=m) \\
&= \frac{2}{I^2 J^{5/2}} \cdot \frac{12^{3/2}}{I^{3/2}(I+1)^{3/2}} \sum_{m=1}^{J} \mathbb{E}\left|(\rho_m(i)-\rho_m(j))^3 S_i^2 \left(S_j - \frac{\sqrt{12}}{\sqrt{I(I+1)J}}(\rho_m(j)-\rho_m(i))\right)\right| \\
&\le \frac{2\cdot 12^{3/2}}{I^{7/2}(I+1)^{3/2}J^{5/2}} \sum_{m=1}^{J} \mathbb{E}|(\rho_m(i)-\rho_m(j))^3 S_i^2 S_j| + \frac{2\cdot 12^2}{I^4(I+1)^2 J^3} \sum_{m=1}^{J} \mathbb{E}|(\rho_m(i)-\rho_m(j))^4 S_i^2| \\
&< \frac{183(I-1)}{I^3 J^{3/2}} + \frac{288}{I^4(I+1)^2 J^{3/2}}(I-1)^4 \mathbb{E}S_i^2 \\
&= \frac{183(I-1)}{I^3 J^{3/2}} + \frac{288(I-1)^5}{I^4(I+1)IJ^2} \\
&< \frac{471(I-1)}{I^3 J^{3/2}}.
\end{aligned}
\tag{6.56}
$$

Using inequalities (6.54), (6.55) and (6.56) gives

$$(D) \leq \frac{IJ}{4} \times I^2 \times \left[\left(4 \cdot \frac{183}{I^2 J^{3/2}} + \frac{471}{I^2 J^{3/2}}\right)\|f^{(3)}\| + \frac{84}{I^2 J^{3/2}}\|f''\|\right] = \frac{I-1}{\sqrt{J}}\left[\frac{1203}{4}\|f^{(3)}\| + 21\|f''\|\right].$$

Summing up our bounds for (A), (B), (C) and (D), and using (2.35) to bound the derivatives of f in terms of the derivatives of the test function h, completes the proof of Theorem 6.24.

We end this subsection with a remark some remarks.

Remark 6.29. In the proof of Theorem 6.24 we made use of inequalities of the type $\mathbb{E}|(\rho_m(k) - \rho_m(i))^3 S_i^3| \leq (I-1)^3 \mathbb{E}|S_i|^3$. This is a very crude approximation, and improving this bound would lead to much smaller multiplicative constants than those in our bound (we would greatly improve on $\frac{105}{2}$ and $\frac{2283}{4}$).

Remark 6.30. An alternative choice of coupling would have been the local approach coupling. We could have introduced the random variables $S_i^{(j)} = S_i - \frac{\sqrt{12}}{\sqrt{I(I+1)J}}\rho_j(i)$, and exploited the fact that $\rho_j(i)$ and the random vector $\mathbf{S}^{(j)} = (S_1^{(j)}, \ldots S_I^{(j)})$ are independent. This would allow us to obtain an order $J^{-1/2}$ type bound, and the calculation would have been simpler and shorter than our exchangeable pair approach. However, preliminary calculations indicate that this approach would lead to a bound of order $I^{5/2}J^{-1/2}$, rather than our $O(I^{1/2}J^{-1/2})$ bound. It would involve a great deal of effort to improve the local approach bound to $O(I^{1/2}J^{-1/2})$. This suggests that the exchangeable pair coupling is a more natural coupling than the local approach coupling when using Stein's method to approximate Friedman's statistic.

Chapter 7

Summary and Future Work

7.1 Summary

In this thesis, motivated by the D_2^* statistic that we considered in Section 1.1, we have developed Stein's method for the class of Variance-Gamma distributions. Our development of Stein's method for Variance-Gamma distributions then lead us to consider the adaptation of Stein's method to the Product Normal distributions and functions of the multivariate normal distribution.

We began this thesis by reviewing Stein's method for normal approximation and demonstrated how the method can be adapted to many other distributions through the use of generators of Markov processes. In Chapter 2 we used this generator approach to obtain new smoothness estimates for the solution of the Gamma Stein equation. Our bounds for the k-th derivative of the solution of the $\Gamma(r, \lambda)$ Stein equation are of order $(r+k)^{-1/2}$, which improve on the $O(r^{-1/2})$ bound of Pickett [55].

In Chapter 3 we adapted Stein's method to the class of Variance-Gamma distributions. A Stein equation for Variance-Gamma distributions was obtained, which reduces to the classical normal and Gamma Stein equations for certain parameter values. We obtained the unique bounded solution to the Stein equation. We then proceeded to obtain uniform bounds on the solution and its first four derivatives for the case of the Symmetric Variance-Gamma Stein equation.

To obtain these smoothness estimates for the solution of the Symmetric Variance-Gamma Stein equation, we needed to bound a number of expressions involving derivatives and integrals of modified Bessel functions $I_\nu(x)$ and $K_\nu(x)$. In Appendix C we obtained new formulas for derivatives of general order for the expressions $x^{-\nu}I_\nu(x)$ and $x^{-\nu}K_\nu(x)$, as well as inequalities for some derivatives and integrals of expressions involving modified Bessel functions. In Appendix D we used these formulas and inequalities to bound a number of expressions involving Bessel functions, which yielded smoothness estimates for the solution of the Symmetric Variance-Gamma Stein equation.

In Chapter 4 we used Stein's method for Variance-Gamma distributions to prove some Symmetric Variance-Gamma limit theorems. We used a conditioning argument and a bound of Pickett [55] for χ^2 approximation to arrive at a Symmetric Variance-Gamma limit theorem for the statistic $U_r = \frac{1}{m}\sum_{k=1}^{r}(\sum_{i=1}^{m} X_{ik})^2 - \frac{1}{n}\sum_{k=1}^{r}(\sum_{i=1}^{n} Y_{ik})^2$, where $(X_{ik})_{1\leq i\leq m,\, 1\leq k\leq r}$ and $(Y_{jk})_{1\leq j\leq n,\, 1\leq k\leq r}$ are collections of i.i.d. random variables with the X_{ik} and Y_{jk} independent, each with zero mean and unit variance. Through the use of symmetry arguments we were able to obtain a bound on the rate of convergence of the statistic $V_r = \frac{1}{\sqrt{mn}}\sum_{i,j,k=1}^{m,n,r} X_{ik}Y_{jk}$ to a $\mathrm{VG}(r,0,1,0)$ random variable of order $m^{-1}+n^{-1}$, for smooth test functions. We then applied this bound to a simple problem in binary sequence comparison. We also used the Variance-Gamma Stein equation to prove a limit theorem for which the random variables were locally dependent, thereby showing that the Stein method's for Variance-Gamma approximation can apply readily even when the assumption of i.i.d. random variables is dropped.

In Chapter 5 we made some progress towards adapting Stein's method to the distribution of the product of r independent mean zero normal random variables. We obtained an elegant Stein equation for this class of distributions. In Chapter 3 we gave bounds for the solution and its first four derivatives in the case $r = 2$, but we have been unable to bound the derivatives for the case $r \geq 3$; this is left as an open problem. Motivated by the Product Normal Stein equation, we obtained a natural generalisation of the zero bias transformation. We established a number of useful properties of this new transformation and demonstrated how this transformation could be used together with the normal product Stein equation to prove Product Normal limit theorems. We ended this chapter by using our generalisation of the zero bias transformation and exchangeable pair couplings to obtain some limit theorems for some statistics that are asymptotically distributed as the product of two independent standard normal random variables.

In the final part of this thesis, in Chapter 6, we demonstrated how the multivariate normal Stein equation can be used to prove limit theorems for statistics that are asymptotically distributed as a function of the multivariate normal distribution. Specifically, we considered the problem of bounding statistics of the form $g(\mathbf{W})$, where $g : \mathbb{R}^d \to \mathbb{R}$ and the entries of the random vector $\mathbf{W}$ are given by $W_k = \frac{1}{\sqrt{n_k}}\sum_{i=1}^{n_k} X_{ik}$ with the X_{ik} being mean zero random variables with non zero, finite variance, by it limiting distribution $g(\mathbf{Z})$, where $\mathbf{Z}$ has the multivariate normal distribution. Using the standard normal Stein equation, we obtained a bound of order $n^{-1/2}$ in the error in this approximation, in a weak convergence sense, for i.i.d. random variables in the case $d = 1$ and

g is differentiable with derivative that has polynomial growth. We showed that this convergence rate can be improved to $O(n^{-1})$ if g is an even function or $\mathbb{E}X_{11}^3 = 0$. We demonstrated how the multivariate normal Stein equation approach can be be used to approximate $g(\mathbf{W})$ under the more general conditions that g has growth rate that is possibly faster than polynomial, g is a function of more than one variable, and that the X_{ik} are locally dependent. We saw that under these weaker assumptions we would expect a convergence rate of order n^{-1} provided that g is an even function. As a consequence, the faster than expected $O(n^{-1})$ limit theorems of Chapter 4 were no accident, but were to be expected. We applied the theory developed in this chapter to prove χ^2, Variance-Gamma and Product Normal limit theorems. Finally, we used the multivariate normal Stein equation together with an exchange pair coupling to obtain an explicit bound for the error in approximating Friedman's χ^2 statistic by its limiting distribution.

7.2 Open Problems

7.2.1 Smoothness for the solutions of the general Variance-Gamma and Product Normal Stein equations

In Chapter 3 we obtained uniform bounds for the solution and its first four derivatives of the solution of $\mathrm{VG}(r, 0, \sigma, \mu)$ Stein equation (3.14). In Theorem 3.21 we obtained bounds that were of order $r^{-1/2}$ for $r \in \mathbb{Z}^+$. However, our bounds for the second, third and fourth order derivative for general $r > 0$, given in Lemma 3.20, perform very poorly for values of r that are very close to but not equal to positive integers (recall that $r = 2\nu + 1$). These poor bounds are an artefact of the series expansion method used in the proofs of Lemma D.16, D.17 and D.18. An open problem is to obtain $O(r^{-1/2})$ bounds for all $r > 0$ (recall $r = 2\nu + 1$) for the expressions given in these lemmas. Such bounds would yield smoothness estimates for the solution of the $\mathrm{VG}(r, 0, \sigma, \mu)$ Stein equation that are of order $O(r^{-1/2})$ for all $r > 0$.

A further open problem is to obtain uniform bounds for derivatives of all order of the solution of $\mathrm{VG}(r, 0, \sigma, \mu)$ Stein equation. The generator approach together with appropriate probabilistic arguments leads to such smoothness estimates for the solutions of the normal and Gamma Stein equations, and it is hoped that this technique will yield similar results for the solution of $\mathrm{VG}(r, 0, \sigma, \mu)$ Stein equation.

Another important problem is the extension of our smoothness estimates for solution of the $\mathrm{VG}(r, 0, \sigma, \mu)$ Stein equation to the solution of the more general $\mathrm{VG}(r, \theta, \sigma, \mu)$ Stein equation. It should be possible, if rather involved, to extend the calculations used for the case $\theta = 0$ to obtain a bound for general θ.

It would also be interesting to obtain smoothness estimates for the solution of the $\mathrm{PN}(r,\sigma^2)$ Stein equation for $r \geq 3$. As was mentioned in Remark 5.7, this is considered to be a difficult open problem. This is because the generator approach is not applicable and obtaining bounds via a direct calculation is perhaps unrealistic, since the solution of the Stein equation for $r \geq 3$ takes a complicated form that involves Meijer -G functions.

7.2.2 Stein's method for Generalised Hyperbolic distributions

The Variance-Gamma distributions form a subclass of a the *Generalised Hyperbolic* distributions. The Generalised Hyperbolic distributions are a superclass of many standard distributions and were introduced by Barndorff-Nielsen [10].

Definition 7.1 (**Generalised Hyperbolic Distribution**). *The random variable X is said to have a* Generalised Hyperbolic *distribution with parameters $\lambda,\alpha,\beta,\delta,\mu$, where $\lambda\in\mathbb{R}$, $\delta>0$, $\mu\in\mathbb{R}$, $\alpha>|\beta|\geq 0$, if and only if it has p.d.f. given by*

$$\begin{aligned} p_{GH}(x;\lambda,\alpha,\beta,\delta,\mu) = &\frac{(\alpha^2-\beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}K_{\lambda}(\delta\sqrt{\alpha^2-\beta^2})}\mathrm{e}^{\beta(x-\mu)}(\delta^2+(x-\mu)^2)^{(\lambda-\frac{1}{2})/2} \\ &\times K_{\lambda-\frac{1}{2}}(\alpha\sqrt{\delta^2+(x-\mu)^2}), \qquad x\in\mathbb{R}. \end{aligned} \tag{7.1}$$

If (7.1) holds then we write $X\sim GH(\lambda,\alpha,\beta,\delta,\mu)$.

For a detailed account on the subclasses of the Generalised Hyperbolic distributions see Eberlein and Hammerstein [21]. They show, for example, that a $GH(-\nu/2,0,0,\sqrt{\nu},0)$ random variable has a Student's t-distribution with ν degrees of freedom, and that a $GH(1,\alpha,\beta,\delta,\mu)$ random variable has a Hyperbolic distribution. Using the asymptotic formula $K_\lambda(x)\sim 2^{\lambda-1}\Gamma(\lambda)x^{-\lambda}$ as $x\downarrow 0$, for $\lambda>0$, we may easily verify that $GH(\lambda,\alpha,\beta,0,\mu)=\mathrm{VG}_2(\lambda,\alpha,\beta,\mu)$.

As the Generalised Hyperbolic distributions are a general class of distributions it would be interesting to extend the work on Chapter 3 for Variance-Gamma distributions to Generalised Hyperbolic distributions. We would hope to find a Stein equation that would reduce to the Variance-Gamma Stein equation (3.12) as $\delta\to 0$. Here we shall not consider in much detail how such a Stein equation could be obtained, however from a brief investigation there are good reasons to suggest that the resulting Hyperbolic Stein equations for the full parametrisation will be complicated.

Whilst a Stein equation for the full parametrisation may be complicated, it is possible to obtain Stein equations for important subclasses of the family. For example, we can obtain a Stein equation for Generalised Hyperbolic distributions in the limit $\delta\to 0$. By considering the asymptotic properties of the modified Bessel function of the second kind we see that p.d.f. of a $GH(-\frac{d}{2},0,0,\delta,\mu)$ distribution, where $d>0$, is given by

$$p(x)=\frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi\delta^2}\Gamma(\frac{d}{2})}\bigg(1+\frac{(x-\mu)^2}{\delta^2}\bigg)^{-\frac{1}{2}(d+1)}, \qquad x\in\mathbb{R}. \tag{7.2}$$

In the case $d=\nu$, $\delta=\sqrt{\nu}$ the density (7.2) is that of Student's t-distribution with ν degrees of freedom. A Stein equation for the $GH(-\frac{d}{2},0,0,\delta,\mu)$ distribution is an almost immediate consequence Theorem 1 of Schoutens [66], which we now state.

Theorem 7.2. *Suppose we have a random variable X on (a,b) with density function $\tilde{\rho}(x)$ and finite second moment and that $\rho(x)$ satisfies $(s(x)\tilde{\rho}(x))'=\tau(x)\tilde{\rho}(x)$, for some polynomials $s(x)$ of degree at most two and $\tau(x)$ of exact degree one. Then $\tilde{\rho}(x)=\rho(x)$ if and only if for all functions $f\in C^2$, $\mathbb{E}[s(X)f'(X)+\tau(X)f(X)]=0$.*

The density of the $GH(-\frac{d}{2},0,0,\delta,\mu)$ distribution satisfies $(s(x)\tilde{\rho}(x))'=\tau(x)\tilde{\rho}(x)$ with $s(x)=1+\frac{(x-\mu)^2}{\delta^2}$ and $\tau(x)=-\frac{(d-1)}{\delta^2}(x-\mu)$, and therefore a Stein equation for this class of distributions is given by

$$\bigg(1+\frac{(x-\mu)^2}{\delta^2}\bigg)f'(x)-\frac{(d-1)}{\delta^2}(x-\mu)f(x)=h(x)-GH^{-d/2,0}_{0,\delta,\mu}h, \tag{7.3}$$

where $GH^{\lambda,\alpha}_{\beta,\delta,\mu}h$ denotes the quantity $\mathbb{E}h(X)$ for $X\sim GH(\lambda,\alpha,\beta,\delta,\mu)$. Taking $d=\nu$, $\delta=\sqrt{\nu}$ and $\mu=0$ in (7.3) gives a Stein operator for a Student's t-distribution with ν degrees of freedom

$$\mathcal{A}f(x)=\bigg(1+\frac{x^2}{\nu}\bigg)f'(x)-\bigg(1-\frac{1}{\nu}\bigg)xf(x). \tag{7.4}$$

As would be expected, the Stein operator (7.4) is in agreement, up to a constant, with the Stein operator for a t-distribution that is given in Schoutens [66].

7.2.3 Limit theorems for the D_2^* statistic

Recall that the D_2^* statistic, which was introduced in Section 1.1, is given by

$$D_2^*=\sum_{\mathbf{w}\in\mathcal{A}^k}\frac{(X_\mathbf{w}-\bar{m}p_\mathbf{w})(Y_\mathbf{w}-\bar{n}p_\mathbf{w})}{\sqrt{\bar{m}\bar{n}}p_\mathbf{w}},$$

where,

$$X_\mathbf{w}=\sum_{i=1}^{\bar{m}}\mathbf{1}(A_i=w_1,\ldots,A_{i+k-1}=w_k),$$

and $\mathbf{Y}_w$ is defined similarly. Therefore D_2^* is of the form $D_2^* = \sum_{l=1}^{d^k} S_l T_l$, where $S_l = \frac{1}{\sqrt{\bar{m}}}\sum_{i=1}^{\bar{m}} X_{il}$, $T_l = \frac{1}{\sqrt{\bar{n}}}\sum_{j=1}^{\bar{n}} Y_{jl}$ and the X_{il} and Y_{jl} are mean zero random variables. Furthermore, the X_{il} and Y_{jl} are independent. Hence, by the central limit theorem, we would expect D_2^* to have an approximate Variance-Gamma distribution provided that the dependence between the amongst the word counts was not 'too large'. It would be interesting to apply Stein's method for Variance-Gamma approximation to obtain an explicit bound on the error in approximating D_2^* by a Variance-Gamma distribution, and therefore quantify what 'too large' means in this context.

We can write $D_2^* = g(\mathbf{W})$, where the entries of the vector $\mathbf{W}$ are a sum of standardised i.i.d. random variables and $g : \mathbb{R}^{2d^k} \rightarrow \mathbb{R}$ is a differentiable even function. Since $\mathbf{W}$ has an asymptotic multivariate normal distribution, we would in principle be able to use the multivariate normal Stein equation approach, which was developed in Chapter 6, to bound the error in approximating D_2^* by its limiting distribution (even when it is not Variance-Gamma distributed). We could also exploit the fact that g is an even function to achieve faster convergence rates than might otherwise be expected, and through the use of appropriate couplings we would hope to disentangle the complex dependence structure of the D_2^* statistic. Although, some rather complex calculations will need to be carried out in order achieve these bounds for the D_2^* statistic.

7.2.4 Further Open Questions

In addition to those problems mentioned above, a number of interesting questions have arisen from this thesis, and we mention some of those here.

Throughout this thesis we have restricted our attention to obtaining bounds in terms of smooth test functions. It would be desirable to obtain explicit error bounds for Variance-Gamma approximation in terms of non-smooth test functions, which would enable us to establish bounds in the classical Kolmogorov-Smirnov distance.

There are many possible ways that the theory developed in Chapter 6 could be extended. In Section 6.2 we considered the possibility of obtaining limit theorems under weakened assumptions, such as that the derivatives of the function $g : \mathbb{R} \rightarrow \mathbb{R}$ have growth that is faster than polynomial. An open problem is to obtain explicit bounds for these cases. When the derivatives of g have polynomial growth and the random variables are i.i.d. we saw that sufficient conditions for $O(n^{-1})$ convergence rates are that g is an even function or that the random variables have a vanishing third moment. An open problem is to determine whether these are also necessary conditions for $O(n^{-1})$ rates.

Finally, it would be desirable to obtain a bound of order J^{-1} for Friedman's χ^2 statistic, thereby improving on the $O(J^{-1/2})$ rate achieved in Theorem 6.24. We consider it possible to obtain such a bound by combining the exchangeable pair approach used in our proof of Theorem 6.24 and the symmetry considerations used by Pickett [55] that were used to establish a bound of order n^{-1} for Pearson's χ^2 statistic.

Appendix A

Proofs from the text

Here we prove some lemmas that appear in Chapter 3. For clarity, we present the statements of the lemmas and then prove them. We begin by Proving Lemma 3.13.

Lemma A.1. *Suppose $-1 < \beta < 1$ and $\nu \geq 0$, then there is at most one bounded solution to the Variance-Gamma Stein equation (3.13). Moreover, if $-1 < \beta < 1$, $\nu > -1/2$ and k is a positive integer then there is at most one solution which has a bounded k-th derivative.*

Proof. Suppose u and v are solutions to Variance-Gamma Stein equation (3.13) 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 ODE

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

This homogeneous ODE has general solution

$$w(x) = A\mathrm{e}^{-\beta x}x^{-\nu}K_\nu(x) + B\mathrm{e}^{-\beta x}x^{-\nu}I_\nu(x),$$

and its k-th derivative is given by

$$w^{(k)}(x) = A\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right) + B\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right).$$

Applying the asymptotic formulas (B.8) and (B.10) we have

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right) \sim \begin{cases} -\log x \quad \text{as } x \downarrow 0, & \nu = 0,\ k = 0 \\ 2^{|\nu|-1}\Gamma(|\nu|)x^{-(|\nu|+\nu+k)} \quad \text{as } x \downarrow 0, & \text{otherwise} \end{cases}$$

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right) \sim \frac{(1-\beta)^k\mathrm{e}^{(1-\beta)x}}{\sqrt{2\pi}x^{\nu+1/2}} \quad \text{as } x\rightarrow\infty.$$

Hence, in order to have a bounded solution or bounded k-th derivative we must take $B=0$. If $\nu\geq 0$ then for $w(x)$ to be bounded we must take $A=0$, and therefore $w=0$ and so $u=v$. If $\nu>-1/2$ then for $w^{(k)}(x)$, where $k\geq 1$, to be bounded we must take $A=0$, and therefore $w=0$ and so $u=v$. □

Here we prove Lemma 3.14.

Lemma A.2. *Let $h:\mathbb{R}\rightarrow\mathbb{R}$ be a measurable function with $\mathbb{E}|h(X)|<\infty$, where $X\sim\mathrm{VG}_2(\nu,1,\beta,0)$, and $\nu>-1/2$ and $-1<\beta<1$. Then a solution $f:\mathbb{R}\rightarrow\mathbb{R}$ to the Variance-Gamma Stein equation (3.13) is given by*

$$\begin{aligned} f(x) &= -\frac{\mathrm{e}^{-\beta x}K_\nu(|x|)}{|x|^\nu}\int_0^x \mathrm{e}^{\beta y}|y|^\nu I_\nu(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \\ &\quad -\frac{\mathrm{e}^{-\beta x}I_\nu(|x|)}{|x|^\nu}\int_x^\infty \mathrm{e}^{\beta y}|y|^\nu K_\nu(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \qquad (A.1) \\ &= -\frac{\mathrm{e}^{-\beta x}K_\nu(|x|)}{|x|^\nu}\int_0^x \mathrm{e}^{\beta y}|y|^\nu I_\nu(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \\ &\quad +\frac{\mathrm{e}^{-\beta x}I_\nu(|x|)}{|x|^\nu}\int_{-\infty}^x \mathrm{e}^{\beta y}|y|^\nu K_\nu(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y. \qquad (A.2) \end{aligned}$$

Suppose further that h is bounded, then f and f' are bounded for all $x\in\mathbb{R}$. Moreover, by Lemma A.1 this is the unique bounded solution for $\nu\geq 0$ and $-1<\beta<1$, and the unique solution with bounded k-th derivative for $\nu>-1/2$ and $-1<\beta<1$, where $k\geq 1$.

Proof. We use the method of variation of parameters (see, for example, Collins [19] for a detailed account of the method) to solve the equation. 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

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

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 $\mathrm{e}^{-\beta x}x^{-\nu}K_\nu(x)$ and $\mathrm{e}^{-\beta x}x^{-\nu}I_\nu(x)$. However, we take $f_1(x)=\mathrm{e}^{-\beta x}|x|^{-\nu}K_\nu(|x|)$ and $f_2(x)=\mathrm{e}^{-\beta x}|x|^{-\nu}I_\nu(|x|)$ 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.21) 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

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

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

Since

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

we have

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

where we used (B.22) to obtain the equality in the above display. Therefore the general solution to the inhomogeneous equation is given by

$$\begin{aligned}f(x)=&-\frac{\mathrm{e}^{-\beta x}K_\nu(|x|)}{|x|^\nu}\int_a^x \mathrm{e}^{\beta y}|y|^\nu I_\nu(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y\\&+\frac{\mathrm{e}^{-\beta x}I_\nu(|x|)}{|x|^\nu}\int_b^x \mathrm{e}^{\beta y}|y|^\nu K_\nu(|y|)[h(y)-\tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y.\end{aligned}$$

The solution is clearly bounded everywhere except possibly for $x=0$ or in the limits $x\to\pm\infty$. We therefore choose a and b to ensure 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 now choose b so that the solution is bounded in the limits $x \to \pm\infty$. If we take $b = \infty$ then we obtain solution (A.1). It now suffices to check that this solution is bounded as $x \to -\infty$. We now note that

$$\begin{aligned} 0 &= M^{-1}(\mathrm{VG}^{\nu,1}_{\beta,0}h - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h) \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y) - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y \\ &= \int_{-\infty}^{x} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y) - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y + \int_{x}^{\infty} \mathrm{e}^{\beta y}|y|^{\nu}K_{\nu}(|y|)[h(y) - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h]\,\mathrm{d}y, \end{aligned}$$

where M is the normalizing constant for the density function. The final equality shows that solutions (A.1) and (A.2) are equal. The second equality shows that the solution is bounded as $x \to -\infty$ and thus we have shown that the solution is bounded.

