

# On Stein’s method for products of normal random variables and zero bias couplings

ROBERT E. GAUNT

*Department of Statistics, University of Oxford, 24–29 St. Giles’, Oxford, OX1 3LB, UK.  
E-mail: gaunt@stats.ox.ac.uk*

In this paper, we extend Stein’s method to the distribution of the product of  $n$  independent mean zero normal random variables. A Stein equation is obtained for this class of distributions, which reduces to the classical normal Stein equation in the case  $n = 1$ . This Stein equation motivates a generalisation of the zero bias transformation. We establish properties of this new transformation, and illustrate how they may be used together with the Stein equation to assess distributional distances for statistics that are asymptotically distributed as the product of independent central normal random variables. We end by proving some product normal approximation theorems.

*Keywords:* coupling; distributional transformation; products of normal random variables; Stein’s method; zero biasing

## 1. Introduction

In 1972, Stein [39] introduced a powerful method for deriving bounds for normal approximation. Since then, Stein’s method has been extended to many other distributions, such as the Poisson [5], beta [8,25], gamma [21,29,31], exponential [4,34], Laplace [9,36] and variance-gamma [13, 20]; for an overview see Reinert [37].

Stein’s method for normal approximation rests on the following characterization of the normal distribution, which can be found in Stein [40], namely  $Z \sim N(0, \sigma^2)$  if and only if

$$\mathbb{E}[\sigma^2 f'(Z) - Zf(Z)] = 0 \tag{1}$$

for all real-valued absolutely continuous functions  $f$  such that  $\mathbb{E}|f'(Z)|$  exists. This gives rise to the following inhomogeneous differential equation, known as the Stein equation:

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

where  $Z \sim N(0, \sigma^2)$ , and the test function  $h$  is real-valued. For any real-valued bounded measurable test function  $h$ , a solution  $f$  to (2) exists. Now, evaluating both sides at any random variable  $W$  and taking expectations gives

$$\mathbb{E}[\sigma^2 f'(W) - Wf(W)] = \mathbb{E}h(W) - \mathbb{E}h(Z). \tag{3}$$

Thus, the problem of bounding the quantity  $\mathbb{E}h(W) - \mathbb{E}h(Z)$  reduces to the bounding the left-hand side of (3).

### 1.1. The product normal distribution

In this paper, we extend Stein’s method to products of independent central normal random variables. The probability density function of the products of independent mean zero normal variables was shown by Springer and Thompson [38] to be a Meijer  $G$ -function (defined in Appendix B.1.1). The probability density function of the product  $Z = X_1 X_2 \cdots X_n$  of independent normal random variables  $N(0, \sigma_i^2), i = 1, 2, \dots, n$ , is given by

$$p_n(x) = \frac{1}{(2\pi)^{n/2}\sigma} G_{0,n}^{n,0} \left( \frac{x^2}{2^n \sigma^2} \middle| 0, \dots, 0 \right), \quad x \in \mathbb{R}, \tag{4}$$

where  $\sigma = \sigma_1 \sigma_2 \cdots \sigma_n$ . If (4) holds, then we say that  $Z$  has a *product normal* distribution, and write  $Z \sim \text{PN}(n, \sigma^2)$ . For the case of two products, (4) simplifies to

$$p_2(x) = \frac{1}{\pi \sigma_1 \sigma_2} K_0 \left( \frac{|x|}{\sigma_1 \sigma_2} \right), \quad x \in \mathbb{R},$$

where  $K_0(x)$  is a modified Bessel function of the second kind (defined in Appendix B.2.1).

The density  $p_n$  satisfies the asymptotic formulas

$$p_n(x) \sim \frac{1}{(2\pi)^{n/2}(n-1)! \sigma} \left[ -\log \left( \frac{|x|}{\sigma} \right) \right]^{n-1}, \quad \text{as } x \rightarrow 0,$$

and

$$p_n(x) \sim \frac{2^{n/2-1}}{\sigma \sqrt{\pi n}} \left( \frac{|x|}{\sigma} \right)^{1/n-1} \exp \left( -\frac{n}{2} \left( \frac{|x|}{\sigma} \right)^{2/n} \right), \quad \text{as } |x| \rightarrow \infty \tag{5}$$

(see Theorem 6 of Springer and Thompson [38], and Luke [30], Section 5.7, Theorem 5). It is interesting to compare these asymptotic formulas with the exact formula for the density of the random variable  $|X|^n \operatorname{sgn}(X)$ , where  $X \sim N(0, \sigma^2)$ , as given by

$$p(x) = \frac{1}{n\sigma \sqrt{2\pi}} \left( \frac{|x|}{\sigma} \right)^{1/n-1} \exp \left( -\frac{1}{2} \left( \frac{|x|}{\sigma} \right)^{2/n} \right), \quad x \in \mathbb{R},$$

which is seen to have a greater singularity at the origin and slower decay at the tails than the  $\text{PN}(n, \sigma^2)$  distribution.

### 1.2. A Stein equation for the product normal distribution

One of the main results of this paper is a Stein equation for the  $\text{PN}(n, \sigma^2)$  distribution:

$$\frac{\sigma^2}{x} \left( x \frac{d}{dx} \right)^n f(x) - x f(x) = h(x) - \text{PN}_n^{\sigma^2} h, \tag{6}$$

where  $\text{PN}_n^{\sigma^2} h$  denotes the quantity  $\mathbb{E}h(X)$  for  $X \sim \text{PN}(n, \sigma^2)$ . For ease of reading, we define the operators  $Tf(x) = xf'(x)$  and  $A_n f(x) = x^{-1}T^n f(x)$ . With this notation, the  $\text{PN}(n, \sigma^2)$  Stein equation (6) can be written as

$$\sigma^2 A_n f(x) - xf(x) = h(x) - \text{PN}_n^{\sigma^2} h. \quad (7)$$

For all  $n \geq 1$ , the operator  $T^n$  satisfies the fundamental identity (Luke [30], page 24)  $T^n f(x) = \sum_{k=1}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^k f^{(k)}(x)$ , where  $\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \frac{1}{k!} \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} j^n$  are Stirling numbers of the second kind (Olver *et al.* [33], Chapter 26). Consequently, (7) is a  $n$ th order linear differential equation with simple coefficients:

$$\sigma^2 \sum_{k=1}^n \left\{ \begin{matrix} n \\ k \end{matrix} \right\} x^{k-1} f^{(k)}(x) - xf(x) = h(x) - \text{PN}_n^{\sigma^2} h. \quad (8)$$

As an  $n$ th order differential equation, this Stein equation is almost unique in the literature. The exceptions being the  $n$ th order Stein equations of Goldstein and Reinert [24] that involve orthogonal polynomials, and the very recent  $n$ th order Stein equations for products of  $n$  independent beta and gamma random variables (see Gaunt [17]) and linear combinations of  $n$  independent gamma random variables (see Arras *et al.* [1]).

The  $\text{PN}(n, \sigma^2)$  Stein equation is a natural generalisation of the normal Stein equation (2) to one for products of independent central normal random variables and has a number of attractive properties; a further discussion is given in Remark 2.1. For the case  $n = 2$ , (8) reduces to

$$\sigma^2 xf''(x) + \sigma^2 f'(x) - xf(x) = h(x) - \text{PN}_2^{\sigma^2} h, \quad (9)$$

and in Lemma 2.3 we obtain the unique bounded solution of (9), as well as bounds on its first four derivatives (see Theorem 2.1), which hold provided that  $h$  and its first three derivatives are bounded. Note that (9) is a second order linear differential equation involving  $f$ ,  $f'$  and  $f''$ . Such Stein equations are uncommon in the literature, although Peköz *et al.* [35] have recently obtained a similar Stein equation for the family of densities

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

where  $U(a, b, x)$  denotes the confluent hypergeometric function of the second kind (see Olver *et al.* [33], Chapter 13). Pike and Ren [36] have also obtained a second order Stein operator for the Laplace distribution. It should also be noted that the  $\text{PN}(2, \sigma^2)$  distribution is a member of the class of variance-gamma distributions, and that (9) is indeed a special case of the variance-gamma Stein equation, recently introduced by Gaunt [16]. Very recently, Gaunt [19] obtained a second order Stein operator for the generalized hyperbolic distribution. A limiting case of this Stein operator is the variance-gamma Stein operator.

In Section 2.3, we solve the product normal Stein equation. For the case of two products, we obtain uniform bounds for the solution and its first four derivatives in terms of supremum norms of derivatives of the test function  $h$ . However, for  $n \geq 3$ , the solution takes on a rather

complicated form, involving integrals of Meijer  $G$ -functions, and we have been unable to obtain estimates for the solution and its derivatives. This is left as an interesting open problem and is discussed further in Section 2.3.2.

### 1.3. A generalised zero bias transformation

Returning to normal approximation, the quantity  $\mathbb{E}[\sigma^2 f'(W) - Wf(W)]$  is typically bounded through the use of a coupling and Taylor expansions. A coupling technique that is commonly used in the case mean zero random variables is the zero bias transformation, which was introduced by Goldstein and Reinert [22]. If  $W$  is a mean zero random variable with finite, non-zero variance  $\sigma^2$ , we say that  $W^*$  has the  $W$ -zero biased distribution if for all differentiable  $f$  for which  $\mathbb{E}Wf(W)$  exists,

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

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  $\sigma^2 E[f'(W) - f'(W^*)]$ . We therefore have

$$\mathbb{E}[\sigma^2 f'(W) - Wf(W)] = \sigma^2 \mathbb{E}[f'(W) - f'(W^*)], \tag{10}$$

and the right-hand side may be bounded by Taylor expanding about  $W$ . Goldstein and Reinert [22] presented a number of interesting properties and obtained some useful constructions, such as:

**Lemma 1.1.** *Let  $X_1, \dots, X_n$  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*

$$W_i = W - X_i = \sum_{j \neq i} X_j.$$

*Then  $W_i + X_i^*$  has the  $W$ -zero biased distribution, where  $X_i^*$  has the  $X_i$ -zero biased distribution.*

Such constructions combined with a Taylor expansion of the right-hand side of (10) often allow simple proofs of limit theorems for normal approximation.

Motivated by the zero bias transformation and the multivariate normal Stein equation, Goldstein and Reinert [23] extended the concept of the zero bias transformation to any finite dimension. In this paper, we introduce another generalisation of the zero bias transformation. The product normal Stein equation and zero bias transformation motivate the following definition.

**Definition 1.1.** *Let  $W$  be a mean zero random variable with finite, non-zero variance  $\sigma^2$ . We say that  $W^{*(n)}$  has the  $W$ -zero biased distribution of order  $n$  if for all  $n$  times differentiable  $f$  for which  $\mathbb{E}Wf(W)$  exists,*

$$\mathbb{E}Wf(W) = \sigma^2 \mathbb{E}A_n f(W^{*(n)}), \tag{11}$$

where  $A_n f(x) = x^{-1} T^n f(x)$  and  $Tf(x) = xf'(x)$ .

The existence of the zero biased distribution of order  $n$  for any  $W$  with zero mean and non-zero, finite variance is established by Lemma 2.2. The zero bias transformation of order  $n$  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, introduced by Goldstein and Reinert [23], is a natural extension to random vectors in  $\mathbb{R}^d$ .

The zero bias transformation of order  $n$  has a number of interesting properties that generalise those of the zero bias transformation. These properties are collected in Propositions 3.2 and 3.3. The application of the zero bias transformation of order  $n$  to assess distributional distances for statistics that have a limiting product normal distribution is analogous to that of the zero bias transformation for normal approximation. We illustrate how such approximation results can be obtained with Theorem 4.1, provided that we have bounds on the relevant derivatives of the solution to the Stein equation (7).

In obtaining these bounds, we use the fact (Proposition 3.3, part (iv)) that the zero bias transformation of order  $n$  of the random variable  $W = \prod_{k=1}^n W_k$ , where the  $W_k$  are independent, is given by  $W^{*(n)} = \prod_{k=1}^n W_k^*$ . One can then construct the zero bias transformations for each of the  $W_k$  by using one of the constructions given by Goldstein and Reinert [22]. 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 1.1.

It is worth comparing this construction with an analogous result, due to Luk [29], involving size bias couplings (for an application of this coupling to normal approximation see Baldi, Rinott and Stein [3]). 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,

$$\mathbb{E}Wf(W) = \mu \mathbb{E}f(W^s).$$

Now, if  $W = \prod_{k=1}^n W_k$ , where the  $W_k$  are positive, independent random variables and  $W_1^s, \dots, W_n^s$  are independent random variables with  $W_k^s$  having the  $W_k$ -size biased distribution, then  $W^s = \prod_{k=1}^n W_k^s$  has the  $W$ -size biased distribution. The constructions are similar, although the product of  $n$  zero bias distributions has the  $W$ -zero bias distribution of order  $n$ , rather than the  $W$ -zero bias distribution.