We now prove that f and its first derivative are bounded for all $x \in \mathbb{R}$. Straightforward calculations (see Lemma 3.17) show that for $x \geq 0$,

$$\begin{aligned} \|f\| &\leq \|\tilde{h}\|\left\|I_{(\nu,\beta,1)}(x)\frac{\mathrm{e}^{-\beta x}K_{\nu}(x)}{x^{\nu}}\right\| + \|\tilde{h}\|\left\|\frac{\mathrm{e}^{-\beta x}I_{\nu}(x)}{x^{\nu}}\int_{x}^{\infty}\mathrm{e}^{\beta y}y^{\nu}K_{\nu}(y)\,\mathrm{d}y\right\|, \\ \|f'\| &\leq \|\tilde{h}\|\left\|I_{(\nu,\beta,1)}(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{-\beta x}K_{\nu}(x)}{x^{\nu}}\right)\right\| + \|\tilde{h}\|\left\|\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{-\beta x}I_{\nu}(x)}{x^{\nu}}\right)\int_{x}^{\infty}\mathrm{e}^{\beta y}y^{\nu}K_{\nu}(y)\,\mathrm{d}y\right\|, \end{aligned}$$

where $\tilde{h} = h(x) - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h$. We may apply Lemmas D.3 and D.4 to bound the supremum norms involving modified Bessel functions that appear in the above display. This gives, for $\nu \geq 1/2$,

$$\begin{aligned} \|f\| &\leq \|h - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h\|\left\{\frac{\sqrt{\pi}\Gamma(\nu+1/2)}{1-|\beta|} + \sup_{x\geq 0}\frac{2}{2\nu+1}xK_{\nu+1}I_{\nu}(x)\right\} < \infty, \\ \|f'\| &\leq \|h - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h\|\left\{\frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{1-|\beta|} + \sup_{x\geq 0}\frac{2(\beta+1)}{2\nu+1}xK_{\nu+1}I_{\nu}(x)\right\} < \infty, \end{aligned}$$

and for $-1/2 < \nu < 1/2$,

$$\begin{aligned} \|f\| &\leq \|h - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h\|\left\{\frac{(\mathrm{e}+1)\sqrt{2}\Gamma(\nu+1/2)}{1-|\beta|} + \sup_{x\geq 0}\frac{2}{2\nu+1}xK_{\nu+1}I_{\nu}(x)\right\} < \infty, \\ \|f'\| &\leq \|h - \tilde{\mathrm{VG}}^{\nu,1}_{\beta,0}h\|\left\{\frac{2(\mathrm{e}+1)\sqrt{2}\Gamma(\nu+1/2)}{1-|\beta|} + \sup_{x\geq 0}\frac{2(\beta+1)}{2\nu+1}xK_{\nu+1}I_{\nu}(x)\right\} < \infty. \end{aligned}$$

Recalling Remark 3.15, it is sufficient to bound these norms 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 $x \in \mathbb{R}$. □

We now bound the third and fourth derivatives of solution of the $\mathrm{VG}_2(\nu, 1, 0, 0)$ Stein equation (3.13) in the region $x \geq 1$. This completes the proof of Lemma 3.19. In the proof we shall use the operator $(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x})^n$, which was defined in Subsection 3.3.2 (see (3.28)). In particular, we shall use the operator for $n=2$, 3 and 4:

$$\begin{aligned}
\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{g(x)}{x}\right)\right)^2 &= \frac{g''(x)}{x^3}-\frac{3g'(x)}{x^4}+\frac{3g(x)}{x^5}, \\
\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{g(x)}{x}\right)\right)^3 &= \frac{g^{(3)}(x)}{x^4}-\frac{6g''(x)}{x^5}+\frac{15g'(x)}{x^6}-\frac{15g(x)}{x^7}, \\
\left(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{g(x)}{x}\right)\right)^4 &= \frac{g^{(4)}(x)}{x^5}-\frac{10g^{(3)}(x)}{x^6}+\frac{45g''(x)}{x^7}-\frac{105g'(x)}{x^8}+\frac{105g(x)}{x^9}.
\end{aligned}$$

Lemma A.3. *Suppose that $h\in C_b^3(\mathbb{R})$. Then for $x\geq 1$ the following bound holds for the third and fourth derivative of the solution to the $\mathrm{VG}_2(\nu,1,0,0)$ Stein equation (3.13):*

$$\begin{aligned}
\sup_{x\geq 1}|f^{(3)}(x)| &\leq 5\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\bigg)\|h''\|+18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\bigg)\|h'\| \\
&\quad +18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|; \\
\sup_{x\geq 1}|f^{(4)}(x)| &\leq 8\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+13}}+\frac{1}{2\nu+4}\bigg)\|h^{(3)}\|+52\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+3}\bigg)\|h''\| \\
&\quad +123\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+2}\bigg)\|h'\|+123\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}}+\frac{1}{2\nu+1}\bigg)\|h-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}h\|.
\end{aligned}$$

Proof. We begin by obtaining the bound for $\|f^{(3)}\|$. Let $\tilde{h}(x)=h(x)-\tilde{\mathrm{VG}}_{0,0}^{\nu,1}$. Substituting the third derivative formulas (B.39) and (B.40) for $x^{-\nu}I_\nu(x)$ and $K_\nu(x)$ into formula (3.17) for $f^{(3)}(x)$ gives

$$\begin{aligned}
f^{(3)}(x) = \frac{h'(x)}{x}-\frac{(2\nu+2)\tilde{h}(x)}{x^2}-\bigg(-\frac{K_{\nu+3}(x)}{x^\nu}+\frac{3K_{\nu+2}(x)}{x^{\nu+1}}\bigg)\int_0^x y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\
-\bigg(\frac{I_{\nu+3}(x)}{x^\nu}+\frac{3I_{\nu+2}(x)}{x^{\nu+1}}\bigg)\int_x^\infty y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y,
\end{aligned}$$

where we took $\beta=0$. Integrating by parts, using that $\frac{\mathrm{d}}{\mathrm{d}x}(x^{\nu+1}I_{\nu+1}(x))=x^{\nu+1}I_\nu(x)$ and

$\frac{\mathrm{d}}{\mathrm{d}x}(x^{\nu+1}K_{\nu+1}(x)) = -x^{\nu+1}K_\nu(x)$ (see (B.29) and (B.30)), gives

$$\begin{aligned}
f^{(3)}(x) &= \frac{h'(x)}{x} - \frac{(2\nu+2)\tilde{h}(x)}{x^2} + \frac{K_{\nu+3}(x)}{x^\nu}\bigg[x^\nu I_{\nu+1}(x)\tilde{h}(x) - x^{\nu+1}I_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\
&\quad + x^{\nu+3}I_{\nu+3}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^2\bigg(\frac{\tilde{h}(x)}{x}\bigg) - \int_0^x y^{\nu+4}I_{\nu+3}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^3\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\
&\quad - \frac{3K_{\nu+2}(x)}{x^{\nu+1}}\bigg[x^\nu I_{\nu+1}(x)\tilde{h}(x) - x^{\nu+1}I_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\
&\quad + \int_0^x y^{\nu+3}I_{\nu+2}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^2\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\
&\quad - \frac{I_{\nu+3}(x)}{x^\nu}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + x^{\nu+1}K_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\
&\quad + x^{\nu+3}K_{\nu+3}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^2\bigg(\frac{\tilde{h}(x)}{x}\bigg) + \int_x^\infty y^{\nu+4}K_{\nu+3}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^3\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\
&\quad - \frac{3I_{\nu+2}(x)}{x^{\nu+1}}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + x^{\nu+1}K_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\
&\quad + \int_x^\infty y^{\nu+3}K_{\nu+2}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^2\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg],
\end{aligned}$$

Cancelling terms and using the Wronskian formula $I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x) = \frac{1}{x}$ and the identity $I_\nu(x)K_{\nu+2}(x) - I_{\nu+2}(x)K_\nu(x) = \frac{2(1+\nu)}{x^2}$ (see (B.22) and (B.23)) gives

$$\begin{aligned}
f^{(3)}(x) = &-\frac{K_{\nu+3}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+3}(y)\bigg(h^{(3)}(y) - \frac{6h''(y)}{y} + \frac{15h'(y)}{y^2} - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y \\
&- \frac{K_{\nu+2}(x)}{x^{\nu+1}}\int_0^x y^\nu I_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\
&- \frac{I_{\nu+3}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+3}(y)\bigg(h^{(3)}(y) - \frac{6h''(y)}{y} + \frac{15h'(y)}{y^2} - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y \\
&- \frac{I_{\nu+2}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y.
\end{aligned}$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^\nu I_{\nu+3}(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(x^{-3}\cdot x^{\nu+3}I_{\nu+3}(x)) = x^\nu I_{\nu+2}(x) - 3x^{\nu-1}I_{\nu+3}(x),$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^\nu K_{\nu+3}(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(x^{-3}\cdot x^{\nu+3}K_{\nu+3}(x)) = -x^\nu K_{\nu+2}(x) - 3x^{\nu-1}K_{\nu+3}(x).$$

Therefore, by integration by parts, we have

$$-\frac{K_{\nu+3}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+3}(y)h^{(3)}(y)\,\mathrm{d}y = h''(x)I_{\nu+3}(x)K_{\nu+3}(x) + \frac{K_{\nu+3}(x)}{x^\nu}\int_0^x h''(y)[y^\nu I_{\nu+2}(y) - 3y^{\nu-1}I_{\nu+3}(y)]\,\mathrm{d}y,$$

and

$$-\frac{I_{\nu+3}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+3}(y)h^{(3)}(y)\,\mathrm{d}y = -h''(x)I_{\nu+3}(x)K_{\nu+3}(x) - \frac{I_{\nu+3}(x)}{x^\nu}\int_x^\infty h''(y)[y^\nu K_{\nu+2}(y) + 3y^{\nu-1}I_{\nu+3}(y)]\,\mathrm{d}y.$$

Thus, we can write $f^{(3)}$ as

$$\begin{aligned} f^{(3)}(x) &= \frac{K_{\nu+3}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+2}(y)h''(y)\,\mathrm{d}y - \frac{K_{\nu+3}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+3}(y)\bigg(-\frac{3h''(y)}{y} + \frac{15h'(y)}{y^2} \\ &\quad - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y - \frac{K_{\nu+2}(x)}{x^{\nu+1}}\int_0^x y^\nu I_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\ &\quad - \frac{I_{\nu+3}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+2}(y)h''(y)\,\mathrm{d}y - \frac{I_{\nu+3}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+3}(y)\bigg(-\frac{3h''(y)}{y} + \frac{15h'(y)}{y^2} \\ &\quad - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y - \frac{I_{\nu+2}(x)}{x^{\nu+1}}\int_x^\infty y^\nu K_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y. \end{aligned}$$

We bound the terms, for $x \geq 1$, using Lemma D.8 to obtain

$$\begin{aligned} \sup_{x\geq 1}|f^{(3)}(x)| &\leq \bigg(\frac{1}{2\nu+3} + \frac{3}{2\nu+3} + \frac{1}{2\nu+3} + \frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{3\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{\sqrt{\pi}}{\sqrt{4\nu+9}}\bigg)\|h''\| \\ &\quad + \bigg(\frac{15}{2\nu+2} + \frac{3}{2\nu+2} + \frac{15\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{3\sqrt{\pi}}{\sqrt{4\nu+9}}\bigg)\|h'\| \\ &\quad + \bigg(\frac{15}{2\nu+1} + \frac{3}{2\nu+1} + \frac{15\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{3\sqrt{\pi}}{\sqrt{4\nu+9}}\bigg)\|\tilde{h}\| \\ &\leq 5\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{1}{2\nu+3}\bigg)\|h''\| + 18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{1}{2\nu+2}\bigg)\|h'\| \\ &\quad + 18\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{1}{2\nu+1}\bigg)\|\tilde{h}\|, \end{aligned}$$

as required.

We now obtain the bound for $\|f^{(4)}\|$. Substituting the fourth derivative formulas (B.41) and

(B.42) for $x^{-\nu}I_\nu(x)$ and $K_\nu(x)$ into formula (3.18) for $f^{(4)}(x)$ gives

$$\begin{aligned}f^{(4)}(x) &= \frac{h''(x)}{x} - \frac{(2\nu+3)h'(x)}{x^2} + \bigg(\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x}\bigg)\tilde{h}(x) \\ &\quad - \bigg(\frac{K_{\nu+4}(x)}{x^\nu} - \frac{6K_{\nu+3}(x)}{x^{\nu+1}} + \frac{3K_{\nu+2}(x)}{x^{\nu+2}}\bigg)\int_0^x y^\nu I_\nu(y)\tilde{h}(y)\,\mathrm{d}y \\ &\quad - \bigg(\frac{I_{\nu+4}(x)}{x^\nu} + \frac{6I_{\nu+3}(x)}{x^{\nu+1}} + \frac{3I_{\nu+2}(x)}{x^{\nu+2}}\bigg)\int_x^\infty y^\nu K_\nu(y)\tilde{h}(y)\,\mathrm{d}y,\end{aligned}$$

where we took $\beta = 0$. Integrating by parts, using the differentiation formulas (B.29) and (B.30), gives the following expression for $f^{(4)}(x)$:

$$\begin{aligned}&\frac{h''(x)}{x} - \frac{(2\nu+3)h'(x)}{x^2} + \bigg(\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x}\bigg)\tilde{h}(x) \\ &- \frac{K_{\nu+4}(x)}{x^\nu}\bigg[x^\nu I_{\nu+1}(x)\tilde{h}(x) - x^{\nu+1}I_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) + x^{\nu+3}I_{\nu+3}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^2\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &- x^{\nu+4}I_{\nu+4}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^3\bigg(\frac{\tilde{h}(x)}{x}\bigg) + \int_0^x y^{\nu+5}I_{\nu+4}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^4\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\ &+ \frac{6K_{\nu+3}(x)}{x^{\nu+1}}\bigg[x^\nu I_{\nu+1}(x)\tilde{h}(x) - x^{\nu+1}I_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) + x^{\nu+3}I_{\nu+3}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^2\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &- \int_0^x y^{\nu+4}I_{\nu+3}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^3\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\ &- \frac{3K_{\nu+2}(x)}{x^{\nu+1}}\bigg[x^\nu I_{\nu+1}(x)\tilde{h}(x) - x^{\nu+1}I_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) + \int_0^x y^{\nu+3}I_{\nu+2}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^2\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\ &- \frac{I_{\nu+4}(x)}{x^\nu}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + x^{\nu+1}K_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) + x^{\nu+3}K_{\nu+3}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^2\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &+ x^{\nu+4}K_{\nu+4}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^3\bigg(\frac{\tilde{h}(x)}{x}\bigg) + \int_x^\infty y^{\nu+5}K_{\nu+4}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^4\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\ &- \frac{6I_{\nu+3}(x)}{x^{\nu+1}}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + x^{\nu+1}K_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) + x^{\nu+3}K_{\nu+3}(x)\bigg(\frac{1}{x}\frac{\mathrm{d}}{\mathrm{d}x}\bigg)^2\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &+ \int_x^\infty y^{\nu+4}K_{\nu+3}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^3\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg] \\ &- \frac{3I_{\nu+2}(x)}{x^{\nu+2}}\bigg[x^\nu K_{\nu+1}(x)\tilde{h}(x) + x^{\nu+1}K_{\nu+2}(x)\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{\tilde{h}(x)}{x}\bigg) \\ &+ \int_x^\infty y^{\nu+3}K_{\nu+2}(y)\bigg(\frac{1}{y}\frac{\mathrm{d}}{\mathrm{d}y}\bigg)^2\bigg(\frac{\tilde{h}(y)}{y}\bigg)\,\mathrm{d}y\bigg].\end{aligned}$$

Cancelling terms and using the identities $I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x) = \frac{1}{x}$, $I_\nu(x)K_{\nu+2}(x) - I_{\nu+2}(x)K_\nu(x) = \frac{2(1+\nu)}{x^2}$ and $I_\nu(x)K_{\nu+3}(x) + I_{\nu+3}(x)K_\nu(x) = \frac{1}{x} + \frac{4(\nu+1)(\nu+2)}{x^3}$ (see (B.22), (B.23) and (B.24)) gives

$$\begin{aligned}
f^{(4)}(x) = &-\frac{K_{\nu+4}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+4}(y)\bigg(h^{(4)}(y) - \frac{10h^{(3)}(y)}{y} + \frac{45h''(y)}{y^2} - \frac{105h'(y)}{y^3} + \frac{105\tilde{h}(y)}{y^4}\bigg)\,\mathrm{d}y \\
&- \frac{K_{\nu+3}(x)}{x^{\nu+1}}\int_0^x y^\nu I_{\nu+3}(y)\bigg(h^{(3)}(y) - \frac{6h''(y)}{y} + \frac{15h'(y)}{y^2} - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y \\
&- \frac{K_{\nu+2}(x)}{x^{\nu+2}}\int_0^x y^\nu I_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\
&- \frac{I_{\nu+4}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+4}(y)\bigg(h^{(4)}(y) - \frac{10h^{(3)}(y)}{y} + \frac{45h''(y)}{y^2} - \frac{105h'(y)}{y^3} + \frac{105\tilde{h}(y)}{y^4}\bigg)\,\mathrm{d}y \\
&- \frac{I_{\nu+3}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+3}(y)\bigg(h^{(3)}(y) - \frac{6h''(y)}{y} + \frac{15h'(y)}{y^2} - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y \\
&- \frac{I_{\nu+2}(x)}{x^{\nu+2}}\int_0^x y^\nu K_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y.
\end{aligned}$$

Note that

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^\nu I_{\nu+4}(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(x^{-4}\cdot x^{\nu+4}I_{\nu+4}(x)) = x^\nu I_{\nu+3}(x) - 4x^{\nu-1}I_{\nu+4}(x),$$

and

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^\nu K_{\nu+4}(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(x^{-4}\cdot x^{\nu+4}K_{\nu+4}(x)) = -x^\nu K_{\nu+3}(x) - 4x^{\nu-1}K_{\nu+4}(x).$$

Therefore, by integration by parts, we have

$$\begin{aligned}
-\frac{K_{\nu+4}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+4}(y)h^{(4)}(y)\,\mathrm{d}y &= h^{(3)}(x)I_{\nu+4}(x)K_{\nu+4}(x) \\
&\quad + \frac{K_{\nu+4}(x)}{x^\nu}\int_0^x h^{(3)}(y)[y^\nu I_{\nu+3}(y) - 4y^{\nu-1}I_{\nu+4}(y)]\,\mathrm{d}y,
\end{aligned}$$

and

$$\begin{aligned}
-\frac{I_{\nu+4}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+4}(y)h^{(4)}(y)\,\mathrm{d}y &= -h^{(3)}(x)I_{\nu+4}(x)K_{\nu+4}(x) \\
&\quad - \frac{I_{\nu+4}(x)}{x^\nu}\int_x^\infty h^{(3)}(y)[y^\nu K_{\nu+3}(y) + 4y^{\nu-1}I_{\nu+4}(y)]\,\mathrm{d}y.
\end{aligned}$$

Thus, we can write $f^{(4)}$ as

$$
\begin{aligned}
f^{(4)}(x) &= \frac{K_{\nu+4}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+3}(y)h^{(3)}(y)\,\mathrm{d}y - \frac{K_{\nu+4}(x)}{x^\nu}\int_0^x y^\nu I_{\nu+4}(y)\bigg(-\frac{10h^{(3)}(y)}{y} \\
&\quad + \frac{45h''(y)}{y^2} - \frac{105h'(y)}{y^3} + \frac{105\tilde{h}(y)}{y^4}\bigg)\,\mathrm{d}y \\
&\quad - \frac{K_{\nu+3}(x)}{x^{\nu+1}}\int_0^x y^\nu I_{\nu+3}(y)\bigg(h^{(3)}(y) - \frac{6h''(y)}{y} + \frac{15h'(y)}{y^2} - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y \\
&\quad - \frac{K_{\nu+2}(x)}{x^{\nu+2}}\int_0^x y^\nu I_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y \\
&\quad - \frac{I_{\nu+4}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+3}(y)h^{(3)}(y)\,\mathrm{d}y - \frac{I_{\nu+4}(x)}{x^\nu}\int_x^\infty y^\nu K_{\nu+4}(y)\bigg(-\frac{10h^{(3)}(y)}{y} \\
&\quad + \frac{45h''(y)}{y^2} - \frac{105h'(y)}{y^3} + \frac{105\tilde{h}(y)}{y^4}\bigg)\,\mathrm{d}y \\
&\quad - \frac{I_{\nu+3}(x)}{x^{\nu+1}}\int_x^\infty y^\nu K_{\nu+3}(y)\bigg(h^{(3)}(y) - \frac{6h''(y)}{y} + \frac{15h'(y)}{y^2} - \frac{15\tilde{h}(y)}{y^3}\bigg)\,\mathrm{d}y \\
&\quad - \frac{I_{\nu+2}(x)}{x^{\nu+2}}\int_0^x y^\nu K_{\nu+2}(y)\bigg(h''(y) - \frac{3h'(y)}{y} + \frac{3\tilde{h}(y)}{y^2}\bigg)\,\mathrm{d}y.
\end{aligned}
$$

We bound the terms, for $x \geq 1$, using Lemma D.8 to obtain

$$
\begin{aligned}
\sup_{x\geq 1}|f^{(4)}(x)| &\leq \bigg(\frac{1}{2\nu+4} + \frac{6}{2\nu+4} + \frac{1}{2\nu+4} + \frac{\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{6\sqrt{\pi}}{\sqrt{4\nu+17}} + \frac{\sqrt{\pi}}{\sqrt{4\nu+13}}\bigg)\|h^{(3)}\| \\
&\quad + \bigg(\frac{45}{2\nu+3} + \frac{6}{2\nu+3} + \frac{1}{2\nu+3} + \frac{45\sqrt{\pi}}{\sqrt{4\nu+17}} + \frac{6\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{\sqrt{\pi}}{\sqrt{4\nu+9}}\bigg)\|h''\| \\
&\quad + \bigg(\frac{105}{2\nu+2} + \frac{15}{2\nu+2} + \frac{3}{2\nu+2} + \frac{105\sqrt{\pi}}{\sqrt{4\nu+17}} + \frac{15\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{3\sqrt{\pi}}{\sqrt{4\nu+9}}\bigg)\|h^{(3)}\| \\
&\quad + \bigg(\frac{105}{2\nu+1} + \frac{15}{2\nu+1} + \frac{3}{2\nu+1} + \frac{\sqrt{105\pi}}{\sqrt{4\nu+17}} + \frac{15\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{3\sqrt{\pi}}{\sqrt{4\nu+9}}\bigg)\|h^{(3)}\| \\
&\leq 8\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+13}} + \frac{1}{2\nu+4}\bigg)\|h^{(3)}\| + 52\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{1}{2\nu+3}\bigg)\|h''\| \\
&\quad + 123\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{1}{2\nu+2}\bigg)\|h'\| + 123\bigg(\frac{\sqrt{\pi}}{\sqrt{4\nu+9}} + \frac{1}{2\nu+1}\bigg)\|\tilde{h}\|,
\end{aligned}
$$

as required. □

Appendix B

Elementary properties of modified Bessel functions

Here we list standard properties of modified Bessel functions that are used throughout this thesis. All these formulas can be found in Olver et al. [52], unless otherwise stated. Inequalities (B.12) and (B.13) can be found in Jones [38] and Nåsell [49]; inequalities (B.14), (B.15) and (B.16) can easily be deduced from (B.4); inequality (B.17) is given in Luke [45]. The integration formulas can be found in Gradshetyn and Ryzhik [34].

B.1 Basic definitions

The *modified Bessel function of the first kind* of order $\nu \in \mathbb{R}$ can defined for $x > 0$ by the following infinite series

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

The series converges for all $x > 0$.

The *modified Bessel function of the second kind* of order $\nu \in \mathbb{R}$ can be defined for $x > 0$ in terms of the modified Bessel function of the first kind as follows

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

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

B.2 Basic properties

The modified Bessel functions $I_\nu(x)$ and $K_\nu(x)$ are both regular functions of x. If $\nu > -1$ then $I_\nu(x)$ is strictly positive for all $x > 0$. If $\nu \in \mathbb{R}$ then $K_\nu(x)$ is strictly positive for all $x > 0$. The modified Bessel functions take the following values at the origin and positive infinity:

$$\begin{aligned} I_\nu(0) &= 0, \quad \nu > 0, \\ I_0(0) &= 1, \\ \lim_{x\to\infty} K_\nu(x) &= 0, \quad \nu \in \mathbb{R}. \end{aligned}$$

B.3 Series expansions

$$K_\nu(x) = \frac{\pi}{2\sin(\nu\pi)} \left(\sum_{k=0}^{\infty} \frac{1}{\Gamma(k-\nu+1)k!} \left(\frac{x}{2}\right)^{2k-\nu} - \sum_{k=0}^{\infty} \frac{1}{\Gamma(\nu+k+1)k!} \left(\frac{x}{2}\right)^{\nu+2k} \right), \quad \nu \notin \mathbb{Z}, \tag{B.2}$$

$$\begin{aligned} K_\nu(x) = (-1)^{\nu-1} \left\{ \log\left(\frac{x}{2}\right) + \gamma \right\} I_\nu(x) + \frac{1}{2} \sum_{k=0}^{\nu-1} \frac{(-1)^k(\nu-k-1)!}{k!} \left(\frac{x}{2}\right)^{2k-\nu} \\ + \frac{(-1)^\nu}{2} \sum_{k=0}^{\infty} \frac{\{\psi(k)+\psi(\nu+k)\}}{k!\,(\nu+k)!} \left(\frac{x}{2}\right)^{\nu+2k}, \quad \nu \in \mathbb{N}, \end{aligned} \tag{B.3}$$

where γ is the Euler-Mascheroni constant and $\psi(k) = \sum_{j=1}^{k} \frac{1}{j}$.

B.4 Integral representations

$$K_\nu(x) = \int_0^\infty \mathrm{e}^{-x\cosh(t)} \cosh(\nu t)\,\mathrm{d}t, \qquad x > 0. \tag{B.4}$$

B.5 Spherical Bessel functions

$$I_{-1/2}(x) = \sqrt{\frac{2}{\pi x}} \cosh(x), \tag{B.5}$$

$$K_{1/2}(x) = K_{-1/2}(x) = \sqrt{\frac{\pi}{2x}} \mathrm{e}^{-x}. \tag{B.6}$$

B.6 Asymptotic expansions

$$
\begin{aligned}
I_\nu(x) &\sim \frac{1}{\Gamma(\nu+1)}\left(\frac{x}{2}\right)^\nu, \qquad x\downarrow 0, && \text{(B.7)}\\
K_\nu(x) &\sim \begin{cases} 2^{|\nu|-1}\Gamma(|\nu|)x^{-|\nu|}, & x\downarrow 0,\ \nu\neq 0,\\ -\log x, & x\downarrow 0,\ \nu=0,\end{cases} && \text{(B.8)}\\
K_\nu(x) &\sim 2^{\nu-1}\Gamma(\nu)x^{-\nu}-2^{\nu-3}\Gamma(\nu-1)x^{-\nu+2}, \qquad x\downarrow 0,\ \nu>1, && \text{(B.9)}\\
I_\nu(x) &\sim \frac{e^x}{\sqrt{2\pi x}}, \quad x\to\infty, && \text{(B.10)}\\
K_\nu(x) &\sim \sqrt{\frac{\pi}{2x}}e^{-x}, \quad x\to\infty. && \text{(B.11)}
\end{aligned}
$$

B.7 Inequalities

Let $x>0$, then following inequalities hold

$$
\begin{aligned}
I_\mu(x) &< I_\nu(x), \quad 0\le\nu<\mu, && \text{(B.12)}\\
I_\nu(x) &< I_{\nu-1}(x), \quad \nu\ge 1/2, && \text{(B.13)}\\
K_\mu(x) &> K_\nu(x), \quad 0\le\nu<\mu, && \text{(B.14)}\\
K_\nu(x) &< K_{\nu-1}(x), \quad \nu<1/2, && \text{(B.15)}\\
K_\nu(x) &\ge K_{\nu-1}(x), \quad \nu\ge 1/2. && \text{(B.16)}
\end{aligned}
$$

We have equality in (B.16) if and only if $\nu=1/2$.

$$\Gamma(\nu+1)\left(\frac{2}{x}\right)^\nu I_\nu(x)<\cosh(x), \qquad x>0, \quad \nu>-\frac{1}{2}. \tag{B.17}$$

B.8 Identities

$$
\begin{aligned}
K_{-\nu}(x) &= K_\nu(x), && \text{(B.18)}\\
I_{\nu+1}(x) &= I_{\nu-1}(x)-\frac{2\nu}{x}I_\nu(x), && \text{(B.19)}\\
K_{\nu+1}(x) &= K_{\nu-1}(x)+\frac{2\nu}{x}K_\nu(x), && \text{(B.20)}\\
K_\nu(-x) &= (-1)^\nu K_\nu(x)-\pi i I_\nu(x), && \text{(B.21)}
\end{aligned}
$$

$$
\begin{aligned}
I_\nu(x)K_{\nu+1}(x) + I_{\nu+1}(x)K_\nu(x) &= \frac{1}{x}, && \text{(B.22)}\\
I_\nu(x)K_{\nu+2}(x) - I_{\nu+2}(x)K_\nu(x) &= \frac{2(\nu+1)}{x^2}, && \text{(B.23)}\\
I_\nu(x)K_{\nu+3}(x) + I_{\nu+3}(x)K_\nu(x) &= \frac{1}{x} + \frac{4(\nu+1)(\nu+2)}{x^3}. && \text{(B.24)}
\end{aligned}
$$

Identity (B.23) can easily be verified using the identities (B.19), (B.20) and (B.22). Identity (B.24) also follows easily from identities (B.19), (B.20), (B.22) and (B.23).

B.9 Differentiation

$$
\begin{aligned}
I_0'(x) &= I_1(x), \\
K_0'(x) &= -K_1(x), \\
\frac{\mathrm{d}}{\mathrm{d}x}(I_\nu(x)) &= \frac{1}{2}(I_{\nu+1}(x) - I_{\nu-1}(x)), && \text{(B.25)}\\
\frac{\mathrm{d}}{\mathrm{d}x}(K_\nu(x)) &= -\frac{1}{2}(K_{\nu+1}(x) + K_{\nu-1}(x)), && \text{(B.26)}\\
\frac{\mathrm{d}}{\mathrm{d}x}(K_\nu(x)) &= -K_{\nu-1}(x) - \frac{\nu}{x}K_\nu(x), && \text{(B.27)}\\
\frac{\mathrm{d}}{\mathrm{d}x}(K_\nu(x)) &= -K_{\nu+1}(x) + \frac{\nu}{x}K_\nu(x), && \text{(B.28)}\\
\frac{\mathrm{d}}{\mathrm{d}x}(x^\nu I_\nu(x)) &= x^\nu I_{\nu-1}(x), && \text{(B.29)}\\
\frac{\mathrm{d}}{\mathrm{d}x}(x^\nu K_\nu(x)) &= -x^\nu K_{\nu-1}(x), && \text{(B.30)}\\
\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{I_\nu(x)}{x^\nu}\right) &= \frac{I_{\nu+1}(x)}{x^\nu}, && \text{(B.31)}\\
\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{K_\nu(x)}{x^\nu}\right) &= -\frac{K_{\nu+1}(x)}{x^\nu}, && \text{(B.32)}\\
\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{I_\nu(x)}{x^\nu}\right) &= \frac{I_\nu(x)}{x^\nu} - \frac{(2\nu+1)I_{\nu+1}(x)}{x^{\nu+1}}, && \text{(B.33)}\\
\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{K_\nu(x)}{x^\nu}\right) &= \frac{K_\nu(x)}{x^\nu} + \frac{(2\nu+1)K_{\nu+1}(x)}{x^{\nu+1}}, && \text{(B.34)}\\
\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{I_\nu(x)}{x^\nu}\right) &= -\frac{(2\nu+1)I_\nu(x)}{x^\nu} + \left(1 + \frac{(2\nu+1)(2\nu+2)}{x^2}\right)\frac{I_{\nu+1}(x)}{x^\nu}, && \text{(B.35)}\\
\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right) &= -\frac{(2\nu+1)K_\nu(x)}{x^\nu} - \left(1 + \frac{(2\nu+1)(2\nu+2)}{x^2}\right)\frac{K_{\nu+1}(x)}{x^\nu}, && \text{(B.36)}\\
\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{I_\nu(x)}{x^\nu}\right) &= \frac{I_{\nu+2}(x)}{x^\nu} + \frac{I_{\nu+1}(x)}{x^{\nu+1}}, && \text{(B.37)}
\end{aligned}
$$

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{K_\nu(x)}{x^\nu}\right) = \frac{K_{\nu+2}(x)}{x^\nu} - \frac{K_{\nu+1}(x)}{x^{\nu+1}}, \tag{B.38}$$

$$\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{I_\nu(x)}{x^\nu}\right) = \frac{I_{\nu+3}(x)}{x^\nu} + \frac{3I_{\nu+2}(x)}{x^{\nu+1}}, \tag{B.39}$$

$$\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right) = -\frac{K_{\nu+3}(x)}{x^\nu} + \frac{3K_{\nu+2}(x)}{x^{\nu+1}}, \tag{B.40}$$

$$\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{I_\nu(x)}{x^\nu}\right) = \frac{I_{\nu+4}(x)}{x^\nu} + \frac{6I_{\nu+3}(x)}{x^{\nu+1}} + \frac{3I_{\nu+2}(x)}{x^{\nu+2}}, \tag{B.41}$$

$$\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right) = \frac{K_{\nu+4}(x)}{x^\nu} - \frac{6K_{\nu+3}(x)}{x^{\nu+1}} + \frac{3K_{\nu+2}(x)}{x^{\nu+2}}. \tag{B.42}$$

It is straightforward to verify formulas (B.33) and (B.35) using the differentiation formula (B.31) and identity (B.19). Formulas (B.34) and (B.36) can also be verified using the differentiation formula (B.32) and identity (B.20). It is easy to verify formulas (B.37), (B.39) and (B.41) using the differentiation formula (B.31). Finally, formulas (B.38), (B.40) and (B.42) can be verified using the differentiation formula (B.32).

B.10 Integration

$$\int_{-\infty}^{\infty} \mathrm{e}^{\beta t}|t|^\nu K_\nu(|t|)\,\mathrm{d}t = \frac{\sqrt{\pi}\Gamma(\nu+1/2)2^\nu}{(1-\beta^2)^{\nu+1/2}}, \quad \nu > -1/2,\ -1<\beta<1, \tag{B.43}$$

$$\int_0^\infty t^{-1}\exp\left(-\frac{1}{2}t - \frac{x^2+y^2}{2t}\right) I_\nu\left(\frac{xy}{t}\right)\,\mathrm{d}t = \begin{cases} 2I_\nu(x)K_\nu(y), & 0<x<y, \\ 2K_\nu(x)I_\nu(y), & 0<y<x, \end{cases} \tag{B.44}$$

$$\int_0^x \mathrm{e}^{-t}t^\nu I_\nu(t)\,\mathrm{d}t = \frac{\mathrm{e}^{-x}x^{\nu+1}}{2\nu+1}[I_\nu(x)+I_{\nu+1}(x)], \quad \nu > -1/2. \tag{B.45}$$

B.11 Modified Bessel differential equation

The modified Bessel differential equation is

$$x^2 f''(x) + xf'(x) - (x^2+\nu^2)f(x) = 0. \tag{B.46}$$

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

Appendix C

Some formulas and inequalities for derivatives and integrals of modified Bessel functions

In this appendix we establish some formulas and inequalities for derivatives and integrals of modified Bessel functions. In Appendix D we shall apply these formulas and inequalities to obtain uniform bounds for the expressions involving modified Bessel functions that appear in the statement of Lemma 3.17 and in the proof of Lemma 3.19.

C.1 Derivative formulas for modified Bessel functions

In this section we obtain formulas for the n-th derivatives of $x^{-\nu}I_{\nu}(x)$ and $x^{-\nu}K_{\nu}(x)$. From these formulas we will deduce simple inequalities for the derivatives of these functions. The formulas will be particularly useful in Section D.3, in which we bound the 'singular' terms that arise in the bounds of for the second, third and fourth order derivatives of the solution of the Variance-Gamma Stein equation, which are given in Lemma 3.17.

Before stating these formulas, we establish a result for the coefficients that are present in the formulas. The coefficients $A_k^n(\nu)$ and $B_k^n(\nu)$ are defined, for $n \in \mathbb{N}$ and all real numbers ν, except the integers $-(k+1), -(k+2), \ldots, -(2k-1), -(2k+1), -(2k+2), \ldots, -(k+n-1)$, and $-(k+2), -(k+3), \ldots, -2k, -(2k+2), -(2k+3), \ldots, -(k+n)$, respectively, as follows:

$$A_k^n(\nu) = \frac{(2n)!\,(\nu+2k)\prod_{j=0}^{k-1}(2\nu+2j+1)}{2^{2n-k}(2k)!\,(n-k)!\prod_{j=0}^{n}(\nu+k+j)}, \qquad k = 0, 1, \ldots n, \tag{C.1}$$

$$B_k^n(\nu) = \frac{(2n+1)!\,(\nu+2k+1)\prod_{j=0}^{k-1}(2\nu+2j+1)}{2^{2n-k}(2k+1)!\,(n-k)!\prod_{j=0}^{n}(\nu+k+j+1)}, \qquad k = 0, 1, \ldots n, \tag{C.2}$$

where we set $\prod_{j=0}^{-1}(2\nu+2j+1)=1$.

Remark C.1. The coefficients $A_k^n(\nu)$ and $B_k^n(\nu)$ are equal to zero if and only if $k\geq 1$ and $\nu=-\frac{1}{2}-l$, where $l=0,1,\ldots,k-1$, and for $\nu>-1/2$ satisfy $0<A_k^n(\nu),B_k^n(\nu)\leq 1$.

The following lemma gives some properties of the coefficients $A_k^n(\nu)$ and $B_k^n(\nu)$ that will be used in the proof of the main result of this section.

Lemma C.2. *Let $n\in\mathbb{N}$, then the coefficients $A_k^n(\nu)$ and $B_k^n(\nu)$ are related as follows*

$$\begin{aligned}
B_k^n(\nu) &= \frac{\nu+k}{\nu+2k}A_k^n(\nu)+\frac{k+1}{\nu+2k+2}A_{k+1}^n(\nu), \quad 0\leq k\leq n-1, && \text{(C.3)}\\
B_n^n(\nu) &= \frac{\nu+n}{\nu+2n}A_n^n(\nu), && \text{(C.4)}\\
A_0^{n+1}(\nu) &= \frac{1}{2(\nu+1)}B_0^n(\nu), && \text{(C.5)}\\
A_{k+1}^{n+1}(\nu) &= \frac{2\nu+2k+1}{2(\nu+2k+1)}B_k^n(\nu)+\frac{2k+3}{2(\nu+2k+3)}B_{k+1}^n(\nu), \quad 0\leq k\leq n-1, && \text{(C.6)}\\
A_{n+1}^{n+1}(\nu) &= \frac{2\nu+2n+1}{2(\nu+2n+1)}B_n^n(\nu), && \text{(C.7)}
\end{aligned}$$

and satisfy

$$\sum_{k=0}^{n}A_k^n(\nu)=\sum_{k=0}^{n}B_k^n(\nu)=1. \tag{C.8}$$

Proof. Identities (C.3)–(C.7) can easily be verified by simply substituting the definitions of $A_{2k}^{2n}(\nu)$ and $B_{2k+1}^{2n+1}(\nu)$, which are given by (C.1) and (C.2), into both sides of the identities.

We now prove that identity (C.8) holds. From identities (C.5), (C.6) and (C.7) we have that

$$\begin{aligned}
\sum_{k=0}^{n+1}A_k^{n+1}(\nu) = \frac{1}{2(\nu+1)}B_0^n(\nu)+\sum_{k=0}^{n-1}\bigg\{\frac{2\nu+2k+1}{2(\nu+2k+1)}B_k^n(\nu)\\
+\frac{2k+3}{2(\nu+2k+3)}B_{k+1}^n(\nu)\bigg\}+\frac{2\nu+2n+1}{2(\nu+2n+1)}B_n^n(\nu).
\end{aligned}$$

Setting $l=k+1$ gives

$$\sum_{k=0}^{n+1}A_k^{n+1}(\nu)=\sum_{k=0}^{n}\frac{2\nu+2k+1}{2(\nu+2k+1)}B_k^n(\nu)+\sum_{l=0}^{n}\frac{2l+1}{2(\nu+2l+1)}B_0^n(\nu)=\sum_{k=0}^{n}B_k^n(\nu).$$

A similar calculation shows that

$$\sum_{k=0}^{n}B_k^n(\nu)=\sum_{k=0}^{n}A_k^n(\nu).$$

Since $A_0^0(\nu)=B_0^0(\nu)=1$ the result follows. $\square$

Before stating our formulas for the n-th derivatives of $x^{-\nu}I_\nu(x)$ and $x^{-\nu}K_\nu(x)$, we introduce some notation. For $N\geq 1$, we write $-[N]$ for the set $\{-1,-2,\ldots,-N\}$. We take $-[0]$ and $-[-1]$ to be the empty set. With this notation we have:

Proposition C.3. *Suppose $n\in\mathbb{N}$. Then,*

$$\begin{aligned}
\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{I_\nu(x)}{x^\nu}\right) &= \sum_{k=0}^{n}A_k^n(\nu)\frac{I_{\nu+2k}(x)}{x^\nu}, \qquad \nu\in\mathbb{R}\setminus(-[2n-1]),\\
\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{K_\nu(x)}{x^\nu}\right) &= \sum_{k=0}^{n}A_k^n(\nu)\frac{K_{\nu+2k}(x)}{x^\nu}, \qquad \nu\in\mathbb{R}\setminus(-[2n-1]), \qquad (C.9)\\
\frac{\mathrm{d}^{2n+1}}{\mathrm{d}x^{2n+1}}\left(\frac{I_\nu(x)}{x^\nu}\right) &= \sum_{k=0}^{n}B_k^n(\nu)\frac{I_{\nu+2k+1}(x)}{x^\nu}, \qquad \nu\in\mathbb{R}\setminus(-[2n]),\\
\frac{\mathrm{d}^{2n+1}}{\mathrm{d}x^{2n+1}}\left(\frac{K_\nu(x)}{x^\nu}\right) &= -\sum_{k=0}^{n}B_k^n(\nu)\frac{K_{\nu+2k+1}(x)}{x^\nu}, \qquad \nu\in\mathbb{R}\setminus(-[2n]). \qquad (C.10)
\end{aligned}$$

Proof. We begin by introducing some notation. Let $\mathscr{L}_\nu(x)$ denote $I_\nu(x)$, $\mathrm{e}^{\nu\pi i}K_\nu(x)$ or any linear combination of these functions, in which the coefficients are independent of ν and x. It therefore suffices to prove

$$\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right) = \sum_{k=0}^{n}A_k^n(\nu)\frac{\mathscr{L}_{\nu+2k}(x)}{x^\nu}, \qquad \nu\in\mathbb{R}\setminus(-[2n-1]), \qquad (C.11)$$

$$\frac{\mathrm{d}^{2n+1}}{\mathrm{d}x^{2n+1}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right) = \sum_{k=0}^{n}B_k^n(\nu)\frac{\mathscr{L}_{\nu+2k+1}(x)}{x^\nu}, \qquad \nu\in\mathbb{R}\setminus(-[2n]). \qquad (C.12)$$

Now, from (B.19) and (B.20) we have

$$\mathscr{L}_{\nu+1}(x)=\mathscr{L}_{\nu-1}(x)-\frac{2\nu}{x}\mathscr{L}_\nu(x), \qquad (C.13)$$

and from (B.31) and (B.32) we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right)=\frac{\mathscr{L}_{\nu+1}(x)}{x^\nu}. \qquad (C.14)$$

We prove formulas (C.11) and (C.12) by induction on n. It is true that (C.11) holds for $n=0$ and (C.12) holds for $n=0$ by (C.14). Suppose now that (C.12) holds for $n=m$, where $m\geq 0$.

We therefore have

$$\frac{\mathrm{d}^{2m+1}}{\mathrm{d}x^{2m+1}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right)=\sum_{k=0}^{m}B_k^m(\nu)\frac{\mathscr{L}_{\nu+2k+1}(x)}{x^\nu},\qquad \nu\in\mathbb{R}\setminus(-[2m]).$$

We may differentiate using (C.14) to obtain

$$\begin{aligned}\frac{\mathrm{d}^{2m+2}}{\mathrm{d}x^{2m+2}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right)&=\sum_{k=0}^{m}B_k^m(\nu)\frac{\mathrm{d}}{\mathrm{d}x}\left(x^{2k+1}\cdot\frac{L_{\nu+2k+1}(x)}{x^{\nu+2k+1}}\right)\\&=\sum_{k=0}^{m}B_k^m(\nu)\left(\frac{\mathscr{L}_{\nu+2k+2}(x)}{x^\nu}+(2k+1)\frac{\mathscr{L}_{\nu+2k+1}(x)}{x^{\nu+1}}\right),\quad \nu\in\mathbb{R}\setminus(-[2m+1]).\end{aligned}\tag{C.15}$$

Applying (C.13) gives

$$(2k+1)\frac{\mathscr{L}_{\nu+2k+1}(x)}{x^{\nu+1}}=\frac{2k+1}{2(\nu+2k+1)}\left(\frac{\mathscr{L}_{\nu+2k}(x)}{x^\nu}-\frac{\mathscr{L}_{\nu+2k+2}(x)}{x^\nu}\right).$$

Substituting into (C.15) gives

$$\begin{aligned}\frac{\mathrm{d}^{2m+2}}{\mathrm{d}x^{2m+2}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right)&=\sum_{k=0}^{m}B_k^m(\nu)\left(\frac{2\nu+2k+1}{2(\nu+2k+1)}\frac{L_{\nu+2k+2}(x)}{x^\nu}+\frac{2k+1}{2(\nu+2k+1)}\frac{\mathscr{L}_{\nu+2k}(x)}{x^\nu}\right)\\&=\sum_{k=0}^{m+1}\tilde{A}_k^{m+1}(\nu)\frac{\mathscr{L}_{\nu+2k}(x)}{x^\nu},\qquad \nu\in\mathbb{R}\setminus(-[2m+1]),\end{aligned}$$

where

$$\tilde{A}_{m+1}^{m+1}(\nu)=\frac{2\nu+2m+1}{2(\nu+2m+1)}B_m^m(\nu),\qquad \tilde{A}_0^{m+1}(\nu)=\frac{1}{2(\nu+1)}B_1^m(\nu),$$

$$\tilde{A}_{k+1}^{m+1}(\nu)=\frac{2\nu+2k+1}{2(\nu+2k+1)}B_k^m(\nu)+\frac{2k+3}{2(\nu+2k+3)}B_{k+1}^m(\nu),\quad k=0,1,\dots,m-1.$$

We see from lemma C.2 that $\tilde{A}_k^{m+1}(\nu)=A_k^{m+1}(\nu)$, for all $k=0,1,\dots,m+1$. It therefore follows that if (C.12) holds for $n=m$ then (C.11) holds for $n=m+1$.

We now suppose that (C.11) holds for $n=m$, where $m\geq 1$. If we can show that it then follows that (C.12) holds for $n=m$ then the proof will be complete. Our inductive hypothesis is therefore that

$$\frac{\mathrm{d}^{2m}}{\mathrm{d}x^{2m}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right)=\sum_{k=0}^{m}A_k^m(\nu)\frac{\mathscr{L}_{\nu+2k}(x)}{x^\nu},\qquad \nu\in\mathbb{R}\setminus(-[2m-1]).$$

Differentiating and applying a similar argument to the first part gives

$$\frac{\mathrm{d}^{2m+1}}{\mathrm{d}x^{2m+1}}\left(\frac{\mathscr{L}_\nu(x)}{x^\nu}\right)=\sum_{k=0}^{m}\tilde{B}_k^m(\nu)\frac{\mathscr{L}_{\nu+2k+1}(x)}{x^\nu},\qquad \nu\in\mathbb{R}\setminus(-[2m]),$$

where

$$\tilde{B}_m^m(\nu)=\frac{\nu+m}{\nu+2m}A_m^m(\nu),\quad \tilde{B}_k^m(\nu)=\frac{\nu+k}{\nu+2k}A_k^m(\nu)+\frac{k+1}{\nu+2k+2}A_{k+1}^m(\nu),$$

and $k=0,1,\dots,m-1$. We see from lemma C.2 that $\tilde{B}_k^m(\nu)=B_k^m(\nu)$, for all $k=0,1,\dots,m$. It therefore follows that if (C.11) holds for $n=m$ then (C.12) holds for $n=m$, which completes the proof. □

In Appendix D we will use the following special cases of the formulas of Proposition C.3.

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}(K_0(x)) = \frac{1}{2}(K_2(x)+K_0(x)), \tag{C.16}$$

$$\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right) = -\frac{2\nu+1}{2(\nu+2)}\frac{K_{\nu+3}(x)}{x^\nu}-\frac{3}{2(\nu+2)}\frac{K_{\nu+1}(x)}{x^\nu}, \tag{C.17}$$

$$\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right) = \frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)}\frac{K_{\nu+4}(x)}{x^\nu}+\frac{3(2\nu+1)}{2(\nu+1)(\nu+3)}\frac{K_{\nu+2}(x)}{x^\nu}+\frac{3}{4(\nu+1)(\nu+2)}\frac{K_\nu(x)}{x^\nu}. \tag{C.18}$$

Proposition C.3 has the following simple corollary. The inequalities given in this corollary will be used throughout Appendix D.

Corollary C.4. *Let $n\in\mathbb{N}$ and $\nu\geq-1/2$, then for all $x>0$ the following inequalities hold*

$$0<\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{K_\nu(x)}{x^\nu}\right)\leq\frac{K_{\nu+2n}(x)}{x^\nu},\quad -\frac{K_{\nu+2n+1}(x)}{x^\nu}\leq\frac{\mathrm{d}^{2n+1}}{\mathrm{d}x^{2n+1}}\left(\frac{K_\nu(x)}{x^\nu}\right)<0, \tag{C.19}$$

and

$$0<\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{I_\nu(x)}{x^\nu}\right)\leq\begin{cases}x^{-\nu}I_{\nu+1}(x), & \text{odd } n,\\ x^{-\nu}I_\nu(x), & \text{even } n.\end{cases} \tag{C.20}$$

Proof. Apply the formulas for the n-th derivative of $x^{-\nu}K_\nu(x)$ and $x^{-\nu}I_\nu(x)$, that are given in Theorem C.3, the fact the coefficients satisfy $\sum_{k=0}^n A_{2k}^{2n}(\nu)=\sum_{k=0}^n B_{2k+1}^{2n+1}(\nu)=1$, and the inequalities $I_\nu(x)\leq I_{\nu+1}(x)$ and $K_\nu(x)\leq K_{\nu+1}(x)$ for $\nu>-1/2$ and $x>0$. □

C.2 Inequalities for modified Bessel functions and their integrals

In this section we present a number of simple inequalities for integrals involving modified Bessel functions, as well as an inequality for an expression involving the modified Bessel function $K_\nu(x)$.

Proposition C.5. *Let $-1<\beta<1$ and $0\leq\gamma<1$, then the following inequalities hold for all $x>0$*

$$\begin{aligned}
\int_0^x t^\nu I_\nu(t)\,\mathrm{d}t &> x^\nu I_{\nu+1}(x), \quad \nu>-1, \\
\int_0^x t^\nu I_\nu(t)\,\mathrm{d}t &< x^\nu I_{\nu}(x), \quad \nu\geq 1/2, && \text{(C.21)} \\
I_{(\nu,0,n+1)}(x) &< I_{(\nu,0,n)}(x), \quad \nu\geq 1/2, && \text{(C.22)} \\
I_{(\nu,-\gamma,n)}(x) &< \frac{1}{(1-\gamma)^n}\mathrm{e}^{-\gamma x}I_{(\nu,0,n)}(x), \quad \nu\geq 1/2,\ n\in\mathbb{N}, && \text{(C.23)} \\
\int_0^x t^\nu I_{\nu+n}(t)\,\mathrm{d}t &< \frac{2(\nu+n+1)}{2\nu+n+1}x^\nu I_{\nu+n+1}(x), \quad \nu>-1/2,\ n\geq 0, && \text{(C.24)} \\
I_{(\nu,0,n)}(x) &< \bigg\{\prod_{k=1}^n\frac{2\nu+2k}{2\nu+k}\bigg\}x^\nu I_{\nu+n}(x), \quad \nu\geq 0,\ n\in\mathbb{Z}^+, && \text{(C.25)} \\
I_{(\nu,-\gamma,n)}(x) &< \frac{1}{(1-\gamma)^n}\bigg\{\prod_{k=1}^n\frac{2\nu+2k}{2\nu+k}\bigg\}\mathrm{e}^{-\gamma x}x^\nu I_{\nu+n}(x), \quad \nu\geq 1/2,\ n\in\mathbb{Z}^+, \\
I_{(\nu,\beta,n)}(x) &< \frac{1}{(1-|\beta|)^n}\bigg\{\prod_{k=1}^n\frac{2\nu+2k}{2\nu+k}\bigg\}\mathrm{e}^{\beta x}x^\nu I_{\nu+n}(x), \quad \nu\geq 1/2,\ n\in\mathbb{Z}^+.
\end{aligned}$$

Proof. (i) From the differentiation formula (B.29) we have that

$$\int_0^x t^\nu I_\nu(t)\,\mathrm{d}t=\int_0^x\frac{1}{t}t^{\nu+1}I_\nu(t)\,\mathrm{d}t>\frac{1}{x}\int_0^x t^{\nu+1}I_\nu(t)\,\mathrm{d}t=x^\nu I_{\nu+1}(x),$$

since by (B.7) we have $\lim_{x\downarrow 0}x^{\nu+1}I_{\nu+1}(x)=0$ for $\nu>-1$.

(ii) Using inequality (B.13) and then applying (B.29) we get

$$\int_0^x t^\nu I_\nu(t)\,\mathrm{d}t<\int_0^x t^\nu I_{\nu-1}(t)\,\mathrm{d}t=x^\nu I_\nu(x).$$

(iii) From inequality (C.21), we have

$$I_{(\nu,0,1)}(x)<I_{(\nu,0,0)}(x).$$

Integrating both sides of the above display n times with respect to x yields the desired inequality.

(iv) We prove the result by induction on n. The result is trivially true for $n = 0$. Suppose the result is true for $n = k$. From the inductive hypothesis we have

$$I_{(\nu,-\gamma,k+1)}(x) = \int_0^x I_{(\nu,-\gamma,k)}(t)\,\mathrm{d}t < \frac{1}{(1-\gamma)^k}\int_0^x \mathrm{e}^{-\gamma t} I_{(\nu,0,k)}(t)\,\mathrm{d}t. \tag{C.26}$$

Integration by parts and inequality (C.22) gives

$$\begin{aligned}\int_0^x \mathrm{e}^{-\gamma t} I_{(\nu,0,k)}(t)\,\mathrm{d}t &= \mathrm{e}^{-\gamma x} I_{(\nu,0,k+1)}(x) + \gamma \int_0^x \mathrm{e}^{-\gamma t} I_{(\nu,0,k+1)}(t)\,\mathrm{d}t \\ &< \mathrm{e}^{-\gamma x} I_{(\nu,0,k+1)}(x) + \gamma \int_0^x \mathrm{e}^{-\gamma t} I_{(\nu,0,k)}(t)\,\mathrm{d}t.\end{aligned}$$

Rearranging we obtain

$$\int_0^x \mathrm{e}^{-\gamma t} I_{(\nu,0,k)}(t)\,\mathrm{d}t < \frac{1}{1-\gamma}\mathrm{e}^{-\gamma x} I_{(\nu,0,k+1)}(x),$$

and substituting into (C.26) gives

$$I_{(\nu,-\gamma,k+1)}(x) < \frac{1}{(1-\gamma)^{k+1}}\mathrm{e}^{-\gamma x} I_{(\nu,0,k+1)}(x).$$

Hence the result has been proved by induction.