In Section 4, we illustrate how the product normal Stein equation may be used together with the zero bias transformation of order  $n$  to assess distributional distances for statistics that are asymptotically distributed as products of independent normal random variables, and we obtain explicit bounds on the convergence rate in the  $n = 2$  case. We are restricted to the case  $n = 2$  because we are unable to obtain bounds for the solution of  $\text{PN}(n, \sigma^2)$  Stein equation for  $n \geq 3$ . In Section 5, we get around this difficulty by using a recent results, due to Gaunt [20], that allow one to obtain bounds on distributional distances when the limit distribution that be represented as a function of a multivariate normal random variable, of which the product normal distribution is an obvious example. We end Section 5 by comparing the bounds obtained by this approach that bypasses the Stein equation and those derived in Section 4 using the Stein equation.

The approach taken in this paper is somewhat classical, although it should be noted that product normal limit theorems have recently been established in the context of Malliavin calculus. Indeed, the Malliavin–Stein method (see Nourdin and Peccati [32] for an introduction), that was

originally developed for Gaussian approximation, is applicable to a wide class of laws; see, for example, Eden and Viens [11] and Eden and Viquez [12]. Eichelsbacher and Thäle [13] used the Malliavin–Stein approach to obtain bounds for variance-gamma approximation of functionals of isonormal Gaussian processes. In particular, they showed that a sequence of random variables in the second Wiener chaos converges to a variance-gamma limit if and only if their second to sixth moments converge to those a variance-gamma random variable. A special case of this general six moment theorem is a six moment theorem for the distribution of the product of two (possibly correlated) standard normal variables (see [13], Theorem 5.8 and Remark 5.9), with explicit bounds on the rate of convergence.

The results of [13] are complemented by those of Azmooden *et al.* [2]. They established necessary and sufficient conditions under which a sequence of random variables living inside a Wiener chaos of arbitrary order converge to limiting random variables, whose distribution is of the form  $\sum_{i=1}^k \alpha_i (Z_i^2 - 1)$ , where  $k$  is a finite integer, the  $\alpha_i$ ,  $i = 1, \dots, k$ , are piecewise distinct and the  $Z_i$  are independent  $N(0, 1)$  variables. The case  $k = 2$ ,  $\alpha_1 = \frac{1}{2} = -\alpha_2$  corresponds to a limit with the same distribution as a  $\text{PN}(2, 1)$  random variable, and other parameter values can yield results for limiting distributions that fall outside the variance-gamma class.

## 1.4. Outline of the paper

We begin Section 2 by obtaining some useful properties for the operator  $A_n$ . We use some of these properties to prove the existence and uniqueness of the zero bias distribution of order  $n$ . We then present a characterising equation for the product normal distribution which motivates the Stein equation (7). Also, for the case  $n = 2$  we obtain the unique bounded solution, as well as bounds on its first four derivatives. The general case  $n \geq 3$  is more difficult, but we are still able to solve the equation. We also discuss how one may extend the approach used in the  $n = 2$  case to this more challenging case. In this section, we also consider an application of our product normal characterisation to the problem of obtaining a formula for the characteristic function of a product normal distributed random variable. In Section 3, we present some of the properties of the zero bias distribution of order  $n$ . In Section 4, we illustrate how the product normal Stein equation may be used together with the zero bias transformation to derive product normal approximation results. In Section 5, we bypass the product normal Stein equation to prove two product normal approximations for general  $n$ . The proof of Theorem 2.1 is given in Appendix A. Appendix B provides a list of some elementary properties of the Meijer  $G$ -function and modified Bessel functions that are used in this paper. Appendix C provides a list of inequalities for expressions involving derivatives and integrals of modified Bessel functions that are used to bound the derivatives of the solution of the  $\text{PN}(2, \sigma^2)$  Stein equation.

**Notation.** Throughout this paper,  $T$  will denote the operator  $Tf(x) = xf'(x)$  and  $A_n$  will denote the operator  $A_n f(x) = x^{-1} T^n f(x) = \frac{d}{dx} (T^{n-1} f(x))$ . We shall write  $C^n(\mathbb{R}^d)$  for the space of  $k$  times differentiable functions on  $\mathbb{R}^d$ . We let  $C_b^n(\mathbb{R}^d)$  denote the space of bounded functions on  $\mathbb{R}^d$  with bounded  $k$ th order derivatives for  $k \leq n$ . We shall also write  $\|f\| = \|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$ .

## 2. A Stein equation for products of independent central normal variables

### 2.1. The product normal Stein equation

In this section, we obtain a characterising equation for the product normal distribution which motivates the  $\text{PN}(n, \sigma^2)$  Stein equation (7). Before presenting that result, we present some useful properties of the operator  $A_n f(x) = x^{-1} T^n f(x)$  and establish the existence of the zero bias distribution of order  $n$  for any  $W$  with zero mean and finite, non-zero variance.

We begin by obtaining an inverse of the operator  $A_n$ . This inverse operator will be used throughout this paper in establishing properties of the zero bias distribution of order  $n$ .

**Lemma 2.1.** *Let  $V_n$  be the product of  $n$  independent  $U(0, 1)$  random variables, and define the operator  $G_n$  by  $G_n f(x) = x \mathbb{E} f(x V_n)$ . Then,  $G_n$  is the right-inverse of the operator  $A_n$  in the sense that*

$$A_n G_n f(x) = f(x).$$

Suppose now that  $f \in C^n(\mathbb{R})$ . Then, for any  $n \geq 1$ ,

$$G_n A_n f(x) = G_1 A_1 f(x) = f(x) - f(0). \quad (12)$$

Therefore,  $G_n$  is the inverse of  $A_n$  when the domain of  $A_n$  is the space of all  $n$  times differentiable functions  $f$  on  $\mathbb{R}$  with  $f(0) = 0$ .

**Proof.** We begin by obtaining a useful formula for  $G_n f(x) = x \mathbb{E} f(x V_n)$ . We have that

$$G_n f(x) = x \int_{(0,1)^n} f(x u_1 \cdots u_n) du_1 \cdots du_n,$$

and by a change of variables we can write

$$G_1 f(x) = \int_0^x f(t_1) dt_1, \quad (13)$$

and for  $n \geq 2$ ,

$$G_n f(x) = \int_0^x \int_0^{t_n} \cdots \int_0^{t_2} \frac{1}{t_2 t_3 \cdots t_n} f(t_1) dt_1 dt_2 \cdots dt_n. \quad (14)$$

We now show that  $G_n$  is the right-inverse of the operator  $A_n$ . It is immediate from (13) that  $A_1 G_1 f(x) = f(x)$ . For  $n \geq 2$ , we differentiate the right-hand side of (14) to obtain

$$T G_n f(x) = x \frac{d}{dx} (G_n f(x)) = G_{n-1} f(x),$$

and hence  $T^k G_n f(x) = G_{n-k} f(x)$ . Using this recursive formula yields

$$A_n G_n f(x) = x^{-1} T^n G_n f(x) = \frac{d}{dx} (T^{n-1} G_n f(x)) = \frac{d}{dx} (G_1 f(x)) = f(x).$$

Finally, we verify relation (12). For  $n \geq 2$ , we have

$$\begin{aligned} G_n A_n f(x) &= \int_0^x \int_0^{t_n} \cdots \int_0^{t_2} \frac{1}{t_2 \cdots t_n} \frac{d}{dt_1} (T^{n-1} f(t_1)) dt_1 dt_2 \cdots dt_n \\ &= \int_0^x \int_0^{t_n} \cdots \int_0^{t_3} \frac{1}{t_2 t_3 \cdots t_n} [T^{n-1} f(t_1)]_{t_1=0}^{t_1=t_2} dt_2 dt_3 \cdots dt_n \\ &= \int_0^x \int_0^{t_n} \cdots \int_0^{t_3} \frac{1}{t_3 \cdots t_n} \cdot t_2^{-1} T^{n-1} f(t_2) dt_2 dt_3 \cdots dt_n \\ &= G_{n-1} A_{n-1} f(x), \end{aligned}$$

where we used that  $A_n f(x) = x^{-1} T^n f(x) = \frac{d}{dx} (T^{n-1} f(x))$ . Iterating gives

$$G_n A_n f(x) = G_1 A_1 f(x) = \int_0^x f'(t_1) dt_1 = f(x) - f(0),$$

as required. □

We are now in a position to prove the existence and uniqueness of the zero bias distribution of order  $n$  for any  $W$  with zero mean and finite, non-zero variance. The proof is a generalisation of the proof of existence of the zero bias distribution that was given in Goldstein and Reinert [22].

**Lemma 2.2.** *Let  $W$  be a mean zero random variable with finite, non-zero variance  $\sigma^2$ . Then there exists a unique random variable  $W^{*(n)}$  such that for all  $f \in C^n(\mathbb{R})$  for which the relevant expectations exist we have*

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

**Proof.** For  $f \in C_c$ , the collection of continuous functions with compact support, define a linear operator  $R$  by

$$Rf = \sigma^{-2} \mathbb{E}W G_n f(W),$$

where  $G_n$  is defined as in Lemma 2.1. Then  $Rf$  exists, since  $\mathbb{E}W^2 < \infty$ . To see, moreover, that  $R$  is positive, take  $f \geq 0$ . Then  $G_n f(x)$  is increasing, and therefore  $W$  and  $G_n f(W)$  are positively correlated. Hence  $\mathbb{E}W G_n f(W) \geq \mathbb{E}W \mathbb{E}f G_n(W) = 0$ , and  $R$  is positive. Using the Riesz representation theorem (see, for example, Folland [14]) we have  $Rf = \int f d\nu$ , for some unique Radon measure  $\nu$ , which is a probability measure as  $R1 = 1$ .

We now take  $f(x) = A_n g(x)$ , where  $g \in C^n(\mathbb{R})$ , with derivatives up to  $n$ th order being continuous with compact support. Then, from (12), we have

$$\mathbb{E}W G_n A_n g(W) = \mathbb{E}W(g(W) - g(0)) = \mathbb{E}Wg(W),$$



which completes the proof.  $\square$

The following proposition motivates the product normal Stein equation (7).

**Proposition 2.1.** *Suppose  $Z \sim \text{PN}(n, \sigma^2)$ . Let  $f \in C^n(\mathbb{R})$  be such that  $\mathbb{E}|Zf(Z)| < \infty$  and  $\mathbb{E}|Z^{k-1}f^{(k)}(Z)| < \infty$ ,  $k = 1, \dots, n$ . Then*

$$\mathbb{E}[\sigma^2 A_n f(Z) - Zf(Z)] = 0. \quad (15)$$

**Proof.** Define  $W_n = \prod_{i=1}^n X_i$  where  $X_i \sim N(0, \sigma_{x_i}^2)$  and the  $X_i$  are independent. We also let  $\sigma_n = \sigma_{x_1} \sigma_{x_2} \cdots \sigma_{x_n}$  and observe that  $(T^n f)(ax) = T^n f_a(x)$  where  $f_a(x) = f(ax)$ .

We prove the result by induction on  $n$ . For  $n = 1$ , the result follows immediately from the characterisation (1) for the normal distribution. By induction assume that  $\mathbb{E}W_n g(W_n) = \sigma_n^2 \mathbb{E}A_n g(W_n) = \sigma_n^2 \mathbb{E}W_n^{-1} T^n g(W_n)$  for all  $g \in C^n(\mathbb{R})$  for some  $n \geq 1$ . Then

$$\begin{aligned} \mathbb{E}W_{n+1} f(W_{n+1}) &= \mathbb{E}[X_{n+1} \mathbb{E}[W_n f_{X_{n+1}}(W_n) | X_{n+1}]] \\ &= \mathbb{E}[X_{n+1} \mathbb{E}[\sigma_n^2 W_n^{-1} (T^n f_{X_{n+1}})(W_n) | X_{n+1}]] \\ &= \sigma_n^2 \mathbb{E}[X_{n+1} W_n^{-1} (T^n f)(W_n X_{n+1})] \\ &= \sigma_n^2 \mathbb{E}[W_n^{-1} \mathbb{E}[X_{n+1} \cdot (T^n f_{W_n})(X_{n+1}) | W_n]] \\ &= \sigma_{n+1}^2 \mathbb{E}[W_n^{-1} \mathbb{E}[X_{n+1}^{-1} (T^{n+1} f_{W_n})(X_{n+1}) | W_n]] \\ &= \sigma_{n+1}^2 \mathbb{E}[W_{n+1}^{-1} T^{n+1} f(W_{n+1})], \end{aligned}$$

as required.  $\square$

**Remark 2.1.** We could have obtained a first order Stein operator for the  $\text{PN}(n, \sigma^2)$  distributions using the density approach of Stein *et al.* [41] (see also Ley *et al.* [28] for an extension of the scope of the density method). However, this approach would lead to complicated operators involving Meijer  $G$ -functions, which, in contrast to our Stein equation, may not be amenable to the use of couplings.