(v) From the differentiation formula (B.29) and identity (B.19) we get that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(t^\nu I_{\nu+n+1}(t)) &= \frac{\mathrm{d}}{\mathrm{d}t}(t^{-(n+1)} \cdot t^{\nu+n+1} I_{\nu+n+1}(t)) \\ &= t^\nu I_{\nu+n}(t) - (n+1)t^{\nu-1} I_{\nu+n+1}(t) \\ &= t^\nu I_{\nu+n}(t) - \frac{n+1}{2(\nu+n+1)} t^\nu I_{\nu+n}(t) + \frac{n+1}{2(\nu+n+1)} t^\nu I_{\nu+n+2}(t) \\ &= \frac{2\nu+n+1}{2(\nu+n+1)} t^\nu I_{\nu+n}(t) + \frac{n+1}{2(\nu+n+1)} t^\nu I_{\nu+n+2}(t).\end{aligned}$$

Integrating both sides over $(0, x)$, applying the fundamental theorem of calculus and rearranging gives

$$\int_0^x t^\nu I_{\nu+n}(t)\,\mathrm{d}t = \frac{2(\nu+n+1)}{2\nu+n+1} x^\nu I_{\nu+n+1}(x) - \frac{n+1}{2\nu+n+1}\int_0^x t^\nu I_{\nu+n+2}(t)\,\mathrm{d}t.$$

The result now follows from the fact that $I_\nu(x) > 0$ for $x > 0$ and by the positivity of the integral.

(vi) From inequality (C.24) we have

$$I_{(\nu,0,1)}(x) = \int_0^x t^\nu I_\nu(t)\,\mathrm{d}t < \frac{2(\nu+1)}{2\nu+1} x^\nu I_{\nu+1}(x),$$

and

$$I_{(\nu,0,2)}(x) = \int_0^x I_{(\nu,0,1)}(t)\,\mathrm{d}t < \frac{2(\nu+1)}{2\nu+1}\int_0^x t^\nu I_{\nu+1}(t)\,\mathrm{d}t < \frac{2(\nu+1)}{2\nu+1}\frac{2(\nu+2)}{2\nu+2} x^\nu I_{\nu+2}(x).$$

Iterating gives the result.

(vii) This follows from inequalities (C.23) and (C.25).

(viii) We have proved the result for $-1 < \beta \leq 0$ in part (vii), so suppose now that $0 < \beta < 1$. Since $\mathrm{e}^{\beta t}$ is an increasing function of t we have $I_{(\nu,\beta,n)}(x) < \mathrm{e}^{\beta x} I_{(\nu,0,n)}(x)$ and as $1 - |\beta| < 1$ the result now follows from part (vi). □

We now establish a simple lemma, which gives a monotonicity result for the ratio $\frac{K_{\nu-1}(x)}{K_\nu(x)}$. The lemma has an immediate corollary, which we will make use of in the proof of Proposition C.9.

Lemma C.6. *Suppose $x > 0$, then the function $\frac{K_{\nu-1}(x)}{K_\nu(x)}$ is strictly monotone increasing for $\nu > 1/2$, is constant for $\nu = 1/2$, and is strictly monotone decreasing for $\nu < 1/2$.*

Proof. To simplify the calculations, we let $\mu = \nu + 1/2$ and define $h_\mu(x) = \frac{K_{\mu-1/2}(x)}{K_{\mu+1/2}(x)}$. It follows from the quotient rule and the differentiation formulas (B.27) and (B.28) that

$$h_\mu'(x) = -1 + \frac{2\mu}{x} h_\mu(x) + h_\mu^2(x). \tag{C.27}$$

Now, Theorem 2 of Segura [67] states that

$$h_\mu'(x) > \frac{x}{\sqrt{x^2+\mu^2}+\mu} = \frac{-\mu + \sqrt{x^2+\mu^2}}{x}, \qquad \mu > 0,\ x > 0, \tag{C.28}$$

and that for $\mu < 0$ the inequality is reversed and for $\mu = 0$ equality holds. Applying inequality (C.28) to equation (C.27) gives, for $\mu > 0$,

$$h_\mu'(x) > -1 + \frac{2\mu}{x}\left(\frac{-\mu+\sqrt{x^2+\mu^2}}{x}\right) + \left(\frac{-\mu+\sqrt{x^2+\mu^2}}{x}\right)^2 = 0$$

Similarly, we see that $h_0'(x) = 0$ and that $h_\mu'(x) < 0$ for $\mu < 0$. This completes the proof. □

Corollary C.7. *For $\nu > 1/2$ and $\alpha > 1$ the equation $K_\nu(x) = \alpha K_{\nu-1}(x)$ has one root in the region $x > 0$.*

Proof. From the asymptotic formulas (B.8) and (B.11), it follows that for $\nu > 1/2$,

$$\lim_{x\downarrow 0}\frac{K_{\nu-1}(x)}{K_\nu(x)}=0, \qquad \text{and} \qquad \lim_{x\to\infty}\frac{K_{\nu-1}(x)}{K_\nu(x)}=1.$$

Since $\frac{K_{\nu-1}(x)}{K_\nu(x)}$ is strictly monotone decreasing on $(0,\infty)$, it follows that for $\alpha>1$ the equation $K_\nu(x)=\alpha K_{\nu-1}(x)$ (i.e. $\frac{K_{\nu-1}(x)}{K_\nu(x)}=\frac{1}{\alpha}$) has one root in the region $x>0$. □

As an aside, we note that Lemma C.7 allows us to easily establish an inequality for the Turánian $\Delta_\nu(x)=K_\nu^2(x)-K_{\nu-1}(x)K_{\nu+1}(x)$ (for more details on the Turánian $\Delta_\nu(x)$ see Baricz [9]).

Proposition C.8. *Suppose $x>0$, then $\Delta_\nu(x)<\Delta_{\nu-1}(x)$ for $\nu>1/2$, $\Delta_{1/2}(x)=\Delta_{-1/2}(x)$, and $\Delta_\nu(x)>\Delta_{\nu-1}(x)$ for $\nu<1/2$.*

Proof. By the quotient rule and differentiation formula (B.26), we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{K_{\nu-1}(x)}{K_\nu(x)}\right)&=-\frac{(K_\nu(x)+K_{\nu-2}(x))K_\nu(x)-(K_{\nu+1}(x)+K_{\nu-1}(x))K_{\nu-1}(x)}{2K_\nu^2(x)}\\&=\frac{K_{\nu-1}^2(x)-K_{\nu-2}(x)K_\nu(x)-(K_\nu^2(x)-K_{\nu-1}(x)K_{\nu+1}(x))}{2K_\nu^2(x)}\\&=\frac{\Delta_{\nu-1}(x)-\Delta_\nu(x)}{2K_\nu^2(x)}.\end{aligned}$$

Since, by Lemma C.7, the function $\frac{K_{\nu-1}(x)}{K_\nu(x)}$ is strictly monotone increasing for $\nu>1/2$, is constant for $\nu=1/2$, and is strictly monotone decreasing for $\nu<1/2$, the result follows. □

With the aid of Corollary C.7 and standard properties of the modified Bessel function $K_\nu(x)$, we can prove at the following proposition.

Proposition C.9. *Let $-1<\beta<1$, then for all $x>0$ the following inequalities hold*

$$\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \;<\; x^\nu K_{\nu+1}(x), \quad \nu\in\mathbb{R}, \tag{C.29}$$

$$\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \;<\; x^\nu K_\nu(x), \quad \nu<1/2, \tag{C.30}$$

$$\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t \;<\; \frac{1}{1-|\beta|}\mathrm{e}^{\beta x}x^\nu K_\nu(x), \quad \nu<1/2, \tag{C.31}$$

$$\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \;\leq\; \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)}x^\nu K_\nu(x), \quad \nu\geq 1/2, \tag{C.32}$$

$$\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t \;\leq\; \frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)}\mathrm{e}^{\beta x}x^\nu K_\nu(x), \quad \nu\geq 1/2. \tag{C.33}$$

Proof. (i) From the differentiation formula (B.30) we have that

$$\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t = \int_x^\infty \frac{1}{t}t^{\nu+1}K_\nu(t)\,\mathrm{d}t < \frac{1}{x}\int_x^\infty t^{\nu+1}K_\nu(t)\,\mathrm{d}t = x^\nu K_{\nu+1}(x),$$

since, by the asymptotic formula (B.11), $\lim_{x\to\infty} x^{\nu+1}K_{\nu+1}(x) = 0$.

(ii) Using inequality (B.15) and then apply the differentiation formula (B.30) we have

$$\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t < \int_x^\infty t^\nu K_{\nu-1}(t)\,\mathrm{d}t = x^\nu K_\nu(x).$$

(iii) Now suppose that $\nu < 1/2$ and $\beta > 0$. Using integration by parts and the differentiation formula (B.30) gives

$$\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t = -\frac{1}{\beta}\mathrm{e}^{\beta x}x^\nu K_\nu(x) + \frac{1}{\beta}\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_{\nu-1}(t)\,\mathrm{d}t.$$

Applying the inequality (B.15) and rearranging gives

$$\left(\frac{1}{\beta}-1\right)\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t < \frac{1}{\beta}\mathrm{e}^{\beta x}x^\nu K_\nu(x).$$

Inequality (C.31) for $\beta > 0$ now follows on rearranging.

The case $\beta \leq 0$ is simple. Since $\mathrm{e}^{\beta t}$ is a non increasing function of t when $\beta \leq 0$ we have

$$\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t \leq \mathrm{e}^{\beta x}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t < \mathrm{e}^{\beta x}x^\nu K_\nu(x) \leq \frac{1}{1-|\beta|}\mathrm{e}^{\beta x}x^\nu K_\nu(x),$$

where we used inequality (C.30) to obtain the the second inequality. Hence inequality (C.31) has been proved.

(iv) The case $\nu = 1/2$ is simple. Using (B.6) we may easily integrate $t^{1/2}K_{1/2}(t)$:

$$\int_x^\infty t^{1/2}K_{1/2}(t)\,\mathrm{d}t = \int_x^\infty \sqrt{\frac{\pi}{2}}\mathrm{e}^{-t}\,\mathrm{d}t = \sqrt{\frac{\pi}{2}}\mathrm{e}^{-x} = x^{1/2}K_{1/2}(x).$$

It therefore follows that inequality (C.32) holds for $\nu = 1/2$ because we have

$$\frac{\sqrt{\pi}\Gamma(1)}{\Gamma(1/2)} = 1,$$

where we used the facts that $\Gamma(1) = 1$ and $\Gamma(1/2) = \sqrt{\pi}$.

Now suppose $\nu > 1/2$. We begin by defining the function $u(x)$ to be

$$u(x) = Mx^{\nu}K_{\nu}(x) - \int_x^{\infty} t^{\nu}K_{\nu}(t)\,\mathrm{d}t,$$

where

$$M = \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)}.$$

We now show that $u(x) \geq 0$ for all $x \geq 0$, which will prove the result. We begin by noting that $\lim_{x\rightarrow 0^+} u(x) = 0$ and $\lim_{x\rightarrow\infty} u(x) = 0$, which are verified by the following calculations, where we make use of the asymptotic formula (B.8) and the definite integral formula (B.43).

$$\begin{aligned} u(0) &= \lim_{x\rightarrow 0^+} \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)} x^{\nu}K_{\nu}(x) - \int_0^{\infty} t^{\nu}K_{\nu}(t)\,\mathrm{d}t \\ &= \sqrt{\pi}\Gamma(\nu+1/2)2^{\nu-1} - \sqrt{\pi}\Gamma(\nu+1/2)2^{\nu-1} \\ &= 0, \end{aligned}$$

and

$$\lim_{x\rightarrow\infty} u(x) = \lim_{x\rightarrow\infty} Mx^{\nu}K_{\nu}(x) - \lim_{x\rightarrow\infty}\int_x^{\infty} t^{\nu}K_{\nu}(t)\,\mathrm{d}t = 0,$$

where we used the asymptotic formula (B.11) to obtain the above equality. We may obtain an expression for the first derivative of $u(x)$ by the use of the differentiation formula (B.30) as follows

$$u'(x) = x^{\nu}[K_{\nu}(x) - MK_{\nu-1}(x)]. \tag{C.34}$$

In the limit $x \rightarrow 0^+$ we have, by the asymptotic formula (B.8), that

$$u'(x) \sim \begin{cases} x^{\nu}\left\{2^{\nu-1}\Gamma(\nu)\frac{1}{x^{\nu}} - M2^{|\nu-1|-1}\Gamma(|\nu-1|)\frac{1}{x^{|\nu-1|}}\right\}, & \nu \neq 1, \\ x^{\nu}\left\{2^{\nu-1}\Gamma(\nu)\frac{1}{x^{\nu}} + M\log x\right\}, & \nu = 1. \end{cases}$$

Since $\nu > |\nu - 1|$ for $\nu > 1/2$ and $\lim_{x\rightarrow 0^+} x^a \log x = 0$, where $a > 0$, we have

$$u'(x) \sim 2^{\nu-1}\Gamma(\nu), \quad \text{as } x \rightarrow 0^+, \quad \text{for } \nu > 1/2.$$

Therefore $u(x)$ is initially an increasing function of x. In the limit $x \rightarrow \infty$ we have, by (B.11),

$$u'(x) \sim \left(1 - \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)}\right)\sqrt{\frac{\pi}{2}}x^{\nu-1/2}\mathrm{e}^{-x} < 0, \quad \text{for } \nu > 1/2.$$

We therefore see that $u(x)$ is an decreasing function of x for large, positive x. From the formula

(C.34) we see that x^* is a turning point of $u(x)$ if and only if

$$K_\nu(x^*) = \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)} K_{\nu-1}(x^*). \tag{C.35}$$

From Corollary C.7, it follows that equation (C.35) has one root for $\nu > 1/2$ (for which $\frac{\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)} > 1$).

Putting these results together, we see that $u(x)$ is non-negative at the origin and initially increases until it reaches it maximum value at x^*, it then decreases and tends to 0 as $x \to \infty$. Therefore $u(x)$ is non-negative for all $x \geq 0$ when $\nu > 1/2$.

(v) The proof for $\beta \leq 0$ is easy and follows immediately from part (iv), since $1 < \frac{2}{(1-\beta^2)^{\nu+1/2}}$ for $\nu \geq 1/2$. So we suppose $\beta > 0$. Again, because $K_{1/2}(x) = \sqrt{\frac{\pi}{2x}}\mathrm{e}^{-x}$, the case $\nu = 1/2$ is straightforward, so we also suppose $\nu > 1/2$. We make use of a similar argument to the one used in the proof of part (iv). We define the function $v(x)$ to be

$$v(x) = N\mathrm{e}^{\beta x}x^\nu K_\nu(x) - \int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t,$$

where

$$N = \frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)}.$$

We now show that $v(x) \geq 0$ for all $x \geq 0$, which will prove the result. We begin by noting that $\lim_{x\to 0^+} v(x) > 0$ and $\lim_{x\to\infty} v(x) = 0$, which are verified by the following calculations, where we make use of the asymptotic formula (B.8) and the definite integral formula (B.43).

$$\begin{aligned}
v(0) &= \lim_{x\to 0^+} \frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)} x^\nu K_\nu(x) - \int_0^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t \\
&= \frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)} \cdot 2^{\nu-1}\Gamma(\nu) - \int_0^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t \\
&> \frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)} \cdot 2^{\nu-1}\Gamma(\nu) - \int_{-\infty}^\infty \mathrm{e}^{\beta t}|t|^\nu K_\nu(|t|)\,\mathrm{d}t \\
&= \frac{\sqrt{\pi}\Gamma(\nu+1/2)2^\nu}{(1-\beta^2)^{\nu+1/2}} - \frac{\sqrt{\pi}\Gamma(\nu+1/2)2^\nu}{(1-\beta^2)^{\nu+1/2}} \\
&= 0,
\end{aligned}$$

and

$$\lim_{x\to\infty} v(x) = \lim_{x\to\infty} N\mathrm{e}^{\beta x}x^\nu K_\nu(x) - \lim_{x\to\infty}\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t = 0,$$

where we used the asymptotic formula (B.11) to obtain the above equality. We may obtain an expression for the first derivative of $v(x)$ by the use of the differentiation formula (B.30) as follows

$$v'(x) = \mathrm{e}^{\beta x} x^{\nu}[(1+N\beta)K_{\nu}(x) - NK_{\nu-1}(x)]. \tag{C.36}$$

In the limit $x \to 0^+$ we have, by the asymptotic formula (B.8), that

$$v'(x) \sim \begin{cases} \mathrm{e}^{\beta x} x^{\nu}\left\{2^{\nu-1}\Gamma(\nu)(1+N\beta)\frac{1}{x^{\nu}} - N\cdot 2^{|\nu-1|-1}\Gamma(|\nu-1|)\frac{1}{x^{|\nu-1|}}\right\}, & \nu \neq 1, \\ \mathrm{e}^{\beta x} x^{\nu}\left\{2^{\nu-1}\Gamma(\nu)(1+N\beta)\frac{1}{x^{\nu}} + N\log x\right\}, & \nu = 1. \end{cases}$$

As in part (iv), we see that $v(x)$ is initially an increasing function of x. In the limit $x \to \infty$ we have

$$v'(x) \sim (1 - N(1-\beta))\sqrt{\frac{\pi}{2}} x^{\nu-1/2}\mathrm{e}^{(\beta-1)x}, \quad \text{for } \nu > 1/2.$$

Now, for $\nu > 1/2$ and $0 < \beta < 1$ we have, by (B.11),

$$N(1-\beta) = \frac{2\sqrt{\pi}\Gamma(\nu+1/2)}{\Gamma(\nu)}\cdot\frac{1}{(1-\beta^2)^{\nu-1/2}}\cdot\frac{1}{1+\beta} > 2\cdot 1\cdot\frac{1}{2} = 1. \tag{C.37}$$

Hence, $v(x)$ is an decreasing function of x for large, positive x. From formula (C.36) we see that x^* is a turning point of $v(x)$ if and only if

$$(1+N\beta)K_{\nu}(x^*) = NK_{\nu-1}(x^*). \tag{C.38}$$

Inequality (C.37) shows that $N > 1 + N\beta$ for all $\nu > 1/2$ and $0 < \beta < 1$. From Corollary C.7, it follows that equation (C.38) has one root for positive x and therefore $v(x)$ has one maximum which occurs at positive x. Putting these results together we see that $v(x)$ is positive at the origin and initially increases until it reaches it maximum value at x^*, it then decreases and tends to 0 as $x \to \infty$. Therefore $v(x)$ is non-negative for all $x \geq 0$ when $\nu > 1/2$, which completes the proof. □

When end this section by obtaining an inequality for the modified Bessel function $K_{\nu}(x)$ that is used in our proof of Lemma D.15.

Proposition C.10. *Let $\mu > 1$ and $x \geq 0$, then*

$$\frac{1}{x^2} - \frac{x^{\mu-2}K_{\mu}(x)}{2^{\mu-1}\Gamma(\mu)}, \tag{C.39}$$

is a monotone decreasing function of x on $(0,\infty)$ and satisfies the inequality

$$0 < \frac{1}{x^2} - \frac{x^{\mu-2}K_{\mu}(x)}{2^{\mu-1}\Gamma(\mu)} \leq \frac{1}{4(\mu-1)}, \quad \textit{for all } x \geq 0.$$

Proof. Applying the differentiation formula (B.26) gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x^2}-\frac{x^{\mu-2}K_\mu(x)}{2^{\mu-1}\Gamma(\mu)}\right)=-\frac{2}{x^3}-\frac{(\mu-2)x^{\mu-3}K_\mu(x)-\frac{1}{2}(K_{\mu-1}(x)+K_{\mu+1}(x))x^{\mu-2}}{2^{\mu-1}\Gamma(\mu)}. \tag{C.40}$$

Using (B.20) we may simplify the numerator as follows

$$\begin{aligned}(\mu-2)K_\mu(x)-\frac{1}{2}x(K_{\mu-1}(x)+K_{\mu+1}(x))&=(\mu-2)K_\mu(x)-\frac{1}{2}x\left(2K_{\mu-1}(x)+\frac{2\mu}{x}K_\mu(x)\right)\\&=-xK_{\mu-1}(x)-2K_\mu(x).\end{aligned}$$

Hence, (C.40) simplifies to

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{x^2}-\frac{x^{\mu-2}K_\mu(x)}{2^{\mu-1}\Gamma(\mu)}\right)=\frac{-2^\mu\Gamma(\mu)+x^{\mu+1}K_{\mu-1}(x)+2x^\mu K_\mu(x)}{2^{\mu-1}\Gamma(\mu)x^3}.$$

Thus, proving that (C.39) is monotone decreasing reduces to proving that, for $x>0$,

$$x^{\mu+1}K_{\mu-1}(x)+2x^\mu K_\mu(x)<2^\mu\Gamma(\mu). \tag{C.41}$$

From (B.30) we get that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(x^{\mu+1}K_{\mu-1}(x)+2x^\mu K_\mu(x)\right)&=\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\cdot x^{\mu-1}K_{\mu-1}(x)+2x^\mu K_\mu(x)\right)\\&=2x^\mu K_{\mu-1}(x)-x^{\mu+1}K_{\mu-2}(x)-2x^\mu K_{\mu-1}(x)\\&=-x^{\mu+1}K_{\mu-2}(x)\\&<0.\end{aligned}$$

So $x^{\mu+1}K_{\mu-1}(x)+2x^\mu K_\mu(x)$ is a strictly monotone decreasing function of x and from the asymptotic formula (B.8) we see that its limit as $x\to0^+$ is

$$\lim_{x\to0^+}(x^{\mu+1}K_{\mu-1}(x)+2x^\mu K_\mu(x))=2\cdot2^{\mu-1}\Gamma(\mu)=2^\mu\Gamma(\mu).$$

Therefore (C.41) is proved, and so (C.39) is monotone decreasing on $(0,\infty)$. It is therefore bounded above and below its values in the limits $x\to\infty$ and $x\to0$. These are calculated using the asymptotic formulas (B.11) and (B.9) and are given below:

$$\begin{aligned}\lim_{x\to\infty}\left(\frac{1}{x^2}-\frac{x^{\mu-2}K_\mu(x)}{2^{\mu-1}\Gamma(\mu)}\right)&=0,\\ \lim_{x\to0^+}\left(\frac{1}{x^2}-\frac{x^{\mu-2}K_\mu(x)}{2^{\mu-1}\Gamma(\mu)}\right)&=\frac{2^{\mu-3}\Gamma(\mu-1)}{2^{\mu-1}\Gamma(\mu)}=\frac{1}{4(\mu-1)}.\end{aligned}$$

This completes the proof. $\square$

Appendix D

Uniform bounds for expressions involving derivatives and integrals of modified Bessel functions

In this appendix we obtain bounds for a number of expressions involving derivatives and integrals of modified Bessel functions. These bounds enable us to show that, for a bounded test function h, the solution of the $\mathrm{VG}_2(\nu,1,\beta,0)$ Stein equation (3.13) and its first derivative are bounded (Lemma 3.14), and also allow us to obtain uniform bounds for the solution of the $\mathrm{VG}_2(\nu,0,\beta,0)$ and its first four derivatives (Lemma 3.18). A simple change of variables then yields uniform bounds for the solution and its first four derivatives of the $\mathrm{VG}(r,0,\sigma,\mu)$ Stein equation (3.14), which are given in Theorems 3.21 and 3.22.

In Lemma 3.17 we gave bounds for the solution of $\mathrm{VG}_2(\nu,1,\beta,0)$ Stein equation (3.13) and its first four derivatives in terms of expressions involving supremum norms of derivatives and integrals of modified Bessel functions. Recall that there are four types of terms that require bounding:

(i). $\left\|I_{(\nu,\beta,1)}(x)\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right\|$;

(ii). $\left\|I_{(\nu,\beta,n)}(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)\right\|$;

(iii). $\left\|\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta y}y^\nu K_\nu(y)[h(y)-\tilde{\mathrm{VG}}_{\beta,0}^{\nu,1}]\,\mathrm{d}y\right\|$;

(iv). $\left\|I_{(\nu,\beta,k)}(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right)+\text{singular terms}\right\|$,

where k and n are non negative integers and $k<n$. Also, recall that we only required bounds on the terms of type (ii) for $n=1,2,3,4$; bounds on expression of type (iii) for $n=0,1,2,3,4$;

and bounds on the term of type (iv) for $1 \leq k < n \leq 4$.

In this appendix we bound expression (i), the terms of type (ii) for $n = 1$ and for all $\nu > -1/2$ and $-1 < \beta < 1$, and those of type (iii) for $n = 0, 1$ and for all $\nu > -1/2$ and $-1 < \beta < 1$. These bounds enable us to show that solution of the $\mathrm{VG}_2(\nu, 1, \beta, 0)$ Stein equation (3.13) and its first derivative are bounded.

We then specialise to the case $\beta = 0$. We obtain uniform bounds for the expressions of type (ii) for all $n \geq 1$ and $\nu > -1/2$, and uniform bounds for those of type (iii) for all $n \geq 0$ and $\nu > -1/2$. We also obtain uniform bounds for terms of of type (iv), but only in the region $0 \leq x \leq 1$. For the case that 2ν is very close but not equal to an integer these bounds perform very poorly – see Remark D.19. In Lemma 3.18 we combine these bounds to obtain a uniform bound on the solution of the $\mathrm{VG}_2(\nu, 1, 0, 0)$ Stein equation and its first four derivatives in the region $0 \leq x \leq 1$.

Finally, we specialise to the case $\nu = \beta = 0$. In Lemma D.20 we are able to bound the terms of type (iv), in the region $0 \leq x < \infty$ for this special case, and using these bounds and a change of parameters yields uniform bounds for the solution of the $\mathrm{VG}(1, 0, 1, 0)$ Stein equation and its first four derivatives for all $x \in \mathbb{R}$ (recall Remark 3.15). These bounds are stated in Theorem 3.22.

To prove Lemma 3.19 we required bounds on the terms

$$\frac{K_{\nu+n+1}(x)}{x^{\nu}} \int_0^x t^{\nu-k} I_{\nu+n}(t) \, \mathrm{d}t \qquad \text{and} \qquad \frac{I_{\nu+n}(x)}{x^{\nu}} \int_x^{\infty} t^{\nu-k} K_{\nu+n}(t) \, \mathrm{d}t,$$

in the region $x \geq 1$, where $0 \leq k \leq n$. These terms are bounded in Lemma D.8. In Lemma 3.19 we use these bounds to bound the second, third and four derivatives of the solution of the $\mathrm{VG}_2(\nu, 1, 0, 0)$ Stein equation in the region $x \geq 1$. Combining our bounds for the derivatives in the regions $0 \leq x \leq 1$ and $x \geq 1$ leads to bounds for the solution of the $\mathrm{VG}_2(\nu, 1, 0, 0)$ Stein equation and its first four derivatives for all $x \in \mathbb{R}$ (recall Remark 3.15).

Before obtaining these bounds, we state a result, which is proved in Baricz [8], that we will use throughout this appendix. The result has a simple corollary (stated below), which we will also make repeated use of.

Proposition D.1. *For positive real argument and $\nu > -1/2$ the product $K_{\nu}(x)I_{\nu}(x)$ is a monotone strictly decreasing function of x.*

Corollary D.2. *Let $\nu > 0$, then for all $x \geq 0$, then the following inequality holds*

$$0 < K_{\nu}(x)I_{\nu}(x) \leq \frac{1}{2\nu}. \tag{D.1}$$

Proof. The proof is follows from Proposition D.1, the asymptotic expansions (B.7) and (B.8) for modified Bessel functions in the limit x tends to 0 and ∞, and the positivity of modified Bessel functions of positive arguments. □

We now set about bounding the expressions involving modified Bessel functions that are required in the proofs of Lemmas 3.18 and 3.19. We list these bounds as a series of lemmas. For clarity, we list these bounds for the norms in the regions $0 \leq x < \infty$, $1 \leq x < \infty$ and $0 \leq x < 1$ in separate sections. An open problem is to extend our bounds that hold in the region $0 \leq x \leq 1$ to the region $x \geq 0$.

D.1 Bounds for $0 \leq x < \infty$

The first two lemmas are used in the proof of Lemma 3.14, and enable us to show that the solution of the $\mathrm{VG}_2(\nu, 1, \beta, 0)$ Stein equation (3.13) and its first derivative are bounded, provided that the test function h is bounded.

Lemma D.3. *Let $-1 < \beta < 1$ and $\nu > -1/2$, then for all $x \geq 0$ the following inequalities hold*

$$\begin{aligned} \frac{\mathrm{e}^{-\beta x} K_{\nu+1}(x)}{x^\nu} \int_0^x \mathrm{e}^{\beta t} t^\nu I_\nu(x)\,\mathrm{d}t &\leq \frac{2}{2\nu+1} x K_{\nu+1}(x) I_\nu(x) < \infty, \\ \frac{\mathrm{e}^{-\beta x} K_{\nu}(x)}{x^\nu} \int_0^x \mathrm{e}^{\beta t} t^\nu I_\nu(x)\,\mathrm{d}t &\leq \frac{2}{2\nu+1} x K_{\nu+1}(x) I_\nu(x) < \infty, \\ \left| \left[\frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\mathrm{e}^{-\beta x} K_{\nu}(x)}{x^\nu} \right) \right] \int_0^x \mathrm{e}^{\beta t} t^\nu I_\nu(x)\,\mathrm{d}t \right| &\leq \frac{2(\beta+1)}{2\nu+1} x K_{\nu+1}(x) I_\nu(x) < \infty. \end{aligned}$$

Proof. (i) Note that

$$\frac{d}{d\beta} \left(\frac{\mathrm{e}^{-\beta x} K_{\nu+1}(x)}{x^\nu} \int_0^x \mathrm{e}^{\beta t} t^\nu I_\nu(x)\,\mathrm{d}t \right) = \frac{\mathrm{e}^{-\beta x} K_{\nu+1}(x)}{x^\nu} \int_0^x (t-x) \mathrm{e}^{\beta t} t^\nu I_\nu(x)\,\mathrm{d}t \leq 0.$$

Therefore, for $-1 < \beta < 1$, we have

$$\begin{aligned} \frac{\mathrm{e}^{-\beta x} K_{\nu+1}(x)}{x^\nu} \int_0^x \mathrm{e}^{\beta t} t^\nu I_\nu(x)\,\mathrm{d}t &\leq \frac{\mathrm{e}^{x} K_{\nu+1}(x)}{x^\nu} \int_0^x \mathrm{e}^{-t} t^\nu I_\nu(x)\,\mathrm{d}t \\ &= \frac{1}{2\nu+1} x K_\nu(x) [I_\nu(x) + I_{\nu+1}(x)] \\ &\leq \frac{2}{2\nu+1} x K_{\nu+1}(x) I_\nu(x), \end{aligned}$$

where we used the integral formula (B.45) to evaluate the integral and inequality (B.13) to obtain the final inequality.

Use of the asymptotic formulas (B.7), (B.8), (B.10) and (B.11) for $K_\nu(x)$ and $I_\nu(x)$ verify that the function $xK_{\nu+1}(x)I_\nu(x)$ is bounded in the limits as x tends to 0 and ∞, and its clearly bounded for all other x, and hence is bounded for all $x \geq 0$. This completes the proof of part (i).

(ii) This follows immediately from (i), since, by inequality (B.16), $K_\nu(x) \leq K_{\nu+1}(x)$ for $\nu > -1/2$.

(iii) By the differentiation formula (B.32), we have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{e}^{-\beta x}K_\nu(x)}{x^\nu}\right) = -\mathrm{e}^{-\beta x}\left(\frac{\beta K_\nu(x)}{x^\nu} + \frac{K_{\nu+1}(x)}{x^\nu}\right).$$

We may then obtain the desired inequality by applying parts (i) and (ii). □

Lemma D.4. *Suppose $-1 < \beta < 1$ and $n = 0, 1, 2, \ldots$, then the following inequalities hold*

$$\left\|\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t\right\| < \frac{2^n\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu+1)}, \quad \nu \geq 1/2,$$

and

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

Proof. We begin by bounding $\frac{\mathrm{d}^n}{\mathrm{d}x^n}(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu})$. By the Leibniz rule for differentiating products we have

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right) = \sum_{k=0}^n \binom{n}{k}(-\beta)^{n-k}\mathrm{e}^{-\beta x}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{I_\nu(x)}{x^\nu}\right).$$

Using that $-1 < \beta < 1$ and inequality (C.20) gives

$$\left|\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right| < \sum_{k=0}^n \binom{n}{k}\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu} = \frac{2^n\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}. \tag{D.2}$$

(i) Suppose that $\nu \geq 1/2$. By inequalities (D.2) and (C.33), and that $I_{\nu+1}(x) \leq I_\nu(x)$ for $\nu \geq 1/2$, we have

$$\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t < \frac{2^{n+1}\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)}I_\nu(x)K_\nu(x).$$

Using inequality (D.1) gives

$$\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t < \frac{2^{n+1}\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu)}\cdot\frac{1}{2\nu} = \frac{2^n\sqrt{\pi}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}\Gamma(\nu+1)},$$

as required.

(ii) Suppose now that $-1/2 < \nu < 1/2$. We begin by proving that the bound holds in the region $x \geq 1/2$. By inequalities (D.2) and (C.31), and $I_{\nu+1}(x) \leq I_\nu(x)$ for $\nu > -1/2$, we have

$$\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t < \frac{2^n}{1-|\beta|}I_\nu(x)K_\nu(x).$$

By Proposition D.1, $I_\nu(x)K_\nu(x)$ is a monotone decreasing function of x for $x > 0$, and therefore we may bound this product for $x \geq 1/2$ by $I_\nu(1/2)K_\nu(1/2)$. In fact, we may produce a bound for all $0 \leq \nu < 1/2$ using (B.12) and (B.14), which gives $I_\nu(1/2)K_\nu(1/2) < I_{-1/2}(1/2)K_{1/2}(1/2) = 1+\mathrm{e}^{-1}$, where we used the formulas (B.5) and (B.6) for $I_{-1/2}(x)$ and $K_{1/2}(x)$, respectively, to obtain the equality. Putting this together we have

$$\mathrm{e}^{-\beta x}\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t < \frac{(1+\mathrm{e}^{-1})2^n}{1-|\beta|}, \quad x \geq 1/2,\ -1/2 < \nu < 1/2.$$

For $-1/2 < \nu < 1/2$, $\Gamma(\nu+1/2) > \Gamma(1) = 1$. Therefore $1+\mathrm{e}^{-1} < (\mathrm{e}+1)2^{2\nu}\Gamma(\nu+1/2)$ for $-1/2 < \nu < 1/2$, and thus the bound holds in the region $x \geq 1/2$.

We now verify that the bound holds in the region $0 \leq x \leq 1/2$. By inequality (D.2), we have

$$\left[\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t < \frac{2^n\mathrm{e}^{-\beta x}I_\nu(x)}{x^\nu}\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t.$$

From the series expansion (B.1) for $I_\nu(x)$ we can easily deduce that $x^{-\nu}I_\nu(x)$ is an increasing function of x. The integral $\int_x^\infty \mathrm{e}^{\beta t}t^\nu K_\nu(t)\,\mathrm{d}t$ is a decreasing function of x, and so we may bound the right-hand side of the previous display by

$$\frac{2^n\mathrm{e}^{-\beta/2}I_\nu(1/2)}{(1/2)^\nu}\int_{-\infty}^\infty \mathrm{e}^{\beta t}|t|^\nu K_\nu(|t|)\,\mathrm{d}t = \frac{\sqrt{\pi}\mathrm{e}^{-\beta/2}I_\nu(1/2)2^{n+\nu}\Gamma(\nu+1/2)}{(1-\beta^2)^{\nu+1/2}}, \tag{D.3}$$

where the integral was evaluated using formula (B.43). For $-1 < \beta < 1$ and $-1/2 < \nu < 1/2$ the following inequalities hold: $\mathrm{e}^{-\beta/2} < \mathrm{e}^{1/2}$, $I_\nu(1/2) < I_{-1/2}(1/2) = \pi^{-1/2}(\mathrm{e}^{1/2}+\mathrm{e}^{-1/2})$ and $(1-\beta^2)^{\nu+1/2} > 1-|\beta|$. With these bounds we may bound the right-hand side of (D.3), and thus obtain, for $0 \leq x \leq 1/2$ and $-1/2 < \nu < 1/2$,

$$\frac{2^n\mathrm{e}^{-\beta/2}I_\nu(1/2)}{(1/2)^\nu}\int_{-\infty}^\infty \mathrm{e}^{\beta t}|t|^\nu K_\nu(|t|)\,\mathrm{d}t < \frac{(\mathrm{e}+1)2^{n+1/2}\Gamma(\nu+1/2)}{1-|\beta|}.$$

Combining this bound with the bound for $x \geq 1/2$, and using that $\Gamma(\nu+1/2) > 1$ for $-1/2 < \nu < 1/2$, completes the proof of part (ii). $\square$

We now specialise to the case $\beta = 0$. The following lemmas give bounds of order ν^{-1} as $\nu \to \infty$ for expressions of type (ii) and (iii). We also obtain a bound for the term (i) (see Lemma D.6), but this bound is of order $\nu^{-1/2}$ as $\nu \to \infty$ (see Olver et al. [52], formula 5.11.12.).

Lemma D.5. *Let $I_{(\nu,0,n)}(x)$ be defined as per equation (3.20). Suppose $\nu > -1/2$, then*

$$\left\| I_{(\nu,0,n)}(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{K_\nu(x)}{x^\nu}\right)\right\| \leq \frac{2^{n-1}}{2\nu+1}, \qquad n \geq 1, \tag{D.4}$$

and

$$\left\| I_{(\nu,0,1)}(x) \frac{K_\nu(x)}{x^\nu}\right\| < \frac{1}{2\nu+1}. \tag{D.5}$$

Proof. (i) Applying inequalities (C.19) and (C.25) gives

$$\left| I_{(\nu,0,n)}(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| \leq \left\{\prod_{k=1}^{n} \frac{2\nu+2k}{2\nu+k}\right\} K_{\nu+n}(x) I_{\nu+n+1}(x).$$

We now use that $K_{\nu+n}(x) < K_{\nu+n+1}(x)$ and inequality (D.1) to obtain

$$\left| I_{(\nu,0,n)}(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \left\{\prod_{k=1}^{n} \frac{2\nu+2k}{2\nu+k}\right\} K_{\nu+n+1}(x) I_{\nu+n+1}(x) \leq \frac{1}{2(\nu+n+1)} \prod_{k=1}^{n} \frac{2\nu+2k}{2\nu+k}.$$

We can simplify the bound given in the above display by noting

$$\frac{1}{2(\nu+n+1)} \prod_{k=1}^{n} \frac{2\nu+2k}{2\nu+k} = \frac{1}{2\nu+1} \cdot \frac{2(\nu+n)}{2(\nu+n+1)} \prod_{k=1}^{n-1} \frac{2\nu+2k}{2\nu+k+1} < \frac{2^{n-1}}{2\nu+1},$$

as $\frac{2\nu+2k}{2\nu+k+1} \leq \frac{2k-1}{k} < 2$, for $\nu > -1/2$ and $k \geq 2$. This completes the proof of part (i).

(ii) We note that, by inequality (B.16) and the differentiation formula (B.31), $x^{-\nu}K_\nu(x) < x^{-\nu}K_{\nu+1}(x) = -\frac{\mathrm{d}}{\mathrm{d}x}(K_\nu(x))$ and apply part (i). □

Lemma D.6. *Let $\nu > -1/2$, then*

$$\left\| \frac{I_\nu(x)}{x^\nu} \int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \right\| = \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{2\Gamma(\nu+1)}. \tag{D.6}$$

Proof. We begin by proving that $\frac{I_\nu(x)}{x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t$ is a decreasing function. By the differentiation formula (B.31) and inequality (C.29), we have, for $x \geq 0$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{I_\nu(x)}{x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t\right) &= \frac{I_{\nu+1}(x)}{x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t - I_\nu(x)K_\nu(x) \\ &\leq I_{\nu+1}(x)K_{\nu+1}(x) - I_\nu(x)K_\nu(x).\end{aligned}$$

Theorem 1 of Segura [67] states that for $x > 0$ and $\mu > -1/2$ the inequality $I_{\mu+1}(x)K_{\mu+1}(x) - I_\mu(x)K_\mu(x) < 0$ holds. Hence, $\frac{I_\nu(x)}{x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t$ is a decreasing function of x.

Using the asymptotic property (B.7) of modified Bessel functions $I_\nu(x)$ and that $\int_0^\infty t^\nu K_\nu(t)\,\mathrm{d}t = \sqrt{\pi}\Gamma(\nu+1/2)2^{\nu-1}$ (see (B.43)), we have

$$\lim_{x\to 0^+}\left(\frac{I_\nu(x)}{x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t\right) = \frac{\sqrt{\pi}\Gamma(\nu+1/2)}{2\Gamma(\nu+1)},$$

proving the result. □

Lemma D.7. *Let $\nu > -1/2$ and $n = 0, 1, 2, \ldots$, then*

$$\left\|\left[\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t\right\| < \frac{1}{2(\nu+2)} + \frac{\sqrt{\pi}}{2(2\nu+1)}, \tag{D.7}$$

and

$$\left\|\left[\frac{\mathrm{d}^{2n+1}}{\mathrm{d}x^{2n+1}}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(y)\,\mathrm{d}t\right\| < \frac{1}{2(\nu+1)}. \tag{D.8}$$

Proof. (i) We begin by obtaining the following inequality:

$$\begin{aligned}\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{I_\nu(x)}{x^\nu}\right) &= \sum_{k=0}^n A_k^n(\nu)\frac{I_{\nu+2k}(x)}{x^\nu} && \text{by (C.11)} \\ &\leq \sum_{k=1}^n A_k^n(\nu)\frac{I_{\nu+2}(x)}{x^\nu} + A_0^n(\nu)\frac{I_\nu(x)}{x^\nu} && \text{by inequality (B.13)} \\ &= (1 - A_0^n(\nu))\frac{I_{\nu+2}(x)}{x^\nu} + A_0^n(\nu)\frac{I_\nu(x)}{x^\nu} && \text{as } \sum_{k=0}^n A_k^n(\nu) = 1 \\ &\leq \frac{I_{\nu+2}(x)}{x^\nu} + \frac{I_\nu(x)}{2(\nu+1)x^\nu} && \text{as } A_0^n(\nu) \leq \frac{1}{2(\nu+1)} \text{ for } n \geq 1.\end{aligned}$$

Therefore

$$\begin{aligned}\left[\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t &\leq \frac{I_{\nu+2}(x)}{x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t + \frac{I_\nu(x)}{2(\nu+1)x^\nu}\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t \\ &< I_{\nu+2}(x)K_{\nu+1}(x) + \frac{1}{2(\nu+1)}\cdot\frac{\sqrt{\pi}\Gamma(\nu+1/2)}{2\Gamma(\nu+1)},\end{aligned}$$

where we used inequalities (C.29) and (D.6) to obtain the second inequality. We now use that $K_{\nu+1} < K_{\nu+2}(x)$ for $\nu > -1/2$ and inequality (D.1) to obtain

$$\begin{aligned}\left[\frac{\mathrm{d}^{2n}}{\mathrm{d}x^{2n}}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t &< I_{\nu+2}(x)K_{\nu+2}(x)+\frac{1}{2(\nu+1)}\cdot\frac{\sqrt{\pi}\Gamma(\nu+1/2)}{2\Gamma(\nu+1)}\\ &\leq \frac{1}{2(\nu+2)}+\frac{\sqrt{\pi}\Gamma(\nu+3/2)}{2(2\nu+1)\Gamma(\nu+2)}\\ &\leq \frac{1}{2(\nu+2)}+\frac{\sqrt{\pi}}{2(2\nu+1)}.\end{aligned}$$

(ii) Applying inequalities (C.20) and (C.29) and then inequality (D.1) gives

$$\left[\frac{\mathrm{d}^{2n+1}}{\mathrm{d}x^{2n+1}}\left(\frac{I_\nu(x)}{x^\nu}\right)\right]\int_x^\infty t^\nu K_\nu(t)\,\mathrm{d}t < \frac{I_{\nu+1}(x)}{x^\nu}\cdot x^\nu K_{\nu+1} = I_{\nu+1}(x)K_{\nu+1}(x) \leq \frac{1}{2(\nu+1)},$$

as required □

D.2 Bounds for $x \geq 1$

The following bounds are used in the proof of Lemma 3.19. Note the expressions in Lemma D.8 are undefined in the limit $x \downarrow 0$. This is easily verified using the asymptotic formulas (B.7) and (B.8).

Lemma D.8. *Let $\nu > -1/2$ and $0 \leq k \leq n$, where $n = 1, 2, 3, \ldots$, then for $x \geq 1$,*

$$\frac{K_{\nu+n+1}(x)}{x^\nu}\int_0^x t^{\nu-k} I_{\nu+n}(t)\,\mathrm{d}t \leq \frac{1}{2\nu+n-k+1}, \tag{D.9}$$

and

$$\frac{I_{\nu+n}(x)}{x^\nu}\int_x^\infty t^{\nu-k} K_{\nu+n}(t)\,\mathrm{d}t < \frac{\sqrt{\pi}}{\sqrt{4(\nu+n)+1}}. \tag{D.10}$$

Proof. (i) Suppose $x \geq 1$. Then,

$$\begin{aligned}\frac{K_{\nu+n+1}(x)}{x^\nu}\int_0^x t^{\nu-k} I_{\nu+n}(t)\,\mathrm{d}t &\leq \frac{2(\nu+n+1)}{2\nu+n-k+1}x^{\nu-k}I_{\nu+n+1}(x)\cdot\frac{K_{\nu+n+1}(x)}{x^\nu} && \text{by (C.24)}\\ &\leq \frac{2(\nu+n+1)}{2\nu+n-k+1}K_{\nu+n+1}(x)I_{\nu+n+1}(x) && \text{as } x \geq 1\\ &\leq \frac{1}{2\nu+n-k+1}, && \text{by (D.1)}\end{aligned}$$

as required.

(ii) Again, we suppose $x \geq 1$. We have

$$\begin{aligned}
\frac{I_{\nu+n}(x)}{x^{\nu}}\int_x^{\infty} t^{\nu-k}K_{\nu+n}(t)\,\mathrm{d}t &\leq \frac{I_{\nu+n}(x)}{x^{\nu+n+k}}\int_x^{\infty} t^{\nu+n}K_{\nu+n}(t)\,\mathrm{d}t \\
&\leq \frac{\sqrt{\pi}\Gamma(\nu+n+1/2)}{\Gamma(\nu+n)}x^{-k}I_{\nu+n}(x)K_{\nu+n}(x) && \text{by (D.1)} \\
&\leq \frac{\sqrt{\pi}\Gamma(\nu+n+1/2)}{\Gamma(\nu+n)}I_{\nu+n}(x)K_{\nu+n}(x) && \text{as } x \geq 1 \\
&\leq \frac{\sqrt{\pi}\Gamma(\nu+n+1/2)}{2\Gamma(\nu+n+1)} && \text{by (C.32)} \\
&< \frac{\sqrt{\pi}}{\sqrt{4(\nu+n)+1}},
\end{aligned}$$

where the final inequality was obtained by an application of the following inequality, which can be found in Elezović et al. [24]:

$$\frac{\Gamma(\mu+1/2)}{\Gamma(\mu+1)} < \frac{1}{\sqrt{\mu+1/4}}, \quad \mu > -1/4.$$

The proof is complete. □

D.3 Bounds for $0 \leq x \leq 1$

In this section we obtain bound for the terms of type (iv) for the case $\beta = 0$ in the region $0 \leq x \leq 1$. In Lemma D.20 we specalise to the case $\nu = 0$ and obtain uniform bounds in the region $0 \leq x < \infty$.

We begin by stating some preliminary lemmas, which are required in the proofs of Lemmas D.16, D.17 and D.18. These preliminary lemmas bound the following expressions:

$$\begin{aligned}
\alpha_{p,q,r}(x) &= \frac{2^{\nu+r-1}\Gamma(\nu+r)}{x^{\nu+r}}\sum_{k=p}^{\infty}\frac{1}{\Gamma(\nu+k+1)k!}\left\{\prod_{i=1}^{q}\frac{1}{2\nu+2k+i}\right\}\frac{x^{\nu+2k+q}}{2^{\nu+2k}}, \\
\beta_{p,q,r}(x) &= \frac{I_{(\nu,0,q)}(x)}{2x^{\nu}}\left|\sum_{k=p}^{\lceil\nu+r\rceil-1}\frac{(-1)^k\Gamma(\nu+r-k)}{k!}\left(\frac{x}{2}\right)^{2k-\nu-r}\right|, \\
\gamma_{q,r}(x) &= \frac{I_{(\nu,0,q)}(x)}{x^{\nu}}\cdot\left|\log\left(\frac{x}{2}\right)+\gamma\right|I_{\nu+r}(x), \\
\delta_{q,r}(x) &= \frac{I_{(\nu,0,q)}(x)}{x^{\nu}}\cdot\frac{1}{2}\sum_{j=0}^{\infty}\frac{\{\psi(j)+\psi(\nu+j+r)\}}{j!\,(\nu+j+r)!}\left(\frac{x}{2}\right)^{\nu+2j+r},
\end{aligned}$$

$$\begin{aligned}\epsilon_{q,r}(x) &= \frac{I_{(\nu,0,q)}(x)}{x^\nu}\cdot\frac{\pi}{2|\sin(\nu+r)\pi|}\sum_{l=\lceil\nu+r\rceil}^{\infty}\frac{1}{\Gamma(l-\nu-r+1)l!}\left(\frac{x}{2}\right)^{2l-\nu-r},\\ \zeta_{q,r}(x) &= \frac{I_{(\nu,0,q)}(x)}{x^\nu}\cdot\frac{\pi}{2|\sin(\nu+r)\pi|}I_{\nu+r}(x),\end{aligned}$$

where p, q and r are positive integers, γ is the EulerMascheroni constant and $\psi(k)=\sum_{j=1}^k \frac{1}{j}$.

Lemma D.9. *Suppose that positive integers p,q and r satisfy $2p+q\geq r$, then for $\nu>-1/2$ and $0\leq x\leq 1$,*

$$\alpha_{p,q,r}(x)<\frac{2^{r-2p}\Gamma(\nu+r)}{p!\,\Gamma(\nu+p+1)}\prod_{i=1}^{q}\frac{1}{2\nu+2p+i},$$

and in particular

$$\begin{aligned}\alpha_{2,1,3}(x) &< \frac{1}{4(2\nu+5)},\\ \alpha_{2,2,4}(x) &< \frac{1}{4(2\nu+5)},\\ \alpha_{3,1,4}(x) &< \frac{1}{24(2\nu+7)}.\end{aligned}$$

Proof. Setting $k=j+p$ gives

$$\begin{aligned}\alpha_{p,q,r}(x) &= 2^{r-2p-1}\Gamma(\nu+r)x^{2p+q-r}\sum_{j=0}^{\infty}\frac{1}{\Gamma(\nu+j+p+1)(j+p)!}\left\{\prod_{i=1}^{q}\frac{1}{2\nu+2j+2p+i}\right\}\left(\frac{x}{2}\right)^{2j}\\ &\leq \frac{2^{r-2p-1}\Gamma(\nu+r)}{p!\,\Gamma(\nu+p+1)}x^{2p+q-r}\left\{\prod_{i=1}^{q}\frac{1}{2\nu+2p+i}\right\}\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{x^2}{4}\right)^{j}\\ &= \frac{2^{r-2p-1}\Gamma(\nu+r)}{p!\,\Gamma(\nu+p+1)}x^{2p+q-r}\mathrm{e}^{x^2/4}\prod_{i=1}^{q}\frac{1}{2\nu+2p+i}\\ &< \frac{2^{r-2p}\Gamma(\nu+r)}{p!\,\Gamma(\nu+p+1)}\prod_{i=1}^{q}\frac{1}{2\nu+2p+i},\end{aligned}$$

where we used that $(j+p)!\geq j!\,p!$ to obtain the first inequality, and the final inequality follows because for $0\leq x\leq 1$ and $2p+q\geq r$ we have that $x^{2p+q-r}\mathrm{e}^{x^2/4}\leq \mathrm{e}^{1/4}<2$.

The bounds for $\alpha_{2,1,3}(x)$, $\alpha_{2,2,4}(x)$ and $\alpha_{3,1,4}(x)$ follow from substituting the relevant values of p, q and r into the bound for $\alpha_{p,q,r}(x)$. □

Lemma D.10. *Suppose that positive integers p,q and r satisfy $2p+q\geq r$, then for $\nu>-1/2$ and $0\leq x\leq 1$,*

$$\beta_{p,q,r}(x)<\frac{\Gamma(\nu+r-p)}{2^{2p-r}p!\,\Gamma(\nu+1)}\prod_{k=1}^{q}\frac{1}{2\nu+k},$$

and in particular

$$\begin{aligned}\beta_{2,1,3}(x) &< \frac{1}{4(2\nu+1)},\\ \beta_{2,2,4}(x) &< \frac{1}{4(2\nu+1)},\\ \beta_{3,1,4}(x) &< \frac{1}{24(2\nu+1)}.\end{aligned}$$

Proof. We bound $\beta_{p,q,r}(x)$ as follows

$$\begin{aligned}\beta_{p,q,r}(x) &\leq \frac{I_{(\nu,0,q)}(x)}{2x^\nu}\cdot\frac{\Gamma(\nu+r-p)}{p!}\left(\frac{x}{2}\right)^{2p-\nu-r}\\ &\leq \left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}I_{\nu+q}(x)\cdot\frac{\Gamma(\nu+r-p)}{p!}\left(\frac{x}{2}\right)^{2p-\nu-r} && \text{by (C.25)}\\ &\leq \frac{\Gamma(\nu+r-p)}{2^{2p+q-r+1}p!\,\Gamma(\nu+q+1)}\left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}x^{2p+q-r}\cosh(x) && \text{by (B.17)}\\ &< \frac{\Gamma(\nu+r-p)}{2^{2p+q-r}p!\,(\Gamma(\nu+q+1)}\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\\ &= \frac{\Gamma(\nu+r-p)}{2^{2p+q-r}p!}\cdot\frac{2^q\prod_{j=1}^{q}(\nu+j)}{\Gamma(\nu+q+1)}\prod_{k=1}^{q}\frac{1}{2\nu+k}\\ &= \frac{\Gamma(\nu+r-p)}{2^{2p-r}p!\,\Gamma(\nu+1)}\prod_{k=1}^{q}\frac{1}{2\nu+k},\end{aligned}$$

where to obtain the third inequality we used that for $2p+q\geq r$ and $0\leq x\leq 1$ we have $x^{2p+q-r}\cosh(x)\leq\cosh(1)=1.54\ldots<2$.