## 2.2. Applications of Proposition 2.1

The main application of Proposition 2.1 that is considered in this paper involves the use of the resulting  $\text{PN}(2, \sigma^2)$  Stein equation in the proofs of the limit theorems of Section 4. It is, however, possible to obtain other interesting results using Proposition 2.1. As an example, we demonstrate how the characterising equation (15) of the product normal distributions can be used to derive a formula for the characteristic function of the  $\text{PN}(n, \sigma^2)$  distribution.

Let  $W_n \sim \text{PN}(n, \sigma^2)$ . We begin by noting that the moment generating function of  $W_n$  is only defined for  $n \leq 2$  (see (5)). We therefore consider the characteristic function of  $W_n$ . On taking

$f(x) = e^{itx}$  in the characterising equation (15) and setting  $\phi_n(t) = \mathbb{E}[e^{itW_n}]$ , we deduce that  $\phi_n(t)$  satisfies the differential equation

$$\sigma^2 t \left( t \frac{d}{dt} + 1 \right)^{n-1} \phi_n(t) + \phi_n'(t) = 0. \tag{16}$$

Solving this differential equation subject to the conditions that  $\phi_n(t)$  is the characteristic function of  $W_n$  would yield a formula for the characteristic function of  $W_n$ . For simplicity, we consider the cases  $n = 2, 3$  (the case  $n = 1$  gives the well-known formula  $\phi_1(t) = e^{-\frac{1}{2}\sigma^2 t^2}$ ).

For  $n = 2$ , the differential equation (16) reduces to

$$(1 + \sigma^2 t^2) \phi_2'(t) + \sigma^2 t \phi_2(t) = 0.$$

Solving this equation subject to the condition  $\phi_2(0) = 1$  gives

$$\phi_2(t) = \frac{1}{\sqrt{1 + \sigma^2 t^2}}.$$

This formula was also obtained in Example 11.22 of Stuart and Ord [42].

For  $n = 3$ , the characteristic function  $\phi_3(t)$  satisfies

$$\sigma^2 t^3 \phi_3''(t) + (3\sigma^2 t^2 + 1) \phi_3'(t) + \sigma^2 t \phi_3(t) = 0.$$

It is straightforward to verify the general solution of this differential equation is given by

$$\phi_3(t) = A|t|^{-1} \exp\left(\frac{1}{4\sigma^2 t^2}\right) I_0\left(\frac{1}{4\sigma^2 t^2}\right) + B|t|^{-1} \exp\left(\frac{1}{4\sigma^2 t^2}\right) K_0\left(\frac{1}{4\sigma^2 t^2}\right),$$

where  $I_0(x)$  and  $K_0(x)$  are modified Bessel functions (see Appendix B.2.1 for definitions), which satisfy the modified Bessel differential equation (54). From the asymptotic formula  $I_0(x) \sim \frac{e^x}{\sqrt{2\pi x}}$  as  $x \rightarrow \infty$  (see (47)), it follows that the solution is unbounded as  $t \rightarrow 0$  unless  $A = 0$ . Since the characteristic function  $\phi_3(t)$  must be bounded for all  $t \in \mathbb{R}$ , we take  $A = 0$ . We also require that  $\phi_3(0) = 1$ , and so on applying the asymptotic formula  $K_0(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x}$  as  $x \rightarrow \infty$  (see (46)), we take  $B = (\sigma\sqrt{2\pi})^{-1}$ . Hence,

$$\phi_3(t) = \frac{1}{\sqrt{2\pi\sigma^2 t^2}} \exp\left(\frac{1}{4\sigma^2 t^2}\right) K_0\left(\frac{1}{4\sigma^2 t^2}\right).$$

In principle, this approach could be used to obtain a formula for  $\phi_n(t)$  for any  $n \geq 1$ . Although, for  $n \geq 4$ , the general solution is given in terms of Meijer  $G$ -functions and it would become increasingly difficult to use the condition that  $\phi_n(t)$  is the characteristic function of the  $PN(n, \sigma^2)$  distribution to determine the values of the constants. We can, however, obtain a formula for  $\phi_n(t)$  by using an integral formula involving the Meijer  $G$ -function, given by formula (18) of Section 5.6 of Luke [30] (see (41)). As far as the author is aware, the formula in the following proposition is new.

**Proposition 2.2.** *The characteristic function of the  $\text{PN}(n, \sigma^2)$  distribution is given by*

$$\phi_n(t) = \frac{1}{\pi^{(n-1)/2}} G_{1,n-1}^{n-1,1} \left( \frac{1}{2^{n-2}\sigma^2 t^2} \middle| \frac{1}{2}, \dots, \frac{1}{2} \right).$$

*In the cases  $n = 2, 3$ , this formula simplifies to*

$$\phi_2(t) = \frac{1}{\sqrt{1 + \sigma^2 t^2}} \quad \text{and} \quad \phi_3(t) = \frac{1}{\sqrt{2\pi\sigma^2 t^2}} \exp\left(\frac{1}{4\sigma^2 t^2}\right) K_0\left(\frac{1}{4\sigma^2 t^2}\right).$$

**Proof.** Let  $W_n \sim \text{PN}(n, \sigma^2)$ . Since the  $\text{PN}(n, \sigma^2)$  distribution is symmetric about the origin, it follows that the characteristic function  $\phi_n(t)$  of  $W_n$  is given by

$$\phi_n(t) = \mathbb{E}[e^{itW_n}] = \mathbb{E}[\cos(tW_n)] = 2 \int_0^\infty \frac{1}{(2\pi)^{n/2}\sigma} \cos(tx) G_{0,n}^{n,0} \left( \frac{x^2}{2n\sigma^2} \middle| 0, \dots, 0 \right) dx.$$

Evaluating the integral using (41) and then simplifying using (39) and (38) gives

$$\begin{aligned} \phi_n(t) &= \frac{1}{2^{n/2-1}\pi^{(n-1)/2}\sigma|t|} G_{2,n}^{n,1} \left( \frac{1}{2^{n-2}\sigma^2 t^2} \middle| \frac{1}{2}, 0 \right) \\ &= \frac{1}{\pi^{(n-1)/2}} G_{2,n}^{n,1} \left( \frac{1}{2^{n-2}\sigma^2 t^2} \middle| \frac{1}{2}, \dots, \frac{1}{2} \right) \\ &= \frac{1}{\pi^{(n-1)/2}} G_{1,n-1}^{n-1,1} \left( \frac{1}{2^{n-2}\sigma^2 t^2} \middle| \frac{1}{2}, \dots, \frac{1}{2} \right), \end{aligned}$$

as required. □

### 2.3. Bounds for the solution

Here we solve the product normal Stein equation. For the case  $n = 2$ , the solution can be represented in terms of integrals of modified Bessel functions and takes a relatively simple form. As a result, we are able to obtain uniform bounds for the solution and its lower order derivatives. These bounds are used in the approximation theorems of Section 4. However, for  $n \geq 3$ , the solution takes a less tractable form and we have been unable to bound the solution or any of its derivatives.

#### 2.3.1. $n = 2$

The  $\text{PN}(2, \sigma^2)$  Stein equation (9) is a second order linear differential equation and the homogeneous equation has  $I_0(x/\sigma)$  and  $K_0(x/\sigma)$  as a pair of linearly independent solutions (see (54)). Therefore, one can obtain a solution to the  $\text{PN}(2, \sigma^2)$  Stein equation through a straightforward application of the method of variation of parameters (see Collins [7] for an account of the method). This was done by Gaunt [16], Lemma 3.3 for the more general variance-gamma Stein equation (recall that the  $\text{PN}(2, \sigma^2)$  Stein equation is a special case of the variance-gamma Stein equation), and the following lemma is a special case of that result. We present the solution for the case  $\sigma = 1$ ; we can recover results for the general case by using a simple change of variables.

**Lemma 2.3.** *Suppose  $h : \mathbb{R} \rightarrow \mathbb{R}$  is bounded. Then the unique bounded solution  $f : \mathbb{R} \rightarrow \mathbb{R}$  to the PN(2, 1) Stein equation (9) is given by*

$$f(x) = -K_0(|x|) \int_0^x I_0(y)[h(y) - \text{PN}_2^1 h] dy - I_0(x) \int_x^\infty K_0(|y|)[h(y) - \text{PN}_2^1 h] dy, \tag{17}$$

where  $I_0(x)$  and  $K_0(x)$  are modified Bessel functions; see Appendix B.2.1 for definitions.

**Remark 2.2.** The equality

$$\int_{-\infty}^x K_0(|y|)[h(y) - \text{PN}_2^1 h] dy = - \int_x^\infty K_0(|y|)[h(y) - \text{PN}_2^1 h] dy$$

is very useful when it comes to obtaining bounds for the derivatives of the solution to the Stein equation. The equality ensures that we can restrict our attention to bounding the derivatives in the region  $x \geq 0$ .

By direct calculations it is possible to obtain bounds for the derivatives of the solution of the PN(2, 1) Stein equation. Bounds for the case of general  $\sigma$  then follow by a simple change of variables. In Theorem 2.1 (see Appendix A for the proof), we present uniform bounds for the solution of the PN(2,  $\sigma^2$ ) Stein equation and its first four derivatives in terms of the supremum norms of the derivatives of the test function  $h$ . It is worth noting that bounds for the solution of the variance-gamma Stein equation and its first four derivatives were derived by Gaunt [16], yielding as a special case estimates for the solution of the PN(2,  $\sigma^2$ ) Stein equation. Also, through a new iterative technique, Döbler *et al.* [10] extended this work to obtain bounds on derivatives of arbitrary order of the solution of variance-gamma Stein equation. However, the constants in Theorem 2.1 improve on these for  $n \leq 4$ , and the estimates for  $\|xf(x)\|$ ,  $\|xf'(x)\|$  and  $\|xf''(x)\|$  are new.

By exploiting properties of the zero bias transformation of order  $n$ , we only require bounds for derivatives up to second order for the limit theorems of Section 4; in fact for Corollary 4.1 we only need estimates for the solution and its first derivative. However, for other coupling choices, such as local couplings, we would require bounds on the third derivative to achieve a  $O(m^{-1/2})$  bound (here  $m$  is the ‘sample size’); and for  $O(m^{-1})$  bounds we may require bounds for the fourth derivative (see, for example, Goldstein and Reinert [22], Corollary 3.1).

**Theorem 2.1.** *Suppose that  $h \in C_b^3(\mathbb{R})$ . Then the solution  $f$  of the PN(2,  $\sigma^2$ ) Stein equation (9) and its first four derivatives are bounded as follows*

$$\begin{aligned} \|f\| &= \frac{3}{\sigma} \|h - \text{PN}_2^{\sigma^2} h\|, \\ \|f'\| &= \frac{3}{2\sigma^2} \|h - \text{PN}_2^{\sigma^2} h\|, \\ \|f''\| &= \frac{2\|h'\|}{\sigma^2} + \frac{5}{\sigma^3} \|h - \text{PN}_2^{\sigma^2} h\|, \end{aligned}$$

$$\begin{aligned} \|f^{(3)}\| &= \frac{4\|h''\|}{\sigma^2} + \frac{5\|h'\|}{\sigma^3} + \frac{4.89}{\sigma^4} \|h - \text{PN}_2^{\sigma^2} h\|, \\ \|f^{(4)}\| &= \frac{8\|h^{(3)}\|}{\sigma^2} + \frac{9\|h''\|}{\sigma^3} + \frac{6.81\|h'\|}{\sigma^4} + \frac{15.75}{\sigma^5} \|h - \text{PN}_2^{\sigma^2} h\|, \end{aligned}$$

and we also have

$$\begin{aligned} \|xf(x)\| &\leq \frac{2}{\sigma} \|h - \text{PN}_2^{\sigma^2} h\|, \\ \|xf'(x)\| &\leq \frac{3}{2\sigma^2} \|h - \text{PN}_2^{\sigma^2} h\|, \\ \|xf''(x)\| &\leq \frac{9}{2\sigma^3} \|h - \text{PN}_2^{\sigma^2} h\|. \end{aligned}$$

We now prove a lemma, which gives a bound on the quantities  $\|(A_n f)^{(k)}\|$  that appear in the bounds of Theorem 4.1. In the lemma, we suppose that certain derivatives of the solution of the  $\text{PN}(n, \sigma^2)$  Stein equation (7) exist and are bounded. However, this might not be the case when  $n \geq 3$ , as we have not been able to obtain an analogue of Theorem 2.1 for  $n \geq 3$ ; this is discussed further in Section 2.3.2.