The bounds for $\beta_{2,1,3}(x)$, $\beta_{2,2,4}(x)$ and $\beta_{3,1,4}(x)$ follow from substituting the relevant values of p, q and r into the bound for $\beta_{p,q,r}(x)$. □

Lemma D.11. *Suppose that q and r are positive integers, then for $\nu\in\mathbb{N}$ and $0\leq x\leq 1$,*

$$\gamma_{q,r}(x) < \frac{1}{2^{2\nu+r-1}\nu!\,(\nu+r)!}\prod_{k=1}^{q}\frac{1}{2\nu+k}.$$

In particular,

$$\begin{aligned}\gamma_{1,3}(x) &< \frac{1}{2^{2\nu+2}\nu!\,(\nu+3)!\,(2\nu+1)},\\ \gamma_{2,4}(x) &< \frac{1}{2^{2\nu+4}(\nu+1)!\,(\nu+4)!\,(2\nu+1)},\end{aligned}$$

and

$$\gamma_{1,4}(x) < \frac{1}{2^{2\nu+3}\nu!\,(\nu+4)!\,(2\nu+1)}.$$

Proof. Using inequalities (C.25) and (B.17) and that $|\log(\frac{x}{2})+\gamma| \leq -\log(\frac{x}{2})$ for $0 \leq x \leq 1$, we have

$$\begin{aligned}\gamma_{q,r}(x) &\leq \left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}I_{\nu+q}(x)\cdot\left|\log\left(\frac{x}{2}\right)+\gamma\right|I_{\nu+r}(x)\\ &\leq -\left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}\log\left(\frac{x}{2}\right)\frac{x^{2\nu+q+r}(\cosh(x))^2}{2^{2\nu+q+r}(\nu+q)!\,(\nu+r)!}.\end{aligned}$$

Since $2\nu+q+r \geq 1$ and $0 \leq x \leq 1$ we have that $-x^{2\nu+q+r}\log(x/2) \leq -\frac{x}{2}\log(\frac{x}{2})$, and by elementary calculus $-\frac{x}{2}\log(\frac{x}{2}) \leq 2\mathrm{e}^{-1}$, and therefore

$$\gamma_{q,r}(x) \leq \left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}\frac{2\mathrm{e}^{-1}(\cosh(1))^2}{2^{2\nu+q+r}(\nu+q)!\,(\nu+r)!} < \frac{1}{2^{2\nu+r-1}\nu!\,(\nu+r)!}\prod_{k=1}^{q}\frac{1}{2\nu+k},$$

where we used that $2\mathrm{e}^{-1}(\cosh(1))^2 = 1.75\ldots < 2$ to obtain the final inequality.

The bounds for $\gamma_{1,3}(x)$, $\gamma_{2,4}(x)$ and $\gamma_{1,4}(x)$ follow from substituting the relevant values of p, q and r into the bound for $\gamma_{q,r}(x)$. □

Lemma D.12. *Suppose that q and r are positive integers, then for $\nu \in \mathbb{N}$ and $0 \leq x \leq 1$,*

$$\delta_{q,r}(x) < \frac{1}{2^{2\nu+r-1}\nu!\,(\nu+r-1)!}\prod_{k=1}^{q}\frac{1}{2\nu+k},$$

and in particular

$$\begin{aligned}\delta_{1,3}(x) &< \frac{1}{2^{2\nu+2}\nu!\,(\nu+2)!\,(2\nu+1)},\\ \delta_{2,4}(x) &< \frac{1}{2^{2\nu+4}(\nu+1)!\,(\nu+3)!\,(2\nu+1)},\\ \delta_{1,4}(x) &< \frac{1}{2^{2\nu+3}\nu!\,(\nu+3)!\,(2\nu+1)}.\end{aligned}$$

Proof. We begin by noting that

$$\frac{\{\psi(j)+\psi(\nu+j+r)\}}{(\nu+j+r)!} \leq \frac{2\psi(\nu+j+r)}{(\nu+j+r)!} \leq \frac{2(\nu+j+r)}{(\nu+j+r)!} = \frac{2}{(\nu+j+r-1)!}.$$

Hence,

$$\begin{aligned}
\delta_{q,r}(x) &\le \left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}I_{\nu+q}(x)\cdot\frac{1}{(\nu+r-1)!}\left(\frac{x}{2}\right)^{\nu+r}\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{x}{2}\right)^{2j} && \text{by (C.25)}\\
&\le \frac{1}{2^{\nu+r}(\nu+r-1)!}\left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}x^{\nu+r}\mathrm{e}^{x^2/4}\cdot\frac{x^{\nu+q}\cosh(x)}{2^{\nu+q}(\nu+q)!} && \text{by (B.17)}\\
&< \frac{1}{2^{2\nu+r+q-1}(\nu+q)!\,(\nu+r-1)!}\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\\
&= \frac{1}{2^{2\nu+r-1}\nu!\,(\nu+r-1)!}\prod_{k=1}^{q}\frac{1}{2\nu+k},
\end{aligned}$$

where we used to obtain the final inequality we used for that, for $2\nu+q+r\ge 0$ and $0\le x\le 1$, we have $x^{2\nu+q+r}\mathrm{e}^{x^2/4}\cosh(x)\le \mathrm{e}^{1/4}\cosh(1)=1.98\ldots<2$.

The bounds for $\delta_{1,3}(x)$, $\delta_{2,4}(x)$ and $\delta_{1,4}(x)$ follow from substituting the relevant values of q and r into the bound for $\delta_{q,r}(x)$. □

Lemma D.13. *Suppose that q and r are positive integers, then for $0\le x\le 1$,*

$$\epsilon_{q,r}(x) < \frac{1}{|\sin(\nu\pi)|2^{2\nu+r-2}\Gamma(\nu+1)\Gamma(\nu+r+1)}\prod_{k=1}^{q}\frac{1}{2\nu+k},$$

and in particular

$$\begin{aligned}
\epsilon_{1,3}(x) &< \frac{1}{|\sin(\nu\pi)|2^{2\nu+1}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1)},\\
\epsilon_{2,4}(x) &< \frac{1}{|\sin(\nu\pi)|2^{2\nu+3}\Gamma(\nu+2)\Gamma(\nu+5)(2\nu+1)},\\
\epsilon_{1,4}(x) &< \frac{1}{|\sin(\nu\pi)|2^{2\nu+2}\Gamma(\nu+1)\Gamma(\nu+5)(2\nu+1)}.
\end{aligned}$$

Proof. Setting $l=\lceil\nu+r\rceil+j$ gives

$$\begin{aligned}
&\frac{\pi}{2|\sin(\nu\pi)|}\sum_{l=\lceil\nu+r\rceil}^{\infty}\frac{1}{\Gamma(l-\nu-r+1)l!}\left(\frac{x}{2}\right)^{2l-\nu-r}\\
&= \frac{\pi}{2|\sin(\nu\pi)|}\sum_{j=0}^{\infty}\frac{1}{\Gamma(k+\lceil\nu+r\rceil+j-\nu-r+1)(\lceil\nu+r\rceil+j)!}\left(\frac{x}{2}\right)^{2j+2\lceil\nu+r\rceil-\nu-r}\\
&\le \frac{\pi}{2|\sin(\nu\pi)|\Gamma(\lceil\nu+r\rceil+j-\nu-r+1)\Gamma(\nu+r+1)}\left(\frac{x}{2}\right)^{2\lceil\nu+r\rceil-\nu-r}\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{x}{2}\right)^{2j}
\end{aligned}$$

$$\begin{aligned}
&= \frac{\pi e^{x^2/4}}{2|\sin(\nu\pi)|\Gamma(\lceil\nu\rceil-\nu+1)\Gamma(\nu+r+1)}\left(\frac{x}{2}\right)^{2\lceil\nu+r\rceil-\nu-r} \\
&< \frac{\pi e^{1/4}}{2^{r+1}\cdot 0.88|\sin(\nu\pi)\Gamma(\nu+r+1)},
\end{aligned}$$

where the final inequality follows from by noting that $(\frac{x}{2})^{2\lceil\nu+r\rceil-\nu-r}e^{x^2/4} \leq (\frac{x}{2})^{\nu+r}e^{x^2/4} \leq 2^{-r-\nu}e^{1/4}$ for $0 \leq x \leq 1$, and that $\Gamma(\lceil\nu\rceil-\nu+1) \geq \inf_{\nu\in[1,2]}\Gamma(\nu) > 0.88$. Therefore

$$\begin{aligned}
\epsilon_{q,r}(x) &< \left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}I_{\nu+q}(x)\cdot\frac{\pi e^{1/4}}{2^r\cdot 1.76|\sin(\nu\pi)\Gamma(\nu+r+1)} && \text{by (C.25)} \\
&\leq \frac{\pi e^{1/4}}{1.76\cdot 2^r\cdot|\sin(\nu\pi)\Gamma(\nu+r+1)}\left\{\prod_{k=1}^{q}\frac{2\nu+2k}{2\nu+k}\right\}\frac{x^{\nu+q}\cosh(x)}{2^{\nu+q}\Gamma(\nu+q+1)} && \text{by (B.17)} \\
&< \frac{1}{|\sin(\nu\pi)|2^{2\nu+r-2}\Gamma(\nu+1)\Gamma(\nu+r+1)}\prod_{k=1}^{q}\frac{1}{2\nu+k},
\end{aligned}$$

where to obtain the final inequality we used that $\frac{\pi e^{1/4}}{1.76}x^{\nu+q}\cosh(x) \leq \frac{\pi e^{1/4}}{1.76}\cosh(1) = 3.53\ldots < 4$.

The bounds for $\epsilon_{1,3}(x)$, $\epsilon_{2,4}(x)$ and $\epsilon_{1,4}(x)$ follow from substituting the relevant values of q and r into the bound for $\epsilon_{q,r}(x)$. □

Lemma D.14. *Suppose that q and r are positive integers, then for $0 \leq x \leq 1$,*

$$\zeta_{q,r}(x) < \frac{1}{|\sin(\nu\pi)|2^{2\nu+r-2}\Gamma(\nu+1)\Gamma(\nu+r+1)}\prod_{k=1}^{q}\frac{1}{2\nu+k},$$

and in particular

$$\begin{aligned}
\zeta_{1,3}(x) &< \frac{1}{|\sin(\nu\pi)|2^{2\nu+1}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1)}, \\
\zeta_{2,4}(x) &< \frac{1}{|\sin(\nu\pi)|2^{2\nu+3}\Gamma(\nu+2)\Gamma(\nu+5)(2\nu+1)}, \\
\zeta_{1,4}(x) &< \frac{1}{|\sin(\nu\pi)|2^{2\nu+2}\Gamma(\nu+1)\Gamma(\nu+5)(2\nu+1)}.
\end{aligned}$$

Proof. The proof is similar to the proof of Lemma D.13, with the only difference being that we don't have to deal with the term $|\log(\frac{x}{2})+\gamma|$, and that we now use that $\frac{\pi}{2}(\cosh(1))^2 = 3.74\ldots < 4$. □

With the preliminary lemmas now stated, we are now in a position to obtain bounds for $0 \leq x \leq 1$ for the expressions of type (iv) for the case $\beta = 0$.

Lemma D.15. *Suppose $\nu > -1/2$ and $n \geq 2$, then for $0 \leq x \leq 1$,*

$$\bigg|\frac{1}{x} - (-1)^n I_{(\nu,0,n-1)}(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\bigg(\frac{K_\nu(x)}{x^\nu}\bigg)\bigg| < \frac{1+2^{n-1}}{2\nu+1}. \tag{D.11}$$

Proof. We begin by proving the result for the case of even n. Let $n = 2m$. Using the differentiation formula (C.9) and the triangle inequality gives

$$\bigg|\frac{1}{x} - I_{(\nu,0,2m-1)}(x)\frac{\mathrm{d}^{2m}}{\mathrm{d}x^{2m}}\bigg(\frac{K_\nu(x)}{x^\nu}\bigg)\bigg| \leq R_1 + R_2,$$

where

$$\begin{aligned} R_1 &= I_{(\nu,0,2m-1)}(x)\sum_{k=0}^{m-1} A_k^m(\nu)\frac{K_{\nu+2k}(x)}{x^\nu}, \\ R_2 &= \bigg|\frac{1}{x} - A_m^m(\nu)\frac{K_{\nu+2m}(x)I_{(\nu,0,2m-1)}(x)}{x^\nu}\bigg|. \end{aligned}$$

We begin by bounding R_1. Using the inequality $K_\nu(x) < K_{\nu+1}(x)$ for $\nu > -1/2$, and the identity $\sum_{k=0}^m A_k^m(\nu) = 1$, we have

$$R_1 < x^{-\nu}K_{\nu+2m-2}(x)I_{(\nu,0,2m-1)}(x)\sum_{k=0}^{m-1} A_k^m(\nu) = (1 - A_m^m(\nu))x^{-\nu}K_{\nu+2m-2}(x)I_{(\nu,0,2m-1)}(x).$$

We therefore have

$$\begin{aligned} R_1 &< (1 - A_m^m(\nu))\bigg\{\prod_{k=1}^{2m-1}\frac{2\nu+2k}{2\nu+k}\bigg\}K_{\nu+2m-2}(x)I_{\nu+2m-1}(x) && \text{by (C.25)} \\ &< \bigg\{\prod_{k=1}^{2m-1}\frac{2\nu+2k}{2\nu+k}\bigg\}K_{\nu+2m-1}(x)I_{\nu+2m-1}(x) && \text{by (B.16) and } 0 \leq A_m^m(\nu) \leq 1 \\ &\leq \frac{1}{2(\nu+2m-1)}\prod_{k=1}^{2m-1}\frac{2\nu+2k}{2\nu+k} && \text{by (D.1)} \qquad \text{(D.12)} \\ &= \frac{1}{2\nu+1}\prod_{k=1}^{2m-2}\frac{2\nu+2k}{2\nu+k+1} \\ &< \frac{2^{n-2}}{2\nu+1} && \text{recall that } n = 2m, \qquad \text{(D.13)} \end{aligned}$$

where to obtain the final inequality we used that $\frac{2\nu+2k}{2\nu+k+1} \leq \frac{2k-1}{k-1} < 2$, for $\nu > -1/2$ and $k \geq 2$. We now bound R_2. By the Mean Value Theorem we have

$$\frac{I_\nu(x)}{x^\nu} = \frac{1}{\Gamma(\nu+1)2^\nu} + x\bigg[\frac{\mathrm{d}}{\mathrm{d}x}\bigg(\frac{I_\nu(x)}{x^\nu}\bigg)\bigg]_{x=\eta},$$

where $0 < \eta < x$, and we used that, by the series expansion (B.1) of $I_\nu(x)$, we have $\lim_{x\downarrow 0} x^{-\nu}I_\nu(x) = \frac{1}{\Gamma(\nu+1)2^\nu}$. As $\frac{\mathrm{d}}{\mathrm{d}x}(x^{-\nu}I_\nu(x)) = x^{-\nu}I_{\nu+1}(x)$ is a positive monotone increasing function of x, it follows that

$$\frac{x^{2\nu}}{\Gamma(\nu+1)2^\nu} \leq x^\nu I_\nu(x) \leq \frac{x^{2\nu}}{\Gamma(\nu+1)2^\nu} + x^{\nu+1}I_{\nu+1}(x).$$

Integrating $n-1$ times with respect to x over the interval $[0,t]$ gives

$$\begin{aligned}\frac{t^{2\nu+n-1}}{\Gamma(\nu+1)2^\nu\prod_{j=0}^{n-1}(2\nu+j+1)} &\leq I_{(\nu,0,n-1)}(t) \\ &\leq \frac{t^{2\nu+n-1}}{\Gamma(\nu+1)2^\nu\prod_{j=0}^{n-1}(2\nu+j+1)} + I_{(\nu+1,0,n-1)}(t).\end{aligned}$$

We may therefore bound R_2 as follows

$$\begin{aligned}R_2 &\leq \left|\frac{1}{x} - \frac{A_m^m(\nu)x^{\nu+2m-1}K_{\nu+2m}(x)}{\Gamma(\nu+1)2^\nu\prod_{j=0}^{2m-1}(2\nu+j+1)}\right| + A_m^m(\nu)x^{-\nu}K_{\nu+2m}(x)I_{(\nu+1,0,2m-1)}(x) \\ &= x\left|\frac{1}{x^2} - \frac{x^{\nu+2m-2}K_{\nu+2m}(x)}{2^{\nu+2m-1}\Gamma(\nu+2m)}\right| + A_m^m(\nu)x^{-\nu}K_{\nu+2m}(x)I_{(\nu+1,0,2m-1)}(x),\end{aligned}$$

where we used inequality (B.16) for $K_\mu(x)$ to obtain the first inequality and to obtain the equality we used that

$$A_m^m(\nu) = \frac{\prod_{j=0}^{m-1}(2\nu+2j+1)}{2^m\prod_{j=0}^{m-1}(\nu+m+j)} = \frac{\prod_{k=0}^{2m-1}(2\nu+k+1)}{2^{2m-1}\prod_{k=0}^{2m-1}(\nu+k)} = \frac{\Gamma(\nu+1)\prod_{k=0}^{2m-1}(2\nu+k+1)}{2^{2m-1}\Gamma(\nu+2m)}.$$

Using Lemma C.10 gives

$$x\left|\frac{1}{x^2} - \frac{x^{\nu+2m-2}K_{\nu+2m}(x)}{2^{\nu+2m-1}\Gamma(\nu+2m)}\right| \leq \frac{x}{4(\nu+2m)} \leq \frac{1}{4(\nu+2m)}, \qquad \text{for } 0 \leq x \leq 1.$$

We also have

$$\begin{aligned}A_m^m(\nu)x^{-\nu}K_{\nu+2m}(x)I_{(\nu+1,0,2m-1)}(x) &\leq x^{-\nu}K_{\nu+2m}(x)I_{(\nu+1,0,2m-1)}(x) && \text{since } A_m^m(\nu) \leq 1 \\ &\leq K_{\nu+2m}(x)I_{\nu+2m}(x)\prod_{k=1}^{2m-1}\frac{2\nu+2k}{2\nu+k} && \text{by (C.25)} \\ &\leq \frac{1}{2(\nu+2m)}\prod_{k=1}^{2m-1}\frac{2\nu+2k}{2\nu+k} \\ &\leq \frac{1}{2(\nu+2m-1)}\prod_{k=1}^{2m-1}\frac{2\nu+2k}{2\nu+k}\end{aligned}$$

$$< \frac{2^{n-2}}{2\nu+1},$$

where the final inequality follows from (D.12) and (D.13). Therefore, for $0 \leq x \leq 1$, we have

$$R_2 < \frac{1}{4(\nu+2m)} + \frac{2^{n-2}}{2\nu+1} < \frac{1}{2\nu+1} + \frac{2^{n-2}}{2\nu+1} = \frac{1+2^{n-2}}{2\nu+1}. \tag{D.14}$$

Summing up the bounds for R_1 and R_2 gives the result for the case of even n. The proof for odd n is similar, the only difference being that we make use of the derivative formula (C.10) for $x^{-\nu}K_\nu(x)$ rather than formula (C.9). □

We now present some series expansions for $K_\mu(x)$ and $I_{(\mu,0,n)}(x)$ which we shall use in our proofs of Lemmas D.16, D.17 and D.18 (below). The series expansion for $K_\mu(x)$ can be written as follows:

$$K_\mu(x) = \frac{1}{2}\sum_{k=0}^{\lceil\mu\rceil-1} \frac{\Gamma(\mu-k)(-1)^k}{k!}\left(\frac{x}{2}\right)^{2k-\mu} + u_\mu(x), \tag{D.15}$$

where

$$u_\mu(x) = \frac{\pi}{2\sin(\pi\mu)}\left(\sum_{k=\lceil\mu\rceil}^{\infty} \frac{1}{\Gamma(k-\mu+1)k!}\left(\frac{x}{2}\right)^{2k-\mu} - \sum_{k=0}^{\infty} \frac{1}{\Gamma(\mu+k+1)k!}\left(\frac{x}{2}\right)^{\mu+2k}\right), \quad \mu \notin \mathbb{Z},$$

and

$$u_\mu(x) = (-1)^{\mu-1}\left\{\log\left(\frac{x}{2}\right) + \gamma\right\}I_\mu(x) + \frac{(-1)^\mu}{2}\sum_{k=0}^{\infty} \frac{\{\psi(k)+\psi(\mu+k)\}}{k!\,(\mu+k)!}\left(\frac{x}{2}\right)^{\mu+2k}, \quad \mu \in \mathbb{N},$$

where $\psi(k) = \sum_{j=1}^{k}\frac{1}{j}$ and the ceiling function $\lceil x\rceil$ is the smallest integer that is not less than x. We can see that (D.15) is true for $\mu \in \mathbb{N}$ by consulting formula (B.3). To see that it holds for $\mu \notin \mathbb{Z}$, we note that the formula $\Gamma(\mu)\Gamma(1-\mu) = \frac{\pi}{\sin(\pi\mu)}$ (see, for example, Olver et al. [52]) yields

$$\frac{\pi}{\sin(\pi\mu)\Gamma(k-\mu+1)} = \frac{\sin(\pi(k-\mu+1))\Gamma(\mu-k)}{\sin(\pi\mu)} = (-1)^k\Gamma(\mu-k),$$

and substituting this formula in (B.2) shows that (D.15) holds for $\mu \notin \mathbb{Z}$.

From the series expansion (B.1) for $I_\mu(x)$ we may deduce the following series expansion for $I_{(\mu,0,n)}(x)$:

$$\begin{aligned} I_{(\mu,0,n)}(x) &= \int_0^x\int_0^{t_{n-1}}\cdots\int_0^{t_1} t_1^\mu I_\mu(t_1)\,\mathrm{d}t_1\cdots\mathrm{d}t_{n-1}\,\mathrm{d}x \\ &= \int_0^x\int_0^{t_{n-1}}\cdots\int_0^{t_1} t_1^\mu\sum_{k=0}^{\infty}\frac{1}{\Gamma(\mu+k+1)k!}\left(\frac{t_1}{2}\right)^{\mu+2k}\,\mathrm{d}t_1\cdots\mathrm{d}t_{n-1}\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{\infty} \int_0^x \int_0^{t_{n-1}} \cdots \int_0^{t_1} \frac{t_1^{2\mu+2k}}{2^{\mu+2k}\Gamma(\mu+k+1)k!}\,\mathrm{d}t_1 \cdots \mathrm{d}t_{n-1}\,\mathrm{d}x \\
&= \sum_{k=0}^{\infty} \frac{x^{\mu+2k+n}}{2^{\mu+2k}\Gamma(\mu+k+1)\prod_{j=1}^{n}(2\mu+2k+j)}, \qquad n \geq 1,
\end{aligned} \tag{D.16}
$$

where the interchange of integration and summation is easily justified by using a corollary of the monotone convergence theorem (see Theorem 17.2 of Priestley [56]).

Lemma D.16. *Let $\nu > -1/2$, then for $0 \leq x \leq 1$ we have*

$$
\left|\frac{2\nu+2}{x^2} + I_{(\nu,0,1)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \frac{25}{12(2\nu+1)} + \frac{1}{v_1(\nu)}, \tag{D.17}
$$

where

$$
v_1(\nu) = \begin{cases} 2^{2\nu+1}\nu!\,(\nu+2)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)| 2^{2\nu}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}. \end{cases}
$$

When $\nu \in \{0, 1/2, 1, 3/2, \ldots\}$ we have, for $0 \leq x \leq 1$,

$$
\left|\frac{2\nu+2}{x^2} + I_{(\nu,0,1)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \frac{3}{2\nu+1}. \tag{D.18}
$$

Proof. Using the differentiation formula (C.17) and the triangle inequality gives

$$
\left|\frac{2\nu+2}{x^2} + I_{(\nu,0,1)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| \leq R_1 + R_2,
$$

where

$$
R_1 = \left|\frac{2\nu+2}{x^2} - \frac{2\nu+1}{2(\nu+2)}\frac{K_{\nu+3}(x)I_{(\nu,0,1)}(x)}{x^\nu}\right|,
$$

and

$$
\begin{aligned}
R_2 &= \frac{3K_{\nu+1}(x)I_{(\nu,0,1)}(x)}{2(\nu+2)x^\nu} \\
&\leq \frac{3}{2(\nu+2)} \cdot \frac{2(\nu+1)}{2\nu+1} K_{\nu+1}(x)I_{\nu+1}(x) && \text{by (C.24)} \\
&\leq \frac{3}{2(\nu+2)} \cdot \frac{2(\nu+1)}{2\nu+1} \cdot \frac{1}{2(\nu+1)} && \text{by (D.1)} \\
&= \frac{3}{(2\nu+1)(2\nu+4)}.
\end{aligned}
$$

We now bound R_1. Using the series expansion (D.15) of $K_{\nu+3}(x)$ and the series expansion

(D.16) of $I_{(\nu,0,1)}(x)$ gives

$$R_1 = \bigg|\frac{2\nu+2}{x^2} - \frac{2\nu+1}{2(\nu+2)}\bigg[\frac{2^{\nu+2}\Gamma(\nu+3)}{x^{\nu+3}} - \frac{2^{\nu}\Gamma(\nu+2)}{x^{\nu+1}} + R_3\bigg] \\ \times\bigg[\frac{x^{\nu+1}}{\Gamma(\nu+1)(2\nu+1)2^{\nu}} + \frac{x^{\nu+3}}{\Gamma(\nu+2)(2\nu+3)2^{\nu+2}} + R_4\bigg]\bigg|,$$

where

$$\begin{aligned} R_3 &= \frac{1}{2}\sum_{k=2}^{\lceil\nu+3\rceil-1}\frac{\Gamma(\nu+3-k)(-1)^k}{k!}\left(\frac{x}{2}\right)^{2k-\nu-3} + u_{\nu+3}(x), \\ R_4 &= \sum_{k=2}^{\infty}\frac{x^{\nu+2k+1}}{\Gamma(\nu+k+1)k!\,(2\nu+2k+1)2^{\nu+2k}}. \end{aligned}$$

Combining terms and simplifying gives

$$\begin{aligned} R_1 &= \bigg|\left(2\nu+2-\frac{2\nu+1}{2(\nu+2)}\cdot\frac{2^{\nu+2}\Gamma(\nu+3)}{\Gamma(\nu+1)(2\nu+1)2^{\nu}}\right)\frac{1}{x^2} \\ &\quad -\frac{2\nu+1}{2(\nu+2)}\left(\frac{2^{\nu+2}\Gamma(\nu+3)}{\Gamma(\nu+2)(2\nu+3)2^{\nu+2}} - \frac{2^{\nu}\Gamma(\nu+2)}{\Gamma(\nu+1)(2\nu+1)2^{\nu}}\right) + R_5\bigg| \\ &= \left|\frac{1}{(2\nu+3)(2\nu+4)} + R_5\right| \\ &\le \frac{1}{(2\nu+3)(2\nu+4)} + |R_5|, \end{aligned}$$

where

$$R_5 = -\frac{2\nu+1}{2(\nu+2)}\bigg[R_3\frac{I_{(\nu,0,1)}(x)}{x^{\nu}} + \left(\frac{2^{\nu+2}\Gamma(\nu+3)}{x^{\nu+3}} - \frac{2^{\nu}\Gamma(\nu+2)}{x^{\nu+1}}\right)R_4 + \frac{x^2}{4(2\nu+3)}\bigg].$$

Applying the triangle inequality, and using that $\frac{2\nu+1}{2(\nu+2)} < 1$, we have

$$|R_5| < R_6 + R_7 + R_8,$$

where

$$\begin{aligned} R_6 &= \frac{|R_3|I_{(\nu,0,1)}(x)}{x^{\nu}}, \\ R_7 &= \left(\frac{2^{\nu+2}\Gamma(\nu+3)}{x^{\nu+3}} - \frac{2^{\nu}\Gamma(\nu+2)}{x^{\nu+1}}\right)|R_4| < \frac{2^{\nu+2}\Gamma(\nu+3)}{x^{\nu+3}}|R_4|, \qquad \text{for } 0 \le x \le 1, \\ R_8 &= \frac{x^2}{4(2\nu+3)} \le \frac{1}{4(2\nu+3)}, \qquad \text{for } 0 \le x \le 1. \end{aligned}$$

By the triangle inequality we have

$$R_6 \leq \frac{I_{(\nu,0,1)}(x)}{x^\nu}\bigg\{\frac{1}{2}\sum_{k=2}^{\lceil\nu+3\rceil-1}\frac{\Gamma(\nu+3-k)(-1)^k}{k!}\left(\frac{x}{2}\right)^{2k-\nu-3}+|u_{\nu+3}(x)|\bigg\}$$

Suppose $\nu\in\mathbb{N}$. Recalling the notation of Lemmas D.10, D.11 and D.12, we have

$$R_6 \leq \beta_{2,1,3}(x)+\gamma_{1,3}(x)+\delta_{1,3}(x).$$

Using the bounds that are given in Lemmas D.10, D.11 and D.12 gives, for $0\leq x\leq 1$,

$$\begin{aligned} R_6 &< \frac{1}{4(2\nu+1)}+\frac{1}{2^{2\nu+2}\nu!\,(\nu+3)!\,(2\nu+1)}+\frac{1}{2^{2\nu+2}\nu!\,(\nu+2)!\,(2\nu+1)} \\ &< \frac{1}{4(2\nu+1)}+\frac{1}{2^{2\nu+1}\nu!\,(\nu+2)!\,(2\nu+1)}. \end{aligned}$$

Suppose now that $\nu\notin\mathbb{Z}$. Then by, Lemmas D.10, D.13 and D.14, we have that for $\nu\notin\mathbb{Z}$ and $0\leq x\leq 1$,

$$R_6 \leq \beta_{2,1,3}(x)+\epsilon_{1,3}(x)+\zeta_{1,3}(x)\leq \frac{1}{4(2\nu+1)}+\frac{1}{|\sin(\pi\nu)|2^{2\nu}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1)}.$$

Let $\alpha_{2,1,3}(x)$ be defined as per Lemma D.9, then by Lemma D.9 we have

$$R_7<\alpha_{2,1,3}(x)<\frac{1}{4(2\nu+5)}, \qquad 0\leq x\leq 1.$$

Summing up the remainder terms gives

$$\begin{aligned} \left|\frac{2\nu+2}{x^2}+I_{(\nu,0,1)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| &\leq \frac{3}{(2\nu+1)(2\nu+4)}+\frac{1}{(2\nu+3)(2\nu+4)}+\frac{1}{4(2\nu+5)} \\ &\quad+\frac{1}{4(2\nu+3)}+\frac{1}{4(2\nu+1)}+\frac{1}{v_1(\nu)} \\ &<\left(1+\frac{1}{3}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}\right)\frac{1}{2\nu+1}+\frac{1}{v_1(\nu)} \\ &=\frac{25}{12(2\nu+1)}+\frac{1}{v_1(\nu)}, \end{aligned}$$

where

$$v_1(\nu)=\begin{cases} 2^{2\nu+1}\nu!\,(\nu+2)!\,(2\nu+1), & \nu\in\mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu}\Gamma(\nu+1)\Gamma(\nu+4)(2\nu+1), & \nu>-1/2 \text{ and } \nu\notin\mathbb{N}, \end{cases}$$

and to obtain the second inequality we used that $\nu>-1/2$ and thus, for example, that $2\nu+4>3$. This completes the proof of inequality (D.17).

We now prove that inequality (D.18) holds. Suppose that $\nu \in \mathbb{N}$, then

$$\frac{1}{v_1(\nu)} \le \frac{1}{2^1 \cdot 2!\,(2\nu+1)} = \frac{1}{4(2\nu+1)}.$$

Suppose now that $\nu = 1/2, 3/2, 5/2, \ldots$, then

$$\frac{1}{v_1(\nu)} \le \frac{1}{2^1\Gamma(3/2)\Gamma(9/2)(2\nu+1)} = \frac{16}{105\pi(2\nu+1)} < \frac{1}{4(2\nu+1)},$$

where we used that $\Gamma(3/2) = \frac{\sqrt{\pi}}{2}$ and $\Gamma(9/2) = \frac{105\sqrt{\pi}}{16}$. Therefore, for $\nu = 0, 1/2, 1, 3/2, \ldots$ we have

$$\left|\frac{2\nu+2}{x^2} + I_{(\nu,0,1)}(x)\frac{\mathrm{d}^3}{\mathrm{d}x^3}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \left(\frac{25}{12} + \frac{1}{4}\right)\frac{1}{2\nu+1} = \frac{7}{3(2\nu+1)} < \frac{3}{2\nu+1},$$

as required. $\square$

Lemma D.17. *Let $\nu > -1/2$, then for $0 \le x \le 1$ we have*

$$\left|\frac{2\nu+3}{x^2} - I_{(\nu,0,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \frac{77}{20(2\nu+1)} + \frac{1}{v_2(\nu)}, \tag{D.19}$$

where

$$v_2(\nu) = \begin{cases} 2^{2\nu+3}(\nu+1)!\,(\nu+3)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu+2}\Gamma(\nu+2)\Gamma(\nu+5)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}. \end{cases}$$

When $\nu \in \{0, 1/2, 1, 3/2, \ldots\}$ we have, for $0 \le x \le 1$,

$$\left|\frac{2\nu+3}{x^2} - I_{(\nu,0,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \frac{4}{2\nu+1}. \tag{D.20}$$

Proof. Using the differentiation formula (C.18) and the triangle inequality gives

$$\left|\frac{2\nu+3}{x^2} - I_{(\nu,0,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| \le R_1 + R_2$$

where

$$R_1 = (2\nu+3)\left|\frac{1}{x^2} - \frac{2\nu+1}{4(\nu+2)(\nu+3)}\frac{K_{\nu+4}(x)I_{(\nu,0,2)}(x)}{x^\nu}\right|$$

and

$$
\begin{aligned}
R_2 &= \left(\frac{3(2\nu+1)K_{\nu+2}(x)}{2(\nu+1)(\nu+3)x^\nu} + \frac{3K_\nu(x)}{4(\nu+1)(\nu+2)x^\nu}\right) I_{(\nu,0,2)}(x) \\
&< \left\{\prod_{k=1}^{3} \frac{2\nu+2k}{2\nu+k}\right\} K_{\nu+2}(x) I_{\nu+2}(x) && \text{by (C.24) and (B.16)} \\
&\le \frac{(2\nu+2)(2\nu+4)(2\nu+6)}{(2\nu+1)(2\nu+2)(2\nu+3)} \cdot \frac{1}{2(\nu+2)} && \text{by (D.1)} \\
&= \frac{2\nu+6}{(2\nu+1)(2\nu+3)} \\
&< \frac{5}{2(2\nu+1)},
\end{aligned}
$$

where the final inequality holds because for $\nu > -1/2$ we have $\frac{2\nu+6}{2\nu+3} < \frac{5}{2}$. Using the series expansion (D.15) for $K_{\nu+4}(x)$ and the series expansion (D.16) for $I_{(\nu,0,2)}(x)$ gives

$$
\begin{aligned}
R_1 = (2\nu+3)\Bigg| \frac{1}{x^2} - \frac{2\nu+1}{4(\nu+2)(\nu+3)} &\left[\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}} + R_3\right] \\
\times &\left[\frac{x^{\nu+2}}{\Gamma(\nu+1)(2\nu+1)(2\nu+2)2^\nu} + \frac{x^{\nu+4}}{\Gamma(\nu+2)(2\nu+3)(2\nu+4)2^{\nu+2}} + R_4\right]\Bigg|,
\end{aligned}
$$

where

$$
\begin{aligned}
R_3 &= \frac{1}{2}\sum_{k=2}^{\lceil \nu+4\rceil-1} \frac{\Gamma(\nu+4-k)(-1)^k}{k!}\left(\frac{x}{2}\right)^{2k-\nu-4} + u_{\nu+4}(x), \\
R_4 &= \sum_{k=2}^{\infty} \frac{x^{\nu+2k+2}}{\Gamma(\nu+k+1)k!\,(2\nu+2k+1)(2\nu+2k+2)2^{\nu+2k}}.
\end{aligned}
$$

Combining terms and simplifying gives

$$
\begin{aligned}
R_1 &= (2\nu+3)\Bigg|\left(1 - \frac{2\nu+1}{4(\nu+2)(\nu+3)} \cdot \frac{2^{\nu+3}\Gamma(\nu+4)}{\Gamma(\nu+1)(2\nu+1)(2\nu+2)2^\nu}\right)\frac{1}{x^2} \\
&\quad - \frac{2\nu+1}{4(\nu+2)(\nu+3)}\left(\frac{2^{\nu+3}\Gamma(\nu+4)}{\Gamma(\nu+2)(2\nu+3)(2\nu+4)2^{\nu+2}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{\Gamma(\nu+1)(2\nu+1)(2\nu+2)2^\nu}\right) + R_5\Bigg| \\
&= \left|\frac{3}{(2\nu+3)(2\nu+4)} + R_5\right| \\
&\le \frac{3}{(2\nu+3)(2\nu+4)} + |R_5|,
\end{aligned}
$$

where

$$
R_5 = -\frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)}\left[R_3\frac{I_{(\nu,0,2)}(x)}{x^\nu} + \left(\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}}\right)R_4 + \frac{x^2}{4(2\nu+3)}\right].
$$

Applying the triangle inequality, and using that $\frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)} < 1$, we have

$$|R_5| < R_6 + R_7 + R_8,$$

where

$$\begin{aligned} R_6 &= \frac{|R_3| I_{(\nu,0,2)}(x)}{x^\nu}, \\ R_7 &= \left(\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}}\right)|R_4| < \frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}}|R_4|, \qquad \text{for } 0 \le x \le 1, \\ R_8 &= \frac{x^2}{4(2\nu+3)} \le \frac{1}{4(2\nu+3)}, \qquad \text{for } 0 \le x \le 1. \end{aligned}$$

We bound R_6 and R_7 by using the same approach as for Lemma D.16. By Lemmas D.10, D.11 and D.12, we have that for $\nu \in \mathbb{N}$ and $0 \le x \le 1$,

$$\begin{aligned} R_6 &\le \beta_{2,2,4}(x) + \gamma_{2,4}(x) + \delta_{2,4}(x) \\ &< \frac{1}{4(2\nu+1)} + \frac{1}{2^{2\nu+4}(\nu+1)!\,(\nu+4)!\,(2\nu+1)} + \frac{1}{2^{2\nu+4}(\nu+1)!\,(\nu+3)!\,(2\nu+1)} \\ &< \frac{1}{4(2\nu+1)} + \frac{1}{2^{2\nu+3}(\nu+1)!\,(\nu+3)!\,(2\nu+1)}. \end{aligned}$$

By Lemmas D.10, D.13 and D.14, we have that for $\nu \notin \mathbb{Z}$ and $0 \le x \le 1$,

$$R_6 \le \beta_{2,2,4}(x) + \epsilon_{2,4}(x) + \zeta_{2,4}(x) < \frac{1}{4(2\nu+1)} + \frac{1}{|\sin(\pi\nu)| 2^{2\nu+2}\Gamma(\nu+2)\Gamma(\nu+5)(2\nu+1)}.$$

We use Lemma D.9 to bound R_7:

$$R_7 \le \alpha_{2,2,4}(x) < \frac{1}{4(2\nu+5)}, \qquad 0 \le x \le 1.$$

Summing up the remainder terms gives

$$\begin{aligned} \left|\frac{2\nu+3}{x^2} - I_{(\nu,0,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| &\le \frac{5}{2(2\nu+1)} + \frac{3}{(2\nu+3)(2\nu+4)} + \frac{1}{4(2\nu+5)} \\ &\quad + \frac{1}{4(2\nu+3)} + \frac{1}{4(2\nu+1)} + \frac{1}{v_2(\nu)} \\ &< \left(\frac{5}{2} + \frac{3}{5} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4}\right)\frac{1}{2\nu+1} + \frac{1}{v_2(\nu)} \\ &= \frac{77}{20(2\nu+1)} + \frac{1}{v_2(\nu)}, \end{aligned}$$

where

$$v_2(\nu) = \begin{cases} 2^{2\nu+3}(\nu+1)!\,(\nu+3)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu+2}\Gamma(\nu+2)\Gamma(\nu+5)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}, \end{cases}$$

and to obtain the second inequality we used that $\nu > -1/2$ and thus, for example, that $2\nu+4 > 3$. This completes the proof of inequality (D.19).

We now prove that inequality (D.20) holds. Suppose that $\nu \in \mathbb{N}$, then

$$\frac{1}{v_2(\nu)} \leq \frac{1}{2^3 \cdot 3!\,(2\nu+1)} = \frac{1}{48(2\nu+1)}.$$

Suppose now that $\nu = 1/2, 3/2, 5/2, \ldots$, then

$$\frac{1}{v_2(\nu)} \leq \frac{1}{2^3\Gamma(5/2)\Gamma(11/2)(2\nu+1)} = \frac{16}{2835\pi(2\nu+1)} < \frac{1}{48(2\nu+1)}.$$

Therefore, for $\nu = 0, 1/2, 1, 3/2, \ldots$ we have

$$\left|\frac{2\nu+3}{x^2} - I_{(\nu,0,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \left(\frac{77}{20} + \frac{1}{48}\right)\frac{1}{2\nu+1} = \frac{929}{240(2\nu+1)} < \frac{4}{2\nu+1},$$

as required. □

Lemma D.18. *Let $\nu > -1/2$, then for $0 \leq x \leq 1$ we have*

$$\left|\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x} - I_{(\nu,0,1)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \frac{2779}{768(2\nu+1)} + \frac{1}{v_3(\nu)}, \tag{D.21}$$

where

$$v_3(\nu) = \begin{cases} 2^{2\nu+2}\nu!\,(\nu+3)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu+1}\Gamma(\nu+1)\Gamma(\nu+5)(2\nu+1), & \nu > -1/2 \textit{ and } \nu \notin \mathbb{N}. \end{cases}$$

When $\nu \in \{0, 1/2, 1, 3/2, \ldots\}$ we have, for $0 \leq x \leq 1$,

$$\left|\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x} - I_{(\nu,0,1)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \frac{4}{2\nu+1}. \tag{D.22}$$

Proof. Using the differentiation formula (C.18) and the triangle inequality gives

$$\left|\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x} - I_{(\nu,0,1)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| \leq R_1 + R_2$$

where

$$R_1 = \bigg| \frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x} - \frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)} \frac{K_{\nu+4}(x) I_{(\nu,0,1)}(x)}{x^\nu} - \frac{3(2\nu+1)}{2(\nu+1)(\nu+3)} \frac{K_{\nu+2}(x) I_{(\nu,0,1)}(x)}{x^\nu} \bigg|$$

and

$$\begin{aligned} R_2 &= \frac{3K_\nu(x) I_{(\nu,0,1)}(x)}{4(\nu+1)(\nu+2)x^\nu} \\ &\le \frac{3}{4(\nu+1)(\nu+2)} \cdot \frac{2\nu+2}{2\nu+1} K_\nu(x) I_{\nu+1}(x) && \text{by (C.24)} \\ &\le \frac{3}{4(\nu+1)(\nu+2)} \cdot \frac{2\nu+2}{2\nu+1} \cdot \frac{1}{2(\nu+1)} && \text{by (B.16) and (D.1)} \\ &= \frac{3}{(2\nu+1)(2\nu+2)(2\nu+4)}. \end{aligned}$$

We now bound R_1. By the triangle inequality

$$R_1 \le R_3 + R_4,$$

where

$$\begin{aligned} R_3 &= \frac{3}{\nu+3} \bigg| \frac{1}{x} - \frac{2\nu+1}{2(\nu+1)} \frac{K_{\nu+2}(x) I_{(\nu,0,1)}(x)}{x^\nu} \bigg|, \\ R_4 &= \bigg| \frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{\nu}{(\nu+3)x} - \frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)} \frac{K_{\nu+4}(x) I_{(\nu,0,1)}(x)}{x^\nu} \bigg|. \end{aligned}$$

In the proof of Lemma D.15 (see inequality (D.14) we showed that:

$$\bigg| \frac{1}{x} - A_m^m(\nu) \frac{K_{\nu+2m}(x) I_{(\nu,0,2m-1)}(x)}{x^\nu} \bigg| < \frac{1+2^{2m-2}}{2\nu+1}, \qquad \text{for } 0 \le x \le 1.$$

Since $A_1^1(\nu) = \frac{2\nu+1}{2(\nu+1)}$, we have the following bound on R_3:

$$R_3 \le \frac{3}{\nu+3} \cdot \frac{2}{2\nu+1} = \frac{6}{(2\nu+1)(\nu+3)}, \qquad \text{for } 0 \le x \le 1.$$

Using the series expansion (D.15) for $K_{\nu+4}(x)$ and the series expansion (D.16) for $I_{(\nu,0,1)}(x)$

gives

$$R_4 = \Bigg|\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{\nu}{(\nu+3)x} - \frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)}\bigg[\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}} + \frac{2^{\nu-2}\Gamma(\nu+2)}{x^{\nu}} + R_5\bigg] \times \bigg[\frac{x^{\nu+1}}{\Gamma(\nu+1)(2\nu+1)2^{\nu}} + \frac{x^{\nu+3}}{\Gamma(\nu+2)(2\nu+3)2^{\nu+2}} + \frac{x^{\nu+5}}{\Gamma(\nu+3)(2\nu+5)2^{\nu+5}} + R_6\bigg]\Bigg|,$$

where

$$\begin{aligned} R_5 &= \frac{1}{2}\sum_{k=2}^{\lceil\nu+4\rceil-1}\frac{\Gamma(\nu+4-k)(-1)^k}{k!}\left(\frac{x}{2}\right)^{2k-\nu-4} + u_{\nu+4}(x), \\ R_6 &= \sum_{k=3}^{\infty}\frac{x^{\nu+2k+1}}{\Gamma(\nu+k+1)k!\,(2\nu+2k+1)2^{\nu+2k}}. \end{aligned}$$

Combining terms and simplifying gives

$$\begin{aligned} R_4 &= \Bigg|\left((2\nu+2)(2\nu+3) - \frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)}\cdot\frac{2^{\nu+3}\Gamma(\nu+4)}{\Gamma(\nu+1)(2\nu+1)2^{\nu}}\right)\frac{1}{x^3} \\ &\quad + \left(\frac{\nu}{\nu+3} - \frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)}\left(\frac{2^{\nu+3}\Gamma(\nu+4)}{\Gamma(\nu+2)(2\nu+3)2^{\nu+2}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{\Gamma(\nu+1)(2\nu+1)2^{\nu}}\right)\right)\frac{1}{x} + R_5\Bigg| \\ &= \left|\frac{x}{(2\nu+1)(2\nu+3)(2\nu+5)} + R_7\right| \\ &\le \frac{1}{(2\nu+1)(2\nu+3)(2\nu+5)} + |R_7|, \end{aligned}$$

where

$$\begin{aligned} R_7 = -\frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)}\bigg[&R_5\frac{I_{(\nu,0,1)}(x)}{x^{\nu}} + \left(\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}} + \frac{2^{\nu-2}\Gamma(\nu+2)}{x^{\nu}}\right)R_6 \\ &+ \left(\frac{2^{\nu-2}\Gamma(\nu+2)}{\Gamma(\nu+2)(2\nu+3)2^{\nu+2}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{\Gamma(\nu+3)(2\nu+5)2^{\nu+5}}\right)x^3 + \frac{2^{\nu-2}(\nu+1)!\,x^5}{\Gamma(\nu+3)(2\nu+5)2^{\nu+5}}\bigg], \end{aligned}$$

and applying the triangle inequality, and using that $\frac{(2\nu+1)(2\nu+3)}{4(\nu+2)(\nu+3)} < 1$, we have

$$|R_7| < R_8 + R_9 + R_{10},$$

where

$$\begin{aligned} R_8 &= \frac{|R_5|I_{(\nu,0,1)}(x)}{x^{\nu}}, \\ R_9 &= \left(\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}} + \frac{2^{\nu-2}\Gamma(\nu+2)}{x^{\nu}}\right)|R_6|, \end{aligned}$$

and

$$R_{10} = \frac{x^3}{8(2\nu+3)(2\nu+5)} + \frac{2^{\nu-2}\Gamma(\nu+2)x^5}{\Gamma(\nu+3)(2\nu+5)2^{\nu+5}}.$$

We can bound R_9 and R_{10}, for $0 \le x \le 1$, as follows

$$\begin{aligned} R_9 &< \left(\frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}} - \frac{2^{\nu+1}\Gamma(\nu+3)}{x^{\nu+2}} + \frac{2^{\nu-2}\Gamma(\nu+2)}{x^{\nu}}\right)|R_6| < \frac{2^{\nu+3}\Gamma(\nu+4)}{x^{\nu+4}}|R_6|, \\ R_{10} &< \left(\frac{1}{8} + \frac{1}{64}\right)\frac{1}{(2\nu+3)(2\nu+5)} = \frac{9}{64(2\nu+3)(2\nu+5)}. \end{aligned}$$

We bound R_6 and R_7 by using the same approach as for Lemmas D.16 and D.17. By Lemmas D.10, D.11 and D.12, we have that for $\nu \in \mathbb{N}$ and $0 \le x \le 1$

$$\begin{aligned} R_8 &\le \beta_{3,1,4}(x) + \gamma_{1,4}(x) + \delta_{1,4}(x) \\ &< \frac{1}{24(2\nu+1)} + \frac{1}{2^{2\nu+3}\nu!\,(\nu+4)!\,(2\nu+1)} + \frac{1}{2^{2\nu+3}\nu!\,(\nu+3)!\,(2\nu+1)} \\ &< \frac{1}{24(2\nu+1)} + \frac{1}{2^{2\nu+2}\nu!\,(\nu+3)!\,(2\nu+1)}, \end{aligned}$$

and by Lemmas D.10, D.13 and D.14 we have that, for $\nu \notin \mathbb{Z}$ and $0 \le x \le 1$,

$$R_8 \le \beta_{3,1,4}(x) + \epsilon_{1,4}(x) + \zeta_{1,4}(x) < \frac{1}{24(2\nu+1)} + \frac{1}{|\sin(\pi\nu)|2^{2\nu+1}\Gamma(\nu+1)\Gamma(\nu+5)(2\nu+1)}.$$

We use Lemma D.9 to bound R_7:

$$R_9 \le \alpha_{3,1,4}(x) < \frac{1}{24(2\nu+7)}, \qquad 0 \le x \le 1.$$

Summing up the remainder terms gives

$$\begin{aligned} &\left|\frac{(2\nu+2)(2\nu+3)}{x^3} + \frac{1}{x} - I_{(\nu,0,1)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| \\ &\le \frac{3}{2(2\nu+1)(2\nu+2)(2\nu+4)} + \frac{6}{(2\nu+1)(2\nu+3)} + \frac{1}{24(2\nu+7)} + \frac{9}{64(2\nu+3)(2\nu+4)} \\ &\quad + \frac{1}{24(2\nu+1)} + \frac{1}{v_3(\nu)} \\ &< \left(\frac{1}{2} + 3 + \frac{1}{24} + \frac{9}{256} + \frac{1}{24}\right)\frac{1}{2\nu+1} + \frac{1}{v_3(\nu)} \\ &= \frac{2779}{768(2\nu+1)} + \frac{1}{v_3(\nu)}, \end{aligned}$$

where

$$v_3(\nu) = \begin{cases} 2^{2\nu+2}\nu!\,(\nu+3)!\,(2\nu+1), & \nu \in \mathbb{N}, \\ |\sin(\pi\nu)|2^{2\nu+1}\Gamma(\nu+1)\Gamma(\nu+5)(2\nu+1), & \nu > -1/2 \text{ and } \nu \notin \mathbb{N}, \end{cases}$$

and to obtain the second inequality we used that $\nu > -1/2$ and thus, for example, that $2\nu+4 > 3$. This completes the proof of inequality (D.21).

We now prove that inequality (D.22) holds. Suppose that $\nu \in \mathbb{N}$, then

$$\frac{1}{v_3(\nu)} \leq \frac{1}{2^2 \cdot 3!\,(2\nu+1)} = \frac{1}{24(2\nu+1)}.$$

Suppose now that $\nu = 1/2, 3/2, 5/2, \ldots$, then

$$\frac{1}{v_3(\nu)} \leq \frac{1}{2^2\Gamma(3/2)\Gamma(11/2)(2\nu+1)} = \frac{16}{945\pi(2\nu+1)} < \frac{1}{24(2\nu+1)}.$$

Therefore, for $\nu = 0, 1/2, 1, 3/2, \ldots$ we have

$$\left|\frac{2\nu+3}{x^2} - I_{(\nu,0,2)}(x)\frac{\mathrm{d}^4}{\mathrm{d}x^4}\left(\frac{K_\nu(x)}{x^\nu}\right)\right| < \left(\frac{2779}{768} + \frac{1}{24}\right)\frac{1}{2\nu+1} = \frac{937}{256(2\nu+1)} < \frac{4}{2\nu+1},$$

as required. $\square$

Remark D.19. The bounds of Lemmas D.16, D.17 and D.18 perform poorly when 2ν is very close but not equal to an integer, due to the presence of a term involving $1/\sin(\pi\nu)$. This term is an artefact of our series expansion method and we consider that an alternative approach should yield bounds of the form $C(2\nu+1)^{-1}$ for all $\nu > -1/2$, where C is a constant not involving ν.