**Lemma 2.4.** *Let  $f$  be the solution of the  $\text{PN}(n, \sigma^2)$  Stein equation (7). Suppose that  $f$  and the test function  $h$  are  $k$  times differentiable. Then*

$$\sigma^2 \|(A_n f)^{(k)}\| \leq \|h^{(k)}\| + \|xf^{(k)}(x)\| + k \|f^{(k-1)}\|, \tag{18}$$

provided that the supremum norms on the right-hand side of (18) are bounded. In particular,

$$\begin{aligned} \sigma^2 \|(A_2 f)'\| &\leq \|h'\| + \left(\frac{3}{\sigma} + \frac{3}{2\sigma^2}\right) \|h - \text{PN}_2^{\sigma^2} h\|, \\ \sigma^2 \|(A_2 f)''\| &\leq \|h''\| + \left(\frac{3}{\sigma^2} + \frac{9}{2\sigma^3}\right) \|h - \text{PN}_2^{\sigma^2} h\|. \end{aligned}$$

**Proof.** The solution  $f$  satisfies the equation  $\sigma^2 A_n f(x) = h(x) - xf(x)$ . By a simple induction,  $\sigma^2 (A_n f)^{(k)}(x) = h^{(k)}(x) + xf^{(k)}(x) + kf^{(k-1)}(x)$ . Applying the triangle inequality now yields (18). To obtain the inequalities for  $\|(A_2 f)'\|$  and  $\|(A_2 f)''\|$ , we use the bounds for  $f$  and its derivatives that were given in Theorem 2.1 to bound the supremum norms on the right-hand side of (18) for the case  $n = 2$  and  $k = 1, 2$ . □

### 2.3.2. $n \geq 3$

We now consider the more challenging  $n \geq 3$  case. Again, for simplicity, we shall take  $\sigma = 1$ ; results for the general case easily follow from a change of variables. As in the case  $n = 2$ , we can obtain a fundamental system for the homogeneous equation, and can thus use variation of parameters to write down a solution to the Stein equation.

**Lemma 2.5.** *Suppose  $h : \mathbb{R} \rightarrow \mathbb{R}$  is bounded, and set  $\tilde{h}(x) = h(x) - \text{PN}_n^1 h$ . Let  $y_k(x) = G_{0,n}^{k,0}((-1)^k x^2/2^n | 0, \dots, 0)$ . Then the function*

$$f(x) = (-1)^n y_1(x) \int_x^\infty \frac{W_1(t)\tilde{h}(t)}{t^{n-1}W(t)} dt + \sum_{k=2}^n (-1)^{n+k} y_k(x) \int_0^x \frac{W_k(t)\tilde{h}(t)}{t^{n-1}W(t)} dt \tag{19}$$

solves the  $\text{PN}(n, 1)$  Stein equation, where  $W$  is the determinant of the matrix

$$\mathbf{M} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{pmatrix} \tag{20}$$

and  $W_k$  is the determinant of the submatrix of  $\mathbf{M}$  obtained by deleting the last row and  $k$ th column.

Moreover, there is at most one bounded solution to the  $\text{PN}(n, 1)$  Stein equation.

The representation of the solution (19), which is given in terms of integrals involving expressions of Meijer  $G$ -functions, does not lend itself easily to the bounding of the solution and its derivatives. This is in contrast to the relatively simple solution (17) that corresponds to the  $n = 2$  case. For this reason, we have not been able to obtain an extension of Theorem 2.1 to  $n \geq 3$ , and this is left as an interesting open problem.

There are various ways one could approach this problem. One would be to use a different method to solve the Stein equation, which may lead to a different (and hopefully simpler) representation of the solution. Alternatively, a first step would be to obtain a simplification of the current representation (19), perhaps by simplifying the determinants  $W_1, \dots, W_n$  and  $W$ . For  $n = 2$ , we have  $y_1(x) = G_{0,2}^{1,0}(-x^2/4 | 0, 0) = I_0(x)$  and  $y_2(x) = G_{0,2}^{2,0}(x^2/4 | 0, 0) = 2K_0(|x|)$  (see (43) and (44)). Therefore, for  $x > 0$ ,

$$\begin{aligned} W(x) &= W(y_1(x), y_2(x)) = y_1(x)y_2'(x) - y_1'(x)y_2(x) \\ &= -2(I_0(x)K_1(x) + I_1(x)K_0(x)) = -\frac{2}{x} \quad (\text{by (36)}) \end{aligned}$$

with a similar formula for  $x < 0$ , and one could attempt to obtain a similar simplification for  $n \geq 3$ . From here, to bound the solution and its derivatives, one could proceed, through rather technical calculations, to bound appropriate expressions involving integrals of Meijer  $G$ -functions. These calculations would be similar, but more technical, than those used to prove Theorem 2.1 (see Appendix A, and Gaunt [20] for the calculations used to obtain the bounds of Appendix C).

If instead of seeking explicit bounds, as was the case in Theorem 2.1, one was content to just prove that the solution and its lower order derivatives were bounded, for sufficiently differentiable test functions  $h$ , then one would only need to investigate the behaviour of the solution (19) in the limits  $x \rightarrow 0$  and  $|x| \rightarrow \infty$ , as the solution and its derivatives are bounded for all other

$x \in \mathbb{R}$ . Asymptotic formulas for the  $G$ -function are available (see, for example, Luke [30], Section 5.10); however, such an analysis may be rather involved, due to occurrence of the Wronskian determinants in the solution (19). Although, the complexity of this analysis would be greatly reduced if a simplification was found for the determinants. Finally, due to Lemma 2.4, if one was using the zero bias transformation of order  $n$  to obtain product normal approximation results, then it would suffice to bound just  $\|f\|$  and  $\|xf'(x)\|$ .

**Proof of Lemma 2.5.** We begin by stating a general result that is given in Kreyszig [27], page 140. Consider the inhomogeneous differential equation

$$q^{(n)}(x) + a_{n-1}(x)q^{(n-1)}(x) + \dots + a_1(x)q'(x) + a_0(x)q(x) = g(x),$$

where the  $a_k(x)$  are  $(n - 1)$ -times differentiable. Suppose that  $q_1(x), \dots, q_n(x)$  form a fundamental solution to the homogeneous equation. Then the general solution is given by

$$q(x) = \sum_{k=1}^n (-1)^{n+k} q_k(x) \int_{a_k}^x \frac{W_k(t)g(t)}{W(t)} dt, \tag{21}$$

where the  $a_k$  are arbitrary constants, and  $W$  and  $W_k$  are determinants of a matrix of the form (20).

We now apply this result to the  $PN(n, 1)$  Stein equation. The homogeneous equation is

$$x^{-1}T^n f(x) - xf(x) = 0 \tag{22}$$

and letting  $f(x) = g((-1)^k x^2/2^n)$  leads to the differential equation

$$T^n g(x) - (-1)^k xg(x) = 0. \tag{23}$$

This is a special case of the Meijer  $G$ -function differential equation (42), and the functions  $G_{0,n}^{k,0}(x|0, \dots, 0)$ ,  $k = 1, \dots, n$ , form a fundamental set of solutions to (23). Consequently, the functions  $y_k(x) = G_{0,n}^{k,0}((-1)^k x^2/2^n|0, \dots, 0)$ ,  $k = 1, \dots, n$ , form a fundamental set of solutions to (22). The differential equation (22) is a  $n$ th order linear differential equation with  $x^{n-1}$  as coefficient of  $f^{(n)}$ , and so from (21) we have that (19) solves the  $PN(n, 1)$  Stein equation.

The values of the arbitrary constants were chosen to allow the solution to be bounded. In fact, in order to have a bounded solution to the  $PN(n, 1)$  Stein equation, one must take  $a_2, \dots, a_n = 0$  and either  $a_1 = \infty$  or  $a_1 = -\infty$ . This is because, due to asymptotic properties of Meijer  $G$ -functions (see Luke [30], Section 5.10), we have that  $y_1(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$  and, for  $k = 2, \dots, n$ ,  $y_k(x) \rightarrow \infty$  as  $x \rightarrow 0$  (asymptotic formulas are only given for  $|x| \rightarrow \infty$  in [30], although asymptotic results for the limit  $x \rightarrow 0$  can be obtained via identity (40)).

Finally, we prove that there is at most one bounded solution to the Stein equation. Suppose  $u$  and  $v$  are bounded solutions to the Stein equation. Define  $w = u - v$ . Then  $w$  is also bounded and is a solution to (22), which has general solution  $w(x) = \sum_{k=1}^n A_k y_k(x)$ , where the  $A_k$  are arbitrary constants. Recall that  $y_1(x) \rightarrow \infty$  as  $|x| \rightarrow \infty$  and, for  $k = 2, \dots, n$ ,  $y_k(x) \rightarrow \infty$  as  $x \rightarrow 0$ . Therefore, for  $w$  to be bounded, we must take  $A_1 = \dots = A_n = 0$ , and thus  $w = 0$  and so there can be at most one bounded solution. □

### 3. The zero bias transformation of order $n$

The zero bias transformation of order  $n$ , given in Definition 1.1, has many useful properties, which we collect in Propositions 3.2 and 3.3, below. We begin by presenting a relationship between the  $W$ -zero bias distribution of order  $n$  and the  $W$ -square bias distribution (see Chen *et al.* [6], pages 34–35, for properties of this distributional transformation). 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  $\mathbb{E}W^2 f(W)$  exists,

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

When  $W$  is non-negative,  $W^\square$  is given by the size bias distribution of  $W^s$ , the size bias distribution of  $W$  (see Pekoz *et al.* [35]). Although, in this section, we will be considering  $W$  with mean zero, and so this attractive result is not available to us.

Before presenting our relationship, we write down a construction of the  $W$ -square bias distribution, for the case that  $W$  is decomposed into a product of independent random variables. This construction will be used in the proof of part (iv) of Proposition 3.3.

**Proposition 3.1.** *Suppose  $W = \prod_{k=1}^n W_k$ , where the  $W_k$  are independent random variables and let  $W_1^\square, \dots, W_n^\square$  be independent random variables with  $W_k^\square$  having the  $W_k$ -square biased distribution, then*

$$W^\square = \prod_{k=1}^n W_k^\square$$

has the  $W$ -square biased distribution.

**Proof.** We prove that the result holds for the case of two products; the extension to a general number of products follows by a straightforward induction. Using independence to obtain the final equality, we have

$$\begin{aligned} \mathbb{E}f(W_1^\square W_2^\square) &= \mathbb{E}[\mathbb{E}[f(W_1^\square W_2^\square) | W_2^\square]] \\ &= \frac{1}{\mathbb{E}W_1^2} \mathbb{E}[W_1^2 \mathbb{E}[f(W_1 W_2^\square) | W_2^\square]] \\ &= \frac{1}{\mathbb{E}W_1^2} \mathbb{E}W_1^2 f(W_1 W_2^\square) \\ &= \frac{1}{\mathbb{E}W_1^2 \mathbb{E}W_2^2} \mathbb{E}W_1^2 W_2^2 f(W_1 W_2) \\ &= \frac{1}{\mathbb{E}W_1^2 W_2^2} \mathbb{E}W_1^2 W_2^2 f(W_1 W_2), \end{aligned}$$

as required. □



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.* [6]. The proof of the relationship utilises the inverse operator  $G_n$  of  $A_n$ , and differs from the approach used in [6].

**Proposition 3.2.** *Let  $W$  be a random variable with zero mean and finite, non-zero variance  $\sigma^2$ , and let  $W^\square$  have the  $W$ -square bias distribution. Let  $U_1, \dots, U_n$  be independent  $U(0, 1)$  random variables, which are independent of  $W^\square$ . Define  $V_n = \prod_{k=1}^n U_k$ . Then, the random variable*

$$W^{*(n)} \stackrel{\mathcal{D}}{=} V_n W^\square$$

has the  $W$ -zero bias distribution of order  $n$ .

**Proof.** Let  $f \in C_c$ , the collection of continuous functions with compact support. In Lemma 2.1, we defined the operator  $G_n g(x) = x \mathbb{E}g(x V_n)$  and showed that  $A_n G_n g(x) = g(x)$  for any  $g$ . We therefore have

$$\sigma^2 \mathbb{E}f(W^{*(n)}) = \sigma^2 \mathbb{E}A_n G_n f(W^{*(n)}) = \mathbb{E}W G_n f(W) = \mathbb{E}W^2 f(V_n W) = \sigma^2 \mathbb{E}f(V_n W^\square).$$

Since the expectation of  $f(W^{*(n)})$  and  $f(V_n W^\square)$  are equal for all  $f \in C_c$ , the random variables  $W^{*(n)}$  and  $V_n W^\square$  must be equal in distribution.  $\square$

In the following proposition, we present some properties of the zero bias transformation of order  $n$ . These properties generalise some of the important properties of the zero bias transformation (see Goldstein and Reinert [22], Lemma 2.1 and Chen *et al.* [6], Proposition 2.1). As was the case for the proof of Proposition 3.2, the proofs of parts (i) and (ii) of the proposition utilise the inverse operator  $G_n$ , and so differ from the approach given in [6] and [22].

**Proposition 3.3.** *Let  $W$  be a mean zero variable with finite, non-zero variance  $\sigma^2$ , and let  $W^{*(n)}$  have the  $W$ -zero biased distribution of order  $n$  in accordance with Definition 1.1.*

(i) *The distribution function of  $W^{*(n)}$  is given by*

$$F_{W^{*(n)}}(w) = \begin{cases} \frac{1}{(n-1)! \sigma^2} \mathbb{E} \left[ W^2 \gamma \left( n, \log \left( \frac{W}{w} \right) \right) \mathbf{1}(W \leq w) \right], & w < 0, \\ 1 - \frac{1}{(n-1)! \sigma^2} \mathbb{E} \left[ W^2 \gamma \left( n, \log \left( \frac{W}{w} \right) \right) \mathbf{1}(W \geq w) \right], & w \geq 0, \end{cases} \tag{24}$$

where  $\gamma(a, x)$  denotes the lower incomplete gamma function  $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$ .

(ii) *The distribution of  $W^{*(n)}$  is unimodal about zero with density*

$$f_{W^{*(n)}}(w) = \begin{cases} -\frac{1}{(n-1)! \sigma^2} \mathbb{E} \left[ W \left( \log \left( \frac{W}{w} \right) \right)^{n-1} \mathbf{1}(W \leq w) \right], & w < 0, \\ \frac{1}{(n-1)! \sigma^2} \mathbb{E} \left[ W \left( \log \left( \frac{W}{w} \right) \right)^{n-1} \mathbf{1}(W \geq w) \right], & w \geq 0. \end{cases} \tag{25}$$

It follows that the support of  $W^{*(n)}$  is the closed convex hull of the support of  $W$  and that  $W^{*(n)}$  is bounded whenever  $W$  is bounded.

Suppose now that  $W$  is symmetric, then the density of  $W^{*(n)}$  is given by

$$f_{W^{*(n)}}(w) = \frac{1}{(n-1)!\sigma^2} \mathbb{E} \left[ W \left( \log \left| \frac{W}{w} \right| \right)^{n-1} \mathbf{1}(W \geq w) \right] \tag{26}$$

$$= -\frac{1}{(n-1)!\sigma^2} \mathbb{E} \left[ W \left( \log \left| \frac{W}{w} \right| \right)^{n-1} \mathbf{1}(W \leq w) \right]. \tag{27}$$

In particular, the zero bias transformation of order  $n$  preserves symmetry.

(iii) For  $p \geq 0$  we have

$$\mathbb{E}(W^{*(n)})^p = \frac{\mathbb{E}W^{p+2}}{\sigma^2(p+1)^n} \quad \text{and} \quad \mathbb{E}|W^{*(n)}|^p = \frac{\mathbb{E}|W|^{p+2}}{\sigma^2(p+1)^n}.$$

(iv) Suppose  $W = \prod_{k=1}^n W_k$ , where the  $W_k$  are independent random variables with zero mean and finite, non-zero variance and  $W_1^*, \dots, W_n^*$  are independent random variables with  $W_k^*$  having the  $W_k$ -zero biased distribution. Then

$$W^{*(n)} \stackrel{\mathcal{D}}{=} \prod_{k=1}^n W_k^* \tag{28}$$

has the  $W$ -zero biased distribution of order  $n$ .

(v) For  $c \in \mathbb{R}$ ,  $cW^{*(n)}$  has the  $cW$ -zero biased distribution of order  $n$ .

**Proof.** (i) In the proof of Proposition 3.2, we showed that  $\sigma^2 \mathbb{E}f(W^{*(n)}) = \mathbb{E}W^2 f(V_n W)$  for all bounded functions  $f$ , where  $V_n$  is the product of  $n$  many independent  $U(0, 1)$  random variables. By taking  $f(x) = \mathbf{1}(x \leq w)$ , we have

$$F_{W^{*(n)}}(w) = \frac{1}{\sigma^2} \mathbb{E}[W^2 \mathbf{1}(V_n W \leq w)] = 1 - \frac{1}{\sigma^2} \mathbb{E}[W^2 \mathbf{1}(V_n W \geq w)],$$

as  $\mathbb{E}W^2 = \sigma^2$ . Formula (24) now follows on noting that  $-\log(V_n)$  follows the Gamma( $n, 1$ ) distribution.

(ii) We verify that (25) holds for  $w < 0$ ; the proof for  $w \geq 0$  is similar. Suppose  $a < 0$ , then the function  $\gamma(n, \log(a/w))$  is differentiable on  $(a, 0)$ , with derivative

$$\frac{d}{dw} \left[ \gamma \left( n, \log \left( \frac{a}{w} \right) \right) \right] = -\frac{1}{a} \left( \log \left( \frac{a}{w} \right) \right)^{n-1}. \tag{29}$$

Using (29) and dominated convergence, we have, for  $w < 0$ ,

$$\begin{aligned} f_{W^{*(n)}}(w) &= \frac{1}{(n-1)!\sigma^2} \mathbb{E} \left[ W^2 \frac{d}{dw} \left[ \gamma \left( n, \log \left( \frac{W}{w} \right) \right) \mathbf{1}(W \leq w) \right] \right] \\ &= -\frac{1}{(n-1)!\sigma^2} \mathbb{E} \left[ W \left( \log \left( \frac{W}{w} \right) \right)^{n-1} \mathbf{1}(W \leq w) \right]. \end{aligned}$$

It follows from (25) that  $f_{W^{*(n)}}$  is increasing for  $w < 0$  and decreasing for  $w > 0$ , and is thus unimodal about zero.

We now consider the case that  $W$  is symmetric, and verify formulas (26) and (27) for  $f_{W^{*(n)}}(w)$ . That (26) and (27) are equal follows since  $W$  is symmetric. Formulas (26) and (27) are certainly true for  $w \geq 0$  and  $w < 0$ , respectively. But (26) and (27) are equal and so both formulas must hold for all  $w \in \mathbb{R}$ . Finally, from (26) and (27) we have  $f_{W^{*(n)}}(w) = f_{W^{*(n)}}(-w)$ , hence  $W^{*(n)}$  is also symmetric.

(iii) Substitute  $w^{p+1}/(p+1)$  and  $w^{p+1} \operatorname{sgn}(w)/(p+1)$  for  $f(w)$  in the characterising equation (11).

(iv) Let  $U_1, \dots, U_n$  and  $V_n$  be defined as in Proposition 3.2. Then, from Propositions 3.1 and 3.2, we have

$$W^{*(n)} \stackrel{\mathcal{D}}{=} V_n W^\square \stackrel{\mathcal{D}}{=} V_n \prod_{k=1}^n W_k^\square = \prod_{k=1}^n U_k W_k^\square \stackrel{\mathcal{D}}{=} \prod_{k=1}^n W_k^*.$$

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

$$\mathbb{E}cWg(cW) = \mathbb{E}W\tilde{g}(W) = \sigma^2 \mathbb{E}A_n \tilde{g}(W^{*(n)}) = (c\sigma)^2 \mathbb{E}A_n g(cW^{*(n)}).$$

Hence,  $cW^{*(n)}$  has the  $cW$ -zero bias distribution of order  $n$ . □

### 4. Zero bias approach bounds for the product of two independent normal distributions

We now illustrate how the product normal Stein equation and zero bias transformation may be used together to assess distributional distances for statistics that are asymptotically distributed as the product of two independent central normal random variables. Theorem 4.1, which is a natural generalisation of Theorem 3.1 of Goldstein and Reinert [22], shows how the distance between an arbitrary mean zero, finite variance random variable  $W$  and a product normal random variable with the same variance can be bounded by the distance between  $W$  and a variate  $W^{*(n)}$  with the  $W$ -zero biased distribution of order  $n$  defined on a joint space. The bounds in Theorem 4.1 only hold if  $(A_n f)'(x)$  and  $(A_n f)''(x)$  exist and are bounded. Of course, this might not be the case when  $n \geq 3$ .

**Theorem 4.1.** *Let  $W$  be a mean zero random variable with variance  $\sigma^2$ . Suppose that  $(W, W^{*(n)})$  is given on a joint probability space so that  $W^{*(n)}$  has the  $W$ -zero biased distribution of order  $n$ . Then*

$$|\mathbb{E}h(W) - \operatorname{PN}_n^{\sigma^2} h| \leq \sigma^2 \|(A_n f)'\| \mathbb{E}|W - W^{*(n)}|, \tag{30}$$

where  $f$  is the solution of the  $\operatorname{PN}(n, \sigma^2)$  Stein equation (7).

Suppose now that  $W = \prod_{k=1}^n W_k$ , where the  $W_k$  are independent. Then

$$\begin{aligned}
 |\mathbb{E}h(W) - \text{PN}_n^{\sigma^2} h| &\leq \sigma^2 \|(A_n f)'\| \sqrt{\mathbb{E}[\mathbb{E}(W - W^{*(n)} | W_1, \dots, W_n)^2]} \\
 &\quad + \frac{\sigma^2}{2} \|(A_n f)''\| \mathbb{E}(W - W^{*(n)})^2.
 \end{aligned}
 \tag{31}$$

The quantities  $\|(A_2 f)'\|$  and  $\|(A_2 f)''\|$  are bounded in Lemma 2.4.

**Proof.** We prove the second bound; the first bound is obtained by a similar but simpler calculation. Using equation (7), we have

$$\mathbb{E}h(W) - \text{PN}_n^{\sigma^2} h = \sigma^2 [A_n f(W) - Wf(W)] = \sigma^2 \mathbb{E}[A_n f(W) - A_n f(W^{*(n)})].$$

By Taylor expansion, we have

$$|\mathbb{E}h(W) - \text{PN}_n^{\sigma^2} h| \leq \sigma^2 |\mathbb{E}[(W - W^{*(n)})(A_n f)'(W)]| + \frac{\sigma^2}{2} \|(A_n f)''\| \mathbb{E}(W - W^{*(n)})^2.$$

For the first term, we condition on  $W_1, \dots, W_n$  and then apply the Cauchy–Schwarz inequality to obtain

$$\begin{aligned}
 |\mathbb{E}[(W - W^{*(n)})(A_n f)'(W)]| &\leq \mathbb{E}[(A_n f)'(W) \mathbb{E}[W - W^{*(n)} | W_1, \dots, W_n]] \\
 &\leq \|(A_n f)'\| \sqrt{\mathbb{E}[\mathbb{E}(W - W^{*(n)} | W_1, \dots, W_n)^2]},
 \end{aligned}$$

which yields the desired bound. □

We now use Theorem 4.1 to obtain two bounds for the error in approximating a statistic that has an asymptotic PN(2, 1) distribution by its limiting distribution.

**Corollary 4.1.** *Suppose  $X, X_1, \dots, X_m, Y, Y_1, \dots, Y_n$  are independent random variables with zero mean, unit variance and bounded absolute third moment, with  $X_i \stackrel{D}{=} X$  for all  $i = 1, \dots, m$  and  $Y_j \stackrel{D}{=} Y$  for all  $j = 1, \dots, n$ . Let  $W_1 = \sum_{i=1}^m X_i, W_2 = \sum_{j=1}^n Y_j$  and set  $W = \frac{1}{\sqrt{mn}} W_1 W_2$ . Then, for  $h \in C_b^1(\mathbb{R})$ , we have*

$$|\mathbb{E}h(W) - \text{PN}_2^1 h| \leq \frac{13}{8} \left( \frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right) \left[ \|h'\| + \frac{9}{2} \|h - \text{PN}_2^1 h\| \right] \mathbb{E}|X|^3 |Y|^3.$$

**Proof.** We will apply bound (30) of Theorem 4.1, and so we just need to bound  $\mathbb{E}|W - W^{*(2)}|$ . By part (iv) of Proposition 3.3 and Lemma 1.1, we have that  $W^{*(2)} = \frac{1}{\sqrt{mn}} W_1^* W_2^*$  where  $W_1^* = W_1 - X_I + X_I^*$  and  $W_2^* = W_2 - Y_J + Y_J^*$ . By the independence of the collections  $X_1, \dots, X_m$

and  $Y_1, \dots, Y_n$ , we have

$$\begin{aligned} \mathbb{E}|W - W^{*(2)}| &= \frac{1}{\sqrt{mn}} \mathbb{E}|(X_I - X_I^*)W_2 + (Y_J - Y_J^*)W_1 - (X_I - X_I^*)(Y_J - Y_J^*)| \\ &\leq \frac{1}{\sqrt{mn}} \{(\mathbb{E}|X_I| + \mathbb{E}|X_I^*|)\mathbb{E}|W_2| + (\mathbb{E}|Y_J| + \mathbb{E}|Y_J^*|)\mathbb{E}|W_1| \\ &\quad + (\mathbb{E}|X_I| + \mathbb{E}|X_I^*|)(\mathbb{E}|Y_J| + \mathbb{E}|Y_J^*|)\}. \end{aligned}$$

By part (iii) of Proposition 3.3, we have that  $\mathbb{E}|X_I| \leq \frac{1}{2}\mathbb{E}|X|^3$ . Using this fact and that  $\mathbb{E}|W_1| \leq \sqrt{m}$  and  $\mathbb{E}|W_2| \leq \sqrt{n}$  gives

$$\begin{aligned} \mathbb{E}|W - W^{*(2)}| &\leq \frac{1}{\sqrt{m}}(1 + \mathbb{E}|X|^3) + \frac{1}{\sqrt{n}}(1 + \mathbb{E}|Y|^3) + \frac{1}{\sqrt{mn}}(1 + \mathbb{E}|X|^3)(1 + \mathbb{E}|Y|^3) \\ &\leq \frac{3\mathbb{E}|X|^3}{2\sqrt{m}} + \frac{3\mathbb{E}|Y|^3}{2\sqrt{n}} + \frac{9\mathbb{E}|X|^3\mathbb{E}|Y|^3}{4\sqrt{mn}} \leq \frac{13}{8} \left( \frac{1}{\sqrt{m}} + \frac{1}{\sqrt{n}} \right) \mathbb{E}|X|^3\mathbb{E}|Y|^3, \end{aligned}$$

where we used the inequalities  $\mathbb{E}|X|^3 \geq 1$  and  $(mn)^{-1/2} \leq \frac{1}{2}(m^{-1/2} + n^{-1/2})$  to simplify the bound. Using Lemma 2.4 to bound  $\|(A_2 f)'\|$  completes the proof.  $\square$

When  $\mathbb{E}X^3 = \mathbb{E}Y^3 = 0$  we can use the second bound of Theorem 4.1 to obtain a bound on the rate of convergence of  $W$  to its limiting 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 [22], 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 4.2.** *Let the  $X_i, Y_j$  and  $W$  be defined as in Corollary 4.1, but with the extra condition that  $\mathbb{E}X^3 = \mathbb{E}Y^3 = 0$ , and  $\mathbb{E}X^4, \mathbb{E}Y^4 < \infty$ . Then, for  $h \in C_b^2(\mathbb{R})$ , we have*

$$|\mathbb{E}h(W) - \text{PN}_2^1 h| \leq \left( \frac{1}{m} + \frac{1}{n} \right) \left[ \frac{7}{2} \|h''\| + \|h'\| + \frac{123}{4} \|h - \text{PN}_2^1 h\| \right] \mathbb{E}X^4 \mathbb{E}Y^4. \quad (32)$$

**Proof.** We make use of the second bound in Theorem 4.1, and so require bounds on the quantities  $\sqrt{\mathbb{E}[\mathbb{E}(W - W^{*(2)}|W_1, W_2)^2]}$  and  $\mathbb{E}(W - W^{*(2)})^2$ . We have

$$\begin{aligned} &\mathbb{E}[W - W^{*(2)}|W_1, W_2] \\ &= \frac{1}{\sqrt{mn}} \mathbb{E}[W_1 W_2 - W_1^* W_2^*|W_1, W_2] \\ &= \frac{1}{\sqrt{mn}} \mathbb{E}[(X_I - X_I^*)W_2 + (Y_J - Y_J^*)W_1 - (X_I - X_I^*)(Y_J - Y_J^*)|W_1, W_2] \\ &= \frac{1}{\sqrt{mn}} \{W_2 \mathbb{E}[X_I|W_1] + W_1 \mathbb{E}[Y_J|W_2] - \mathbb{E}[X_I|W_1] \mathbb{E}[Y_J|W_2]\} \\ &= \frac{1}{\sqrt{mn}} \left( \frac{1}{m} + \frac{1}{n} - \frac{1}{mn} \right) W_1 W_2, \end{aligned}$$

where we used that  $X_I^*$  and  $W_1$  are independent and from part (iii) of Proposition 3.3 that  $\mathbb{E}X_I^* = \frac{1}{2}\mathbb{E}X^3 = 0$  to obtain the third equality, and that  $\mathbb{E}(X_I|W_1) = \frac{1}{m}W_1$  to obtain the final equality. As  $\mathbb{E}W_1^2 = m$  and  $\mathbb{E}W_2^2 = n$ , we have

$$\sqrt{\mathbb{E}[\mathbb{E}(W - W^{*(2)}|W_1, W_2)^2]} = \frac{1}{m} + \frac{1}{n} - \frac{1}{mn} < \frac{1}{m} + \frac{1}{n}.$$

We now bound the second term:

$$\begin{aligned} \mathbb{E}(W - W^{*(2)})^2 &= \frac{1}{mn} \mathbb{E}[\{(X_I - X_I^*)W_2 + (Y_J - Y_J^*)W_1 - (X_I - X_I^*)(Y_J - Y_J^*)\}^2] \\ &\leq \frac{3}{mn} [\mathbb{E}(X_I - X_I^*)^2 \mathbb{E}W_2^2 + \mathbb{E}(Y_J - Y_J^*)^2 \mathbb{E}W_1^2 \\ &\quad + \mathbb{E}(X_I - X_I^*)^2 \mathbb{E}(Y_J - Y_J^*)^2], \end{aligned}$$

where we used that  $(a + b + c)^2 \leq 3(a^2 + b^2 + c^2)$  to obtain the inequality. By part (iii) of Proposition 3.3,

$$\mathbb{E}(X_I - X_I^*)^2 = \mathbb{E}X_I^2 + \mathbb{E}(X_I^*)^2 = 1 + \frac{1}{3}\mathbb{E}X^4 \leq \frac{4}{3}\mathbb{E}X^4,$$

and therefore

$$\begin{aligned} \mathbb{E}(W - W^{*(2)})^2 &\leq \frac{4\mathbb{E}X^4}{m} + \frac{4\mathbb{E}Y^4}{n} + \frac{16\mathbb{E}X^4\mathbb{E}Y^4}{3mn} \\ &\leq \left(4 + \frac{8}{3}\right) \left(\frac{1}{m} + \frac{1}{n}\right) \mathbb{E}X^4\mathbb{E}Y^4 < 7 \left(\frac{1}{m} + \frac{1}{n}\right) \mathbb{E}X^4\mathbb{E}Y^4, \end{aligned}$$

where we used the inequalities  $\mathbb{E}X^4 \geq 1$  and  $(mn)^{-1} \leq \frac{1}{2}(m^{-1} + n^{-1})$  to obtain the second inequality. Applying Lemma 2.4 to bound  $\|(A_2 f)'\|$  and  $\|(A_2 f)''\|$  and then using the inequality  $\mathbb{E}X^4 \geq 1$  to simplify the resulting bound yields (32). □

### 5. Bounds for a general $n$ via the multivariate normal Stein equation

In the previous section, we saw that the product normal Stein equation and the zero bias transformation of order  $n$  can be applied together to derive bounds for product normal approximation, provided we have bounds for the appropriate lower order derivatives of the solution of the Stein equation. However, in this paper, we have only been able to achieve this for the case of  $n = 2$  products. A similar problem was encountered by Arras *et al.* [1], in which an  $n$ th order Stein equation was obtained for a general linear combination of  $n$  independent gamma random variables. They were also unable to bound the appropriate lower order derivatives of the solution, but were still able to prove approximation theorems by bypassing the Stein equation (see Section 3 of [1]).

In this section, we take the same philosophy as [1] and, by bypassing the product normal Stein equation, we are able to prove product normal approximation theorems for general  $n$ . In a recent paper, Gaunt [18] introduced a general method for proving approximation theorems in which the limit distribution can be represented as a functions of multivariate normal random variables (see also Gaunt *et al.* [21], in which the technique is applied for the case of a chi-square limit).

Let  $X_{1,1}, \dots, X_{n,1}, \dots, X_{1,d}, \dots, X_{n,d}$  be independent random variables with mean zero and unit variance. Define  $W_j = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_{ij}$  and let  $\mathbf{W} = (W_1, \dots, W_d)^T$ , which, by the central limit theorem, converges to the standard  $d$ -dimensional multivariate normal random variable  $\mathbf{Z}$ . In [18], general bounds were given for distributional distance between  $g(\mathbf{W})$  and  $g(\mathbf{Z})$ , where  $g: \mathbb{R}^d \rightarrow \mathbb{R}$  satisfies certain differentiability and growth rate conditions. The approach bypasses the Stein equation for the limit distribution  $g(\mathbf{Z})$  by using the multivariate normal Stein equation (see Goldstein and Rinott [26]) to approximate the random vector  $\mathbf{W}$  and then links this approximation to one for  $g(\mathbf{W})$  by a suitably chosen test function. This approach allows a large class of limit distributions to be treated within one framework, including products of independent normal random variables. We now state general bounds (Theorems 3.2 and 3.4 of Gaunt [18]), which we shall then apply to obtain some  $\text{PN}(n, 1)$  approximation theorems that hold for general  $n \geq 1$ .

**Theorem 5.1 ([18], Theorem 3.2).** *Let  $X_{1,1}, \dots, X_{n,1}, \dots, X_{1,d}, \dots, X_{n,d}$  be independent random variables with  $\mathbb{E}X_{ij}^k = \mathbb{E}Z^k$  for all  $1 \leq i \leq n_j, 1 \leq j \leq d, 1 \leq k \leq p$ , where  $Z \sim N(0, 1)$ . Suppose also that  $\mathbb{E}|X_{ij}|^{r_i+p+1} < \infty$  for all  $i, j$  and  $l$ . Let  $P(\mathbf{w}) = A + B \sum_{i=1}^d |w_i|^{r_i}$ , where  $A, B$  and  $r_1, \dots, r_d$  are non-negative constants. Suppose  $g \in C^p(\mathbb{R}^d)$  is such that  $|\frac{\partial^k g(\mathbf{w})}{\partial w_j^k}|^{p/k} \leq P(\mathbf{w})$  for all  $1 \leq j \leq d, 1 \leq k \leq p$ . Then, for  $h \in C_b^p(\mathbb{R})$ ,*

$$\begin{aligned} & |\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z}))| \\ & \leq \frac{p+1}{p!} h_p \sum_{j=1}^d \sum_{i=1}^{n_j} \frac{1}{n_j^{(p+1)/2}} \left[ A \mathbb{E}|X_{ij}|^{p+1} \right. \\ & \quad \left. + B \sum_{k=1}^d 2^{r_k} \left( 2^{r_k} \mathbb{E}|X_{ij}|^{p+1} \mathbb{E}|W_j|^{r_k} + \frac{2^{r_k}}{n_j^{r_k/2}} \mathbb{E}|X_{ij}|^{r_k+p+1} + \mathbb{E}|Z|^{r_k+1} \mathbb{E}|X_{ij}|^{p+1} \right) \right], \end{aligned}$$

where  $h_p = \sum_{k=1}^p \left\{ \begin{matrix} p \\ k \end{matrix} \right\} \|h^{(k)}\|$  and the Stirling number of the second kind are given by  $\left\{ \begin{matrix} p \\ k \end{matrix} \right\} = \frac{1}{k!} \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} j^p$  (see Olver *et al.* [33]).

**Theorem 5.2 ([18], Theorem 3.4).** *Let  $X_{1,1}, \dots, X_{n,1}, \dots, X_{1,d}, \dots, X_{n,d}$  be independent random variables with  $\mathbb{E}X_{ij}^k = \mathbb{E}Z^k$  for all  $1 \leq i \leq n_j, 1 \leq j \leq d, 1 \leq k \leq p$ . Suppose also that  $\mathbb{E}|X_{ij}|^{r_i+p+2} < \infty$  for all  $i, j$  and  $l$ . Let  $P(\mathbf{w}) = A + B \sum_{i=1}^d |w_i|^{r_i}$ , where  $A, B$  and  $r_1, \dots, r_d$  are non-negative constants. Suppose  $g \in C^{p+2}(\mathbb{R}^d)$  is such that  $|\frac{\partial^k g(\mathbf{w})}{\partial w_j^k}|^{(p+2)/k} \leq P(\mathbf{w})$  for all  $1 \leq j \leq d, 1 \leq k \leq p+2$ . Suppose further that  $g$  is an even function ( $g(\mathbf{w}) = g(-\mathbf{w})$  for all*

$\mathbf{w} \in \mathbb{R}^d$ ). Then, for  $h \in C_b^{p+2}(\mathbb{R})$ ,

$$\begin{aligned} & |\mathbb{E}h(g(\mathbf{W})) - \mathbb{E}h(g(\mathbf{Z}))| \\ & \leq \frac{h_{p+2}}{p!} \left\{ \frac{1}{p+2} \sum_{j=1}^d \sum_{i=1}^{n_j} \frac{1}{n_j^{p/2+1}} \left( \frac{p+2}{p+1} + |\mathbb{E}X_{ij}^{p+1}| \right) \left[ A\mathbb{E}|X_{ij}|^{p+2} \right. \right. \\ & \quad \left. \left. + B \sum_{k=1}^d 2^{r_k} \left( 2^{r_k} \mathbb{E}|X_{ij}|^{p+2} \mathbb{E}|W_j|^{r_k} + \frac{2^{r_k}}{n_j^{r_k/2}} \mathbb{E}|X_{ij}|^{r_k+p+2} + \mathbb{E}|Z|^{r_k} \mathbb{E}|X_{ij}|^{p+2} \right) \right] \right. \\ & \quad \left. + \frac{3}{2} \sum_{j=1}^d \sum_{i=1}^{n_j} \frac{|\mathbb{E}X_{ij}^{p+1}|}{n_j^{(p+1)/2}} \sum_{k=1}^d \sum_{l=1}^{n_k} \frac{1}{n_k^{3/2}} \left[ A\mathbb{E}|X_{ij}|^3 \right. \right. \\ & \quad \left. \left. + B \sum_{k=1}^d 3^{r_k} \left( 2^{r_k} \mathbb{E}|X_{ij}|^3 \mathbb{E}|W_j|^{r_k} + \frac{2^{r_k}}{n_j^{r_k/2}} \mathbb{E}|X_{ij}|^{r_k+3} + 2\mathbb{E}|Z|^{r_k+1} \mathbb{E}|X_{ij}|^3 \right) \right] \right\}. \end{aligned}$$

A  $\text{PN}(d, 1)$  random variable can be represented as  $g(\mathbf{Z})$ , where  $g(\mathbf{w}) = \prod_{j=1}^d w_j$  is infinitely often differentiable and has derivatives of polynomial growth as  $|\mathbf{w}| \rightarrow \infty$ . Also, when  $d$  is even, the function  $g$  is even. We can therefore apply Theorems 5.1 and 5.2 to obtain the following bounds for product normal approximation.

**Corollary 5.1.** Fix  $d > 1$ . Let  $X, X_{1,1}, \dots, X_{n,1}, \dots, X_{1,d}, \dots, X_{n,d}$  be i.i.d. random variables with  $\mathbb{E}X^k = \mathbb{E}Z^k$  for all  $1 \leq k \leq p$ , and such that  $\mathbb{E}|X|^{2p+1} < \infty$ . Define  $W = \prod_{j=1}^d W_j$ , where  $W_j = \frac{1}{\sqrt{n_j}} \sum_{i=1}^{n_j} X_{ij}$ . Then, for  $h \in C_b^p(\mathbb{R})$ ,

$$\begin{aligned} |\mathbb{E}h(W) - \text{PN}_d^1 h| & \leq \frac{2^p(p+1)}{p!} h_p \sum_{j=1}^d \frac{1}{n_j^{(p-1)/2}} \left[ 2^p \mathbb{E}|X|^{p+1} \mathbb{E}|W_1|^p \right. \\ & \quad \left. + \mathbb{E}|Z|^{p+1} \mathbb{E}|X|^{p+1} + \frac{2^p}{n_j^{p/2}} \mathbb{E}|X|^{2p+1} \right]. \end{aligned} \tag{33}$$

**Proof.** We apply Theorem 5.1 with  $g(\mathbf{w}) = \prod_{j=1}^d w_j$ . For any  $j = 1, \dots, d$  we have that  $\frac{\partial g(\mathbf{w})}{\partial w_j} = \prod_{l \neq j} w_l$  and  $\frac{\partial^2 g(\mathbf{w})}{\partial w_j^2} = 0$ . Therefore, we can take  $P(\mathbf{w}) = \frac{1}{d} \sum_{j=1}^d |w_j|^p$ . Applying the bound of Theorem 5.1 with  $A = 0$ ,  $B = \frac{1}{d}$  and  $r_1 = \dots = r_d = p$  yields (33).  $\square$

**Corollary 5.2.** Let  $d \geq 2$  be even. Let  $X, X_{1,1}, \dots, X_{n,1}, \dots, X_{1,d}, \dots, X_{n,d}$  be i.i.d. random variables with  $\mathbb{E}X^k = \mathbb{E}Z^k$  for all  $1 \leq k \leq p$ , and such that  $\mathbb{E}|X|^{2p+4} < \infty$ . Let  $W$  be defined as



in Corollary 5.1. Then, for  $h \in C_b^{p+2}(\mathbb{R})$ ,

$$\begin{aligned}
 |\mathbb{E}h(W) - \text{PN}_d^1 h| &\leq \frac{h_{p+2}}{p!} \left\{ \frac{2^{p+2}}{p+2} \sum_{j=1}^d \frac{1}{n_j^{p/2}} \left( \frac{p+2}{p+1} + |\mathbb{E}X^{p+1}| \right) \left[ 2^{p+2} \mathbb{E}|X|^{p+2} \mathbb{E}|W_1|^{p+2} \right. \right. \\
 &\quad \left. \left. + \mathbb{E}|Z|^{p+2} \mathbb{E}|X|^{p+2} + \frac{2^{p+2}}{n_j^{p/2+1}} \mathbb{E}|X|^{2p+4} \right] \right. \\
 &\quad \left. + \frac{3^{p+3}}{2} |\mathbb{E}X^{p+1}| \sum_{j,k=1}^d \frac{1}{n_j^{(p-1)/2} n_k^{1/2}} \left[ 2^{p+2} \mathbb{E}|X|^3 \mathbb{E}|W_1|^{p+2} \right. \right. \\
 &\quad \left. \left. + 2 \mathbb{E}|Z|^{p+3} \mathbb{E}|X|^3 + \frac{2^{p+2}}{n_j^{p/2+1}} \mathbb{E}|X|^{p+5} \right] \right\}. \tag{34}
 \end{aligned}$$

**Proof.** Here  $d$  is even, so we apply Theorem 5.2. Arguing as in the proof of Corollary 5.1, we take  $P(\mathbf{w}) = \frac{1}{d} \sum_{j=1}^d |w_j|^{p+2}$ . On applying the bound of Theorem 5.1 with  $A = 0$ ,  $B = \frac{1}{d}$  and  $r_1 = \dots = r_d = p + 2$  we obtain (34).  $\square$

**Remark 5.1.** From Corollary 5.1, we have a bound on the rate of convergence of  $W$  to the  $\text{PN}(d, 1)$  of order  $n_1^{-(p-1)/2} + \dots + n_d^{-(p-1)/2}$  for smooth test functions, provided that first  $p$  moments of  $X$  and the standard normal distribution agree. From Corollary 5.2, we see that, for even  $d$ , in which case  $g(\mathbf{w}) = \prod_{j=1}^d w_j$  is an even function, this rate of convergence improves to order  $n_1^{-p/2} + \dots + n_d^{-p/2}$ .

**Remark 5.2.** It is instructive to compare the bound of Corollary 4.2 that was obtained using the zero bias coupling approach and the bounds of Corollaries 5.1 and 5.2 (in the case  $d = 2$ ) that were obtained by bypassing the product normal Stein equation. The bound of Corollary 4.2 is of order  $n_1^{-1} + n_2^{-1}$  and was derived under the assumption that the  $X_i$  and  $Y_i$  had third moments equal to zero. Applying (33) with this setup gives a bound of the same order, although we must impose stronger moment assumptions (bounded absolute seventh moment, rather than bounded fourth moment) and stronger conditions on the test functions (we require  $h \in C_b^3(\mathbb{R})$ , instead of  $h \in C_b^2(\mathbb{R})$ ). Whilst the bound of Corollary 4.2 performs better in this example, the proof relies heavily on the assumption of third moments equal to zero. However, even when the third moments are non-zero, we can obtain an  $O(n_1^{-1} + n_2^{-1})$  bound from (34), under the assumption of bounded eighth moments are tests functions from the class  $C_b^4(\mathbb{R})$ .

## Appendix A: Proof of Theorem 2.1

We begin by obtaining formulas for the first four derivatives of the solution (17).

**Lemma A.1.** Suppose  $h \in C_b^3(\mathbb{R})$  and let  $\tilde{h}(x) = h(x) - \text{PN}_2^1 h$ . Then the first four derivatives of the solution (17) of the  $\text{PN}(2, 1)$  Stein equation (9), in the region  $x > 0$ , are given by

$$\begin{aligned} f'(x) &= -K_0'(x) \int_0^x I_0(y)\tilde{h}(y) \, dy - I_0'(x) \int_x^\infty K_0(y)\tilde{h}(y) \, dy, \\ f''(x) &= \frac{\tilde{h}(x)}{x} - K_0''(x) \int_0^x I_0(y)\tilde{h}(y) \, dy - I_0''(x) \int_x^\infty K_0(y)\tilde{h}(y) \, dy, \\ f^{(3)}(x) &= \frac{h'(x)}{x} - \frac{2\tilde{h}(x)}{x^2} - K_0^{(3)}(x) \int_0^x I_0(y)\tilde{h}(y) \, dy - I_0^{(3)}(x) \int_x^\infty K_0(y)\tilde{h}(y) \, dy, \\ f^{(4)}(x) &= \frac{h''(x)}{x} - \frac{3h'(x)}{x^2} + \left(\frac{6}{x^3} + \frac{1}{x}\right)\tilde{h}(x) - K_0^{(4)}(x) \int_0^x I_0(y)\tilde{h}(y) \, dy \\ &\quad - I_0^{(4)}(x) \int_x^\infty K_0(y)\tilde{h}(y) \, dy. \end{aligned}$$

**Proof.** We will make repeated use 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  $a(x) \leq y \leq b(x)$ ,  $x_0 \leq x \leq x_1$ , and the functions  $a(x)$  and  $b(x)$  are continuous and have continuous derivatives for  $x_0 \leq x \leq x_1$ , then for  $x_0 \leq x \leq x_1$ ,

$$\frac{d}{dx} \int_{a(x)}^{b(x)} u(y, x) \, dy = \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} u(y, x) \, dy + u(b, x) \frac{db}{dx} - u(a, x) \frac{da}{dx}. \tag{35}$$

We will also make use of the identity

$$I_0(x)K_1(x) + I_1(x)K_0(x) = \frac{1}{x} \tag{36}$$

(see Olver *et al.* [33]), as well as the formulas (48)–(53) for the first three derivatives of  $I_0(x)$  and  $K_0(x)$ .

It is easy to compute the first and second derivatives by applying (35), the differentiation formulas (48) and (49) and identity (36). The calculation of the third derivative is still straightforward but a little longer. We differentiate the formula for the second derivative using (35) to obtain

$$\begin{aligned} f^{(3)}(x) &= \frac{h'(x)}{x} - \frac{\tilde{h}(x)}{x^2} - K_0^{(3)}(x) \int_0^x I_0(y)\tilde{h}(y) \, dy \\ &\quad - I_0^{(3)}(x) \int_x^\infty K_0(y)\tilde{h}(y) \, dy + \tilde{h}(x) [-I_0(x)K_0''(x) + K_0(x)I_0''(x)]. \end{aligned}$$

Using the differentiation formulas (50) and (51) and identity (36) allows us to calculate the term in the brackets (\*) from the above expression

$$(*) = -x^\nu I_0(x)K_0''(x) + x^\nu K_0(x)I_0''(x) = -\frac{1}{x} [I_0(x)K_1(x) + I_1(x)K_0(x)] = \frac{1}{x^2}.$$

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 (35) to obtain

$$f^{(4)}(x) = \frac{h''(x)}{x} - \frac{3h'(x)}{x^2} + \frac{4\tilde{h}(x)}{x^3} - K_0^{(4)}(x) \int_0^x I_0(y)\tilde{h}(y) dy - I_0^{(4)}(x) \int_x^\infty K_0(y)\tilde{h}(y) dy + \tilde{h}(x)[-I_0(x)K_0^{(3)}(x) + K_0(x)I_0^{(3)}(x)].$$

Using the differentiation formulas (52) and (53) and identity (36) allows us to calculate the term in the brackets (\*\*) from the above expression

$$\begin{aligned} (**) &= -x^\nu I_0(x)K_0^{(3)}(x) + x^\nu K_0(x)I_0^{(3)}(x) \\ &= \left(\frac{2}{x^2} + 1\right)[I_0(x)K_1(x) + I_1(x)K_0(x)] = \frac{2}{x^3} + \frac{1}{x}. \end{aligned}$$

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 repeated integral of the function  $I_0(x)$  will be used in the next lemma. It is consistent with the notation in Gaunt [20].

$$I_{(0,0,0)}(x) = I_0(x), \quad I_{(0,0,n)}(x) = \int_0^x I_{(0,0,n-1)}(y) dy, \quad n = 1, 2, 3 \dots \quad (37)$$

**Lemma A.2.** *Suppose  $h \in C_b^3(\mathbb{R})$  and let  $\tilde{h}(x) = h(x) - \text{PN}_2^1 h$ . Then the solution (17) of the PN(2, 1) Stein equation (9) and its first four derivatives, in the region  $x > 0$ , may be bounded as follows*

$$\begin{aligned} |f(x)| &\leq \|\tilde{h}\| |I_{(0,0,1)}(x)K_0(x)| + \|\tilde{h}\| \left| I_0(x) \int_x^\infty K_0(y) dy \right|, \\ |f'(x)| &\leq \|\tilde{h}\| |I_{(0,0,1)}(x)K_0'(x)| + \|\tilde{h}\| \left| I_0'(x) \int_x^\infty K_0(y) dy \right|, \\ |f''(x)| &\leq \|\tilde{h}\| \left| \frac{1}{x} - I_{(0,0,1)}(x)K_0''(x) \right| + \|h'\| |I_{(0,0,2)}(x)K_0''(x)| + \|\tilde{h}\| \left| I_0''(x) \int_x^\infty K_0(y) dy \right|, \\ |f^{(3)}(x)| &\leq \|\tilde{h}\| \left| \frac{2}{x^2} + I_{(0,0,1)}(x)K_0^{(3)}(x) \right| + \|h'\| \left| \frac{1}{x} + I_{(0,0,2)}(x)K_0^{(3)}(x) \right| \\ &\quad + \|h''\| |I_{(0,0,3)}(x)K_0^{(3)}(x)| + \|\tilde{h}\| \left| I_0^{(3)}(x) \int_x^\infty K_0(y) dy \right|, \end{aligned}$$

$$\begin{aligned}
 |f^{(4)}(x)| &\leq \|\tilde{h}\| \left| \frac{6}{x^3} + \frac{1}{x} - I_{(0,0,1)}(x)K_0^{(4)}(x) \right| + \|h'\| \left| \frac{3}{x^2} - I_{(0,0,2)}(x)K_0^{(4)}(x) \right| \\
 &\quad + \|h''\| \left| \frac{1}{x} - I_{(0,0,3)}(x)K_0^{(4)}(x) \right| + \|h^{(3)}\| |I_{(0,0,4)}(x)K_0^{(4)}(x)| \\
 &\quad + \|\tilde{h}\| \left| I_0^{(4)}(x) \int_x^\infty K_0(y) dy \right|.
 \end{aligned}$$

**Proof.** The first two bounds are immediate from the formulas for  $f(x)$  and  $f'(x)$  that are given in Lemma A.1. Integrating by parts and using the notation for the repeated integral of  $I_0(x)$ , which is defined above, gives

$$\begin{aligned}
 f''(x) &= \frac{\tilde{h}(x)}{x} - K_0''(x) \left[ \tilde{h}(x) \int_0^x I_0(y) dy - \int_0^x h'(y) \left( \int_0^y I_0(u) du \right) dy \right] \\
 &\quad - I_0''(x) \int_x^\infty K_0(y)\tilde{h}(y) dy \\
 &= \tilde{h}(x) \left( \frac{1}{x} - I_{(0,0,1)}(x)K_0''(x) \right) - K_0''(x) \int_0^x h'(y)I_{(0,0,1)}(y) dy \\
 &\quad - I_0''(x) \int_x^\infty K_0(y)\tilde{h}(y) dy.
 \end{aligned}$$

The bound now follows from the triangle inequality and (37). The bounds for the third and fourth derivatives are obtained in a similar manner, in which we apply integration by parts to the integrals  $I_{(0,0,n)}(x)$ . □

The expressions involving modified Bessel functions that appear in Lemma A.2 are bounded by Gaunt [20], and we list these bounds in Appendix C. We are now in a position to prove Theorem 2.1. Applying the inequalities of Appendix C to the bounds of Lemma A.2 and a simple change of variables leads to our bounds for the solution of the PN(2,  $\sigma^2$ ) Stein equation.

**Proof of Theorem 2.1.** Let  $\psi_g(x)$  denote the solution (17) of the PN(2, 1) Stein equation with test function  $g$ . Recalling Remark 2.2, it suffices to bound  $\psi_g$  and its first four derivatives in the region  $x \geq 0$ . Hence, applying the bounds, that are given in Appendix C, for expressions involving modified Bessel functions given in Lemma A.2 leads to the following bounds on the derivatives of the solution of the PN(2, 1) Stein equation:

$$\begin{aligned}
 \|\psi_g\| &\leq \left( 1 + \frac{\pi}{2} \right) \|g - \text{PN}_2^1 g\| \leq 3 \|g - \text{PN}_2^1 g\|, \\
 \|\psi_g'\| &\leq \frac{3}{2} \|g - \text{PN}_2^1 g\|, \\
 \|\psi_g''\| &\leq 2 \|g'\| + \left( \frac{13}{4} + \frac{\sqrt{\pi}}{2} \right) \|g - \text{PN}_2^1 g\| \leq 2 \|g'\| + 5 \|g - \text{PN}_2^1 g\|,
 \end{aligned}$$

$$\begin{aligned} \|\psi_g^{(3)}\| &\leq 4\|g''\| + 5\|g'\| + 4.89\|g - \text{PN}_2^1 g\|, \\ \|\psi_g^{(4)}\| &\leq 8\|g^{(3)}\| + 9\|g''\| + 6.81\|g'\| + 15.75\|g - \text{PN}_2^1 g\|, \\ \|x\psi_g(x)\| &\leq (0.615 + 1)\|h - \text{PN}_2^1 h\| \leq 2\|h - \text{PN}_2^1 h\|, \\ \|x\psi_g'(x)\| &\leq \frac{3}{2}\|h - \text{PN}_2^1 h\|, \end{aligned}$$

where in establishing the bound for  $\|\psi_g^{(4)}\|$  we used that  $14.61 + \frac{1}{4} + \frac{\sqrt{\pi}}{2} < 15.75$ . From the  $\text{PN}(2, 1)$  Stein equation, we have

$$\|x\psi_g''(x)\| \leq \|h - \text{PN}_2^1 h\| + \|\psi_g'\| + \|x\psi_g(x)\| \leq \frac{9}{2}\|h - \text{PN}_2^1 h\|.$$

We now make a change of variables to obtain bounds for the derivatives of the solution of the  $\text{PN}(2, \sigma^2)$  Stein equation. The function  $f_h(x) = \frac{1}{\sigma}\psi_g(\frac{x}{\sigma})$  solves the  $\text{PN}(2, \sigma^2)$  Stein equation

$$\sigma^2 x f''(x) + \sigma^2 f'(x) - x f(x) = h(x) - \text{PN}_2^{\sigma^2} h,$$

where  $h(x) = g(\frac{x}{\sigma})$ , since  $\text{PN}_2^{\sigma^2} h = \text{PN}_2^1 g$ . We verify that  $\text{PN}_2^{\sigma^2} h = \text{PN}_2^1 g$  with the following calculation:

$$\text{PN}_2^{\sigma^2} h = \int_{-\infty}^{\infty} \frac{1}{\pi\sigma} K_0\left(\frac{|x|}{\sigma}\right) h(x) dx = \int_{-\infty}^{\infty} K_0(|u|) g(u) du = \text{PN}_2^1 g,$$

where we made the change of variables  $u = \frac{x}{\sigma}$ . We have that  $\|f_h^{(k)}\| = \sigma^{-k-1}\|\psi_g^{(k)}\|$  and  $\|x f_h^{(k)}(x)\| = \sigma^{-k-1}\|x\psi_g^{(k)}(x)\|$  for  $k \geq 0$ , and  $\|g - \text{PN}_2^1 g\| = \|h - \text{PN}_2^{\sigma^2} h\|$  and  $\|g^{(k)}\| = \sigma^k \|h^{(k)}\|$  for  $k \geq 1$ . This completes the proof of Theorem 2.1.  $\square$

## Appendix B: Elementary properties of the Meijer $G$ -function and modified Bessel functions

Here we define the Meijer  $G$ -function and modified Bessel functions and state some of their elementary properties. For further properties of these functions see, for example, Olver *et al.* [33] and Luke [30]. The formulas for the Meijer  $G$ -function are given in Luke [30]. The modified Bessel function formulas can be found in Olver *et al.* [33], except for the second and third order derivative formulas which are given in Gaunt [15].

### B.1. The Meijer G-function

#### B.1.1. Definition

The Meijer G-function is defined, for  $z \in \mathbb{C} \setminus \{0\}$ , by the contour integral:

$$G_{p,q}^{m,n} \left( z \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} z^{-s} \frac{\prod_{j=1}^m \Gamma(s + b_j) \prod_{j=1}^n \Gamma(1 - a_j - s)}{\prod_{j=n+1}^p \Gamma(s + a_j) \prod_{j=m+1}^q \Gamma(1 - b_j - s)} ds,$$

where  $c$  is a real constant defining a Bromwich path separating the poles of  $\Gamma(s + b_j)$  from those of  $\Gamma(1 - a_j - s)$  and where we use the convention that the empty product is 1.

#### B.1.2. Basic properties

The Meijer G-function is symmetric in the parameters  $a_1, \dots, a_n; a_{n+1}, \dots, a_p; b_1, \dots, b_m$ ; and  $b_{m+1}, \dots, b_q$ . Thus, if one the  $a_j$ 's,  $j = n + 1, \dots, p$ , is equal to one of the  $b_k$ 's,  $k = 1, \dots, m$ , the Meijer G-function reduces to one of lower order. For example,

$$G_{p,q}^{m,n} \left( z \left| \begin{matrix} a_1, \dots, a_{p-1}, b_1 \\ b_1, \dots, b_q \end{matrix} \right. \right) = G_{p-1,q-1}^{m-1,n} \left( z \left| \begin{matrix} a_1, \dots, a_{p-1} \\ b_2, \dots, b_q \end{matrix} \right. \right), \quad m, p, q \geq 1. \tag{38}$$

The Meijer G-function satisfies the identities

$$z^c G_{p,q}^{m,n} \left( z \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right) = G_{p,q}^{m,n} \left( z \left| \begin{matrix} a_1 + c, \dots, a_p + c \\ b_1 + c, \dots, b_q + c \end{matrix} \right. \right), \quad z \in \mathbb{R} \tag{39}$$

and

$$G_{p,q}^{m,n} \left( z \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right) = G_{p,q}^{m,n} \left( \frac{1}{z} \left| \begin{matrix} 1 - b_1, \dots, 1 - b_q \\ 1 - a_1, \dots, 1 - a_p \end{matrix} \right. \right), \quad z \in \mathbb{R}. \tag{40}$$

#### B.1.3. Integration

For  $\alpha > 0, \gamma > 0, a_j < 1$  for  $j = 1, \dots, n$ , and  $b_j > -\frac{1}{2}$  for  $j = 1, \dots, m$ , we have

$$\int_0^\infty \cos(\gamma x) G_{p,q}^{m,n} \left( \alpha x^2 \left| \begin{matrix} a_1, \dots, a_p \\ b_1, \dots, b_q \end{matrix} \right. \right) dx = \sqrt{\pi} \gamma^{-1} G_{p+2,q}^{m,n+1} \left( \frac{4\alpha}{\gamma^2} \left| \begin{matrix} \frac{1}{2}, a_1, \dots, a_p, 0 \\ b_1, \dots, b_q \end{matrix} \right. \right). \tag{41}$$

#### B.1.4. Differential equation

The Meijer G-function satisfies a differential equation of order  $\max(p, q)$ :

$$\left[ (-1)^{p-m-n} z \prod_{j=1}^p \left( z \frac{d}{dz} - a_j + 1 \right) - \prod_{j=1}^q \left( z \frac{d}{dz} - b_j \right) \right] G(z) = 0. \tag{42}$$

## B.2. Modified Bessel functions

### B