It would also be of interest for future research to establish $O(\nu^{-1})$ bounds for all $x \in [0,\infty)$ for each of the expressions in Lemmas D.15, D.16, D.17 and D.18. Whilst our series expansion method that we used in our proofs leads to good bounds in the region $0 \leq x \leq 1$, it would give poor bounds for large x, and so a different approach would be needed to obtain $O(\nu^{-1})$ bounds for large x.

For the case $\nu = 0$ we can actually use a relatively straightforward approach to achieve good bounds for all $x \in [0,\infty)$, as we shall see in the following lemma. The approach could be easily extended to general $\nu > -1/2$, but would lead to poor bounds for large ν.

Lemma D.20. *For $x \geq 0$ we have*

$$\begin{aligned}
\left\| \frac{1}{x} - I_{(0,0,1)}(x) K_0''(x) \right\| &< 3, && \text{(D.23)} \\
\left\| \frac{1}{x} + I_{(0,0,2)}(x) K_0^{(3)}(x) \right\| &< 5, && \text{(D.24)} \\
\left\| \frac{1}{x} - I_{(0,0,3)}(x) K_0^{(4)}(x) \right\| &< 9, && \text{(D.25)} \\
\left\| \frac{2}{x^2} + I_{(0,0,1)}(x) K_0^{(3)}(x) \right\| &< 4.39, && \text{(D.26)} \\
\left\| \frac{3}{x^2} - I_{(0,0,2)}(x) K_0^{(4)}(x) \right\| &< 6.81, && \text{(D.27)} \\
\left\| \frac{6}{x^3} + \frac{1}{x} - I_{(0,0,1)}(x) K_0^{(4)}(x) \right\| &< 14.61. && \text{(D.28)}
\end{aligned}$$

Proof. In Lemmas (D.15)-(D.18) we established bounds for these expressions in the region $0 \leq x \leq 1$:

$$\begin{aligned}
\left| \frac{1}{x} - I_{(0,0,1)}(x) K_0''(x) \right| &< 3, && \text{(D.29)} \\
\left| \frac{1}{x} + I_{(0,0,2)}(x) K_0^{(3)}(x) \right| &< 5, && \text{(D.30)} \\
\left| \frac{1}{x} - I_{(0,0,3)}(x) K_0^{(4)}(x) \right| &< 9, && \text{(D.31)} \\
\left| \frac{2}{x^2} + I_{(0,0,1)}(x) K_0^{(3)}(x) \right| &< 3, && \text{(D.32)} \\
\left| \frac{3}{x^2} - I_{(0,0,2)}(x) K_0^{(4)}(x) \right| &< 4, && \text{(D.33)} \\
\left| \frac{6}{x^3} + \frac{1}{x} - I_{(0,0,1)}(x) K_0^{(4)}(x) \right| &< 4. && \text{(D.34)}
\end{aligned}$$

We now obtain bounds in the region $x \geq 1$, and combining these bounds will give us bounds in the region $0 \leq x \leq \infty$.

(i) For $x \geq 1$ we have

$$\begin{aligned}
\left| \frac{1}{x} - I_{(0,0,1)}(x) K_0''(x) \right| &\leq 1 + \left| I_{(0,0,1)}(x) K_0''(x) \right| \\
&= 1 + \frac{1}{2} I_{(0,0,1)}(x) [K_2(x) + K_0(x)] && \text{by (C.16)} \\
&< 1 + I_1(x) [K_2(x) + K_0(x)] && \text{by (C.25)} \\
&< 1 + I_1(x) [K_2(x) + K_1(x)], && \text{by (B.16)} \qquad \text{(D.35)}
\end{aligned}$$

Now, Theorem 1.1 of Laforgia and Natalini [40] states that, for $x > 0$ and $\mu \geq 0$,

$$\frac{I_{\mu-1}(x)}{I_\mu(x)} < \frac{x}{-\mu + \sqrt{\mu^2 + x^2}} = \frac{\mu + \sqrt{\mu^2 + x^2}}{x}.$$

Straightforward calculus shows that for $\mu \geq 0$ the function $\frac{\mu+\sqrt{\mu^2+x^2}}{x}$ is strictly monotone decreasing in the region $(0, \infty)$. Hence, in the region $x \geq 1$ the following inequality holds

$$I_{\mu-1}(x) < (\mu + \sqrt{\mu^2 + 1})I_\mu(x), \qquad \mu \geq 0. \tag{D.36}$$

Applying inequality (D.36) to (D.35) gives, for $x \geq 1$,

$$\begin{aligned} \left|\frac{1}{x} - I_{(0,0,1)}(x)K_0''(x)\right| &< 1 + (2 + \sqrt{5})I_2(x)K_2(x) + I_1(x)K_1(x) \\ &< 1 + (2 + \sqrt{5})I_2(1)K_2(1) + I_1(1)K_1(1) \\ &= 2.274\ldots, \end{aligned} \tag{D.37}$$

where we used that $I_\nu(x)K_\nu(x)$ is a monotone decreasing function of x in $(0, \infty)$ (see Proposition D.1) to obtain the second inequality, and the values of $I_1(1)$, $I_2(1)$, $K_1(1)$ and $K_2(1)$ were calculated using Table 9.8 of Abramowitz and Stegun [1]. Combining inequalities (D.29) and (D.37) yields (D.23).

(ii) The proof is similar to that of inequality (D.23). For $x \geq 1$ we have

$$\begin{aligned} \left|\frac{1}{x} + I_{(0,0,2)}(x)K_0^{(3)}(x)\right| &\leq 1 + \left|I_{(0,0,2)}(x)K_0^{(3)}(x)\right| \\ &= 1 + \frac{1}{4}I_{(0,0,2)}(x)[K_3(x) + 3K_1(x)] && \text{by (C.17)} \\ &< 1 + I_2(x)[K_3(x) + 3K_1(x)] && \text{by (C.25)} \\ &< 1 + I_2(x)[K_3(x) + 3K_2(x)], && \text{by (B.16)} \\ &< 1 + (3 + \sqrt{10})I_3(x)K_3(x) + 3I_2(x)K_2(x) && \text{by (D.36)} \\ &< 1 + (3 + \sqrt{10})I_3(1)K_3(1) + I_2(1)K_2(1) \\ &= 2.631\ldots. \end{aligned} \tag{D.38}$$

Combining inequalities (D.30) and (D.38) yields (D.24).

(iii) For $x \geq 1$ we have

$$\left|\frac{1}{x} - I_{(0,0,3)}(x)K_0^{(4)}(x)\right| \leq 1 + \left|I_{(0,0,3)}(x)K_0^{(4)}(x)\right|$$

$$\begin{aligned}
&= 1 + \frac{1}{8} I_{(0,0,3)}(x)[K_4(x) + 4K_2(x) + 3K_0(x)] && \text{by (C.18)} \\
&< 1 + I_3(x)[K_4(x) + 4K_2 + 3K_0(x)] && \text{by (C.25)} \\
&< 1 + I_3(x)[K_4(x) + 7K_3(x)], && \text{by (B.16)} \\
&< 1 + (4 + \sqrt{17}) I_4(x) K_4(x) + 7 I_3(x) K_3(x) && \text{by (D.36)} \\
&< 1 + (4 + \sqrt{17}) I_4(1) K_4(1) + 7 I_3(1) K_3(1) \\
&= 3.085 \ldots. && \text{(D.39)}
\end{aligned}$$

Combining inequalities (D.31) and (D.39) yields (D.25).

(iv) For $x \geq 1$ we have

$$\begin{aligned}
\left| \frac{2}{x^2} + I_{(0,0,1)}(x) K_0^{(3)}(x) \right| &\leq 2 + \left| I_{(0,0,1)}(x) K_0^{(3)}(x) \right| \\
&= 2 + \frac{1}{4} I_{(0,0,1)}(x)[K_3(x) + 3K_1(x)] && \text{by (C.17)} \\
&< 2 + \frac{1}{2} I_1(x)[K_3(x) + 3K_1(x)] && \text{by (C.25)} \\
&< 2 + \frac{1}{2}(2 + \sqrt{5}) I_2(x) K_3(x) + \frac{3}{2} I_1(x) K_1(x) && \text{by (D.36)} \\
&< 2 + \frac{1}{2}(2 + \sqrt{5})(3 + \sqrt{10}) I_3(x) K_3(x) + \frac{3}{2} I_1(x) K_1(x) && \text{by (D.36)} \\
&< 2 + \frac{1}{2}(2 + \sqrt{5})(3 + \sqrt{10}) I_3(1) K_3(1) + \frac{3}{2} I_1(1) K_1(1) \\
&= 4.385 \ldots. && \text{(D.40)}
\end{aligned}$$

Combining inequalities (D.32) and (D.40) yields (D.26).

(v) For $x \geq 1$ we have

$$\begin{aligned}
\left| \frac{3}{x^2} - I_{(0,0,2)}(x) K_0^{(4)}(x) \right| &\leq 3 + \left| I_{(0,0,2)}(x) K_0^{(4)}(x) \right| \\
&= 3 + \frac{1}{8} I_{(0,0,2)}(x)[K_4(x) + 4K_2(x) + 3K_0(x)] && \text{by (C.18)} \\
&< 3 + \frac{1}{2} I_2(x)[K_4(x) + 4K_2(x) + 3K_0(x)] && \text{by (C.25)} \\
&< 3 + \frac{1}{2} I_2(x)[K_4(x) + 7K_2(x)], && \text{by (B.16)} \\
&< 3 + \frac{1}{2}(3 + \sqrt{10})(4 + \sqrt{17}) I_4(x) K_4(x) + \frac{7}{2} I_2(x) K_2(x) && \text{by (D.36)} \\
&< 3 + \frac{1}{2}(3 + \sqrt{10})(4 + \sqrt{17}) I_4(1) K_4(1) + \frac{7}{2} I_2(1) K_2(1) \\
&= 6.802 \ldots. && \text{(D.41)}
\end{aligned}$$

Combining inequalities (D.33) and (D.41) yields (D.27).

(vi) For $x \geq 1$ we have

$$\begin{aligned}
\left|\frac{6}{x^3} + \frac{1}{x} - I_{(0,0,1)}(x)K_0^{(4)}(x)\right| &\leq 7 + \left|I_{(0,0,1)}(x)K_0^{(4)}(x)\right| \\
&= 7 + \frac{1}{8}I_{(0,0,1)}(x)[K_4(x) + 4K_2(x) + 3K_0(x)] && \text{by (C.18)} \\
&< 7 + \frac{1}{4}I_1(x)[K_4(x) + 4K_2(x) + 3K_0(x)] && \text{by (C.25)} \\
&< 7 + \frac{1}{4}I_1(x)[K_4(x) + 4K_2(x) + 3K_1(x)], && \text{by (B.16)} \\
&< 7 + \frac{1}{4}(2+\sqrt{5})(3+\sqrt{10})(4+\sqrt{17})I_4(x)K_4(x) \\
&\quad + (2+\sqrt{5})I_2(x)K_2(x) + \frac{3}{4}I_1(x)K_1(x) && \text{by (D.36)} \\
&< 7 + \frac{1}{4}(2+\sqrt{5})(3+\sqrt{10})(4+\sqrt{17})I_4(1)K_4(1) \\
&\quad + (2+\sqrt{5})I_2(1)K_2(1) + \frac{3}{4}I_1(1)K_1(1) \\
&= 14.607\ldots. && \text{(D.42)}
\end{aligned}$$

Combining inequalities (D.34) and (D.42) yields (D.28). □

Index of Notation

$\mathcal{A}$	Infinitesimal generator of a stochastic process
B_t	Standard Brownian motion
$\mathcal{B}$	Borel σ–algebra defined on $\mathbb{R}$
$C^i(\mathbb{R})$	Space of all real–valued i–times continuously differentiable functions on $\mathbb{R}$
$C_b(\mathbb{R})$	Space of all bounded real–valued continuous functions
$C_b^i(\mathbb{R})$	Subspace of $C^i(\mathbb{R})$ containing absolutely bounded functions with absolutely bounded k–th order derivatives for $k \leq i$
$\mathbf{C}_b^i(\mathbb{R})$	Subspace of $C^i(\mathbb{R})$ containing functions with $\|f^{(k)}\| \leq M_f$ for $k \leq i$ and $M_f \in \mathbb{R}_+$
C_c	Space of all continuous functions with compact support
$\mathcal{C}_\lambda$	Space of real-valued functions on $\mathbb{R}_+$ dominated by ce^{ax} ($c > 0$, $a < \lambda$)
$d_K(\mu_X, \mu_Y)$	Kolmogorov-Smirnov distance between probability measures μ_X and μ_Y
$d_{TV}(\mu_X, \mu_Y)$	Total Variational distance between probability measures μ_X and μ_Y
$d_W(\mu_X, \mu_Y)$	Wasserstein distance between probability measures μ_X and μ_Y
$I_\nu(x)$	Modified Bessel function of the first kind of order ν
$I_{(\nu,\beta,n)}(x)$	$I_{(\nu,\beta,0)}(x) = \mathrm{e}^{\beta x} x^\nu I_\nu(x)$ and $I_{(\nu,\beta,n+1)}(x) = \int_0^x I_{(\nu,\beta,n)}(y)\,\mathrm{d}y$ for $n \geq 0$
I_n	$n \times n$ identity matrix
i.i.d.	Independent and identically distributed
$K_\nu(x)$	Modified Bessel function of the second kind of order ν
$K_{\nu,\beta}(\,\cdot\,)$	Distribution function for a $\mathrm{VG}_2(\nu, 1, \beta, 0)$ distributed random variable
$\mathscr{L}(\,\cdot\,)$	Law of a random variable
$\mathscr{L}_\nu(x)$	Denotes $I_\nu(x)$, $\mathrm{e}^{\nu\pi i} K_\nu(x)$ or any linear combination of these functions in which the coefficients are independent of ν and x
$\log(x)$	The logarithm to base e
$\mathbb{N}$	Set of natural numbers, $\{0, 1, 2, \ldots\}$
$o(f(n))$	$g(n)$ is so if and only if $\frac{g(n)}{f(n)} \to 0$ as $n \to \infty$
$O(f(n))$	$g(n)$ is so if and only if $\exists\, c, N > 0$ s.t. $n \geq N \Rightarrow g(n) \leq cf(n)$

p.d.f.	Probability density function
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of non-negative real numbers
$\mathbb{R}^n$	n–dimensional Euclidean space
$T_t f(x)$	Transition semigroup: $\mathbb{E}\{f(X_t)\,\|\,X_0 = x\}$ for X_t a Markov process
$\tilde{\mathrm{VG}}^{\nu,\alpha}_{\beta,\mu}h$	$\mathbb{E}h(X)$, where $X \sim \mathrm{VG}_2(\nu,\alpha,\beta,\mu)$
$\mathrm{VG}^{r,\theta,}_{\sigma,\mu}h$	$\mathbb{E}h(X)$, where $X \sim \mathrm{VG}(r,\theta,\sigma,,\mu)$
W^*	Random variable with the W-zero bias distribution
$W^{*(r)}$	Random variable with the W-zero bias distribution or order r
W^s	Random variable with the W-size bias distribution
$W^\square$	Random variable with the W-square bias distribution
W'	Random variable for which (W, W') is an exchangeable pair
$\Gamma_{\alpha,\beta}h$	$\mathbb{E}h(X)$, where $X \sim \Gamma(\alpha,\beta)$
$\gamma(\,\cdot\,;\alpha,\beta)$	Gamma density function with shape α and scale β
$\hat{\gamma}(\,\cdot\,;\alpha,\beta,\theta)$	Non-central Gamma density function with shape α, scale β and non-centrality θ
$\Phi(\,\cdot\,)$	Standard normal distribution function
Φh	$\mathbb{E}h(X)$, where $X \sim N(0,1)$
$\chi_I(x)$	Indicator function of an interval $I \subset \mathbb{R}$: $\chi_I(x) = 1$ if $x \in I$, otherwise $\chi_I(x) = 0$
$\chi^2_{(p)}h$	$\mathbb{E}h(X)$, where $X \sim \chi^2_{(p)}$
$\mathbb{Z}^+$	Set of positive integers, $\{1,2,3,\ldots\}$
$\mathbf{1}[A]$	Indicator function of an event A
M^T	Transpose of a matrix M
$\lceil x \rceil$	Ceiling function : $\lceil x \rceil = \min\{n \in \mathbb{Z}\,\|\,n \geq x\}$
$\perp\!\!\!\perp$	Independence relation symbol
$\gg$	We write $f(n) \gg g(n)$ to mean $g(n) = o(f(n))$
$\stackrel{\mathcal{D}}{=}$	Equal in distribution
$\Rightarrow$	Weak convergence
$\stackrel{\mathcal{D}}{\to}$	Convergence in distribution
$\|\cdot\|_\infty$	Supremum norm : $\|f\|_\infty = \sup_{x\in\mathbb{R}} \|f(x)\|$

Bibliography

[1] ABRAMOWITZ, M. and STEGUN, I. A. *Handbook of Mathematical Functions with Formulas Graphs and Mathematical Tables.* Dover Publications, Inc., New York, 1965.

[2] APOSTOL, T. *Mathematical Analysis.* second ed. Addison–Wesley, 1974.

[3] ARFKEN, G. *Mathematical Methods for Physicists.* 3rd ed. Orlando, FL: Academic Press, 1985.

[4] BALDI, P., RINOTT, Y. and STEIN, C. A normal approximation for the number of local maxima of a random function on a graph. In *Probability, Statistics and Mathematics, Papers in Honor of Samuel Karlin* (T. W. Anderson, K. B. Athreya and D. L. Iglehart, eds.) Academic Press, New York. (1989), pp. 59–81.

[5] BARBOUR, A. D. Stein's method and Poisson process convergence. *J. Appl. Probab.* **25A** (1988), pp. 175–184.

[6] BARBOUR, A. D. Stein's method for diffusion approximations. *Probab. Theory Related Fields* **84** (1990), pp. 297–322.

[7] BARBOUR, A. D. and CHEN, L. H. Y. *An Introduction to Stein's Method.* Singapore University Press, 2005.

[8] BARICZ, Á. On a product of modified Bessel functions. *Proc. Amer. Math. Soc* **137** (2009), pp. 189–193.

[9] BARICZ, Á. Turán type inequalities for modified Bessel functions. *Bull. Aust. Math. Soc.* **82** (2010), pp. 254-264.

[10] BARNDORFF-NIELSEN, O. E. Exponentially decreasing distributions for the logarithm of particle size. *Proc. Roy. Soc. London A* **353** (1977), pp. 401-419.

[11] BARNDORFF-NIELSEN, O. E., KENT, J. and SØRENSEN, M. Normal Variance-Mean Mixtures and z Distributions. *International Statistical Review* **50** (1982), pp. 145–159.

[12] H. BATEMAN, *Tables of Integral Transforms.* vol. I, McGraw-Hill, New York, 1954.

[13] BERRY, A. C. The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** (1941), pp. 126–136.

[14] 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.

[15] BILLINGSLEY, P. *Convergence of Probability Measures.* second ed. Wiley, 1999.

[16] CHATTERJEE, S., DIACONIS, P. and MECKES, E. Exchangeable pairs and Poisson approximation. *Probab. Surv.* **2** (2005), pp. 64–106.

[17] CHATTERJEE, S. and MECKES, E. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** (2008), pp. 257-283.

[18] CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q–M. *Normal Approximation By Stein's Method.* Springer, 2011.

[19] COLLINS, P. J. *Differential and Integral Equations.* Oxford University Press, 2006.

[20] DURRETT, R. *Stochastic Calculus: a Practical Introduction.* CRC Press, 1996.

[21] 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.

[22] ETHIER, P. and KURTZ, T. *Markov Processes - characterisation and Convergence.* Wiley, 1986.

[23] ERDÉLYI, A., MAGNUS, W., OBERHETTINGER, F. and TRICOMI, F. G. *Higher Transcendental Functions. Vol. I.* McGraw-Hill, 1953.

[24] ELEZOVIĆ, N., GIORDANO, C. and PEČARIĆ, J. The best bounds in Gautschi's inequality. *Math. Inequal. Appl.* **3** (2000), pp. 239-252.

[25] ESSÉEN, C–G. Fourier analysis of distribution functions, a mathematical study of the Laplace-Gaussian law. *Acta Math.* **77** (1945), pp. 1–125.

[26] FINLAY, R. and SENETA, E. Option pricing with VG-like models, *International Journal of Theoretical and Applied Finance.* **11**, (2008), pp. 943-955.

[27] FOLLAND, G. *Real Analysis: Modern Techniques and Their Applications.* Wiley, 1984.

[28] FRIEDMAN, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Statist. Assoc.* **32**, (1937), pp. 675–701.

[29] 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.

[30] GOLDSTEIN, L. and REINERT, G. Zero biasing in one and higher dimensions, and applications. In *Steins Method and Applications. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** Singapore Univ. Press, Singapore. (2005), pp. 1-18.

[31] GOLDSTEIN, L. and REINERT, G. Distributional transformations, orthogonal polynomials, and Stein characterisations. *J. Theoret. Probab.* **18** (2005), pp. 237–260.

[32] GOLDSTEIN, L. and RINOTT, Y. Multivariate normal approximations by Steins method and size bias couplings. *J. Appl. Probab.* **33** (1996), pp. 1–17.

[33] GÖTZE, F. On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** (1991), pp. 724–739.

[34] GRADSHETYN, I. S. and RYZHIK, I. M. *Table of Integrals, Series and Products.* seventh ed. Academic Press, 2007.

[35] HALL, P. *The Bootstrap and Edgeworth Expansion.* Springer–Verlag, 1991.

[36] 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.

[37] JENSEN, D. R. On approximating the distributions of Friedman's χ_r^2 and related statistics. *Metrika* **24** (1977), pp. 75-86.

[38] JONES, A. L. An extension of an inequality involving modified Bessel functions. *J. Math. and Phys.* **47** (1968), pp. 220–221.

[39] KENT, J. Some Probabilistic Properties of Bessel Functions. *Ann. Probab.* **6** (1978), pp. 760–770.

[40] LAFORGIA, A. and NATALINI, P. Some inequalities for modified Bessel functions. *J. Inequal. Appl.* (2010), Art. 253035, 10pp.

[41] LEY, C. and SWAN, Y. A unified approach to Stein characterizations. *http://arxiv.org/abs/1105.4925*, 2011.

[42] LINETSKY, V. The spectral representation of Bessel processes with drift: Applications in queueing and finance. *J. Appl. Probab.* **41** (2004), pp. 327–344.

[43] 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.

[44] LUK, H. *Stein's Method for the Gamma Distribution and Related Statistical Applications.* PhD thesis, University of Southern California, 1994.

[45] LUKE, Y. L. Inequalities for Generalized Hypergeometric Functions, *J. Approx. Theory* **5** (1972), pp. 41–65.

[46] MADAN, D. B. and SENETA, E. The Variance Gamma (V.G.) Model for Share Market Returns. *Journal of Business* **63** (1990), pp. 511–524.

[47] MARDIA, K. V., KENT, J. T. and BIBBY, J. M. *Multivariate Analysis.* Academic Press Inc, 1979.

[48] MECKES, E. S. On Stein's method for multivariate normal approximation. *IMS Collect.* **5** (2009), pp. 153-178.

[49] NÅSELL, I. Inequalities for Modified Bessel Functions. *Mathematics of Computation* **28** (1974), pp. 253–256.

[50] NUALART, D. *The Malliavin Calculus and Related Topics.* Springer-Verlag, Berlin, 2nd edition, 2006.

[51] ØKSENDAL, B. *Stochastic Differential Equations: an Introduction with Applications.* fifth ed. Springer–Verlag, 2000.

[52] OLVER, F. W. J., LOZIER, D. W., BOISVERT, R. F. and CLARK, C. W. *NIST Handbook of Mathematical Functions.* Cambridge University Press, 2010.

[53] PATNAIK, P. B. The noncentral chi-square and F-distributions and their applications. *Biometrika,* **36** 1949, pp. 202–232.

[54] PEKÖZ, E., RÖLLIN, A. and ROSS, N. Degree asymptotics with rates for preferential attachment random graphs. *http://arxiv.org/abs/ 1108.5236*, 2011.

[55] PICKETT, A. *Rates of Convergence of χ^2 Approximations via Stein's Method.* DPhil thesis, University of Oxford, 2004.

[56] PRIESTLEY, H. A. *Introduction to Integration.* Oxford University Press, 1997.

[57] RAIC, M. Normal approximation by Steins method. *In: Proceedings of the 7th Young Statisticians Meeting* (2003), pp. 71-97.

[58] REINERT, G. Couplings for Normal Approximations with Stein's Method. In *Microsurveys in Discrete Probability*, vol. of *DIMACS series. AMS,* (1998), pp. 193-207.

[59] REINERT, G. Three general approaches to Steins method. In *An Introduction to Steins Method. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* Singapore Univ. Press, Singapore. **4** (2005), pp. 183-221.

[60] REINERT, G., SCHBATH, S. and WATERMAN, M. S. in LOTHAIRE, M. *Applied Combinatorics on Words.* Cambridge University Press, 2005.

[61] 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.

[62] 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.

[63] RINOTT, Y. and ROTAR, V. On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.* **7** (1997), pp. 1080–1105.

[64] ROBBINS, H. A Remark of Stirling's Formula. *Amer. Math. Monthly* **62** (1955), pp. 26–29.

[65] RÖLLIN, A. Translated Poisson approximation using exchangeable pair couplings. *Ann. Appl. Prob.* **17** (2007), pp. 1596–1614.

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

[67] SEGURA, J. Bounds for ratios of modified Bessel functions and associated Turán-type inequalities. *J. Math. Anal. Appl.* **374** (2011), pp. 516–528.

[68] SHORACK, G. R. The weighted process of row independent random variables with arbitrary distribution functions. *Statist. Neerlandica* **33** (1979), pp.169–189.

[69] SONI, R. P. On an inequality for modified Bessel functions. *Journal of Mathematical Physics* **44** (1965), pp. 406–407.

[70] SPRINGER, M. D. and THOMPSON, W. E. The distribution of products of Beta, Gamma and Gaussian random variables. *SIAM J. Appl. Math.* **18** (1970), pp. 721–737.

[71] STADJE, W. Probabilistic proofs of some formulas for Bessel functions. *Indagationes Mathematicae Proceedings* **86** (1983), pp. 343–359.

[72] 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. Statis. Prob.* (1972), vol. 2, Univ. California Press, Berkeley, pp. 583–602.

[73] STEIN, C. *Approximate Computation of Expectations.* IMS, Hayward, California, 1986.

[74] STEIN, C., DIACONIS, S, HOLMES, S. and REINERT, G. Use of exchangeable pairs in the analysis of simulations. In Steins method: expository lectures and applications, volume 46 of IMS Lecture Notes Monogr. Ser., Inst. Math. Statist., Beachwood, OH (2004), pp. 1–26.

[75] WINKELBAUER, A. Moments and absolute moments of the normal distribution. *http://arxiv.org/abs/1209.4340*, 2012.

[76] ZWILLINGER, D. *CRC Standard Mathematical Tables and Formulae.* 31st ed. Chapman and Hall/CRC, 2002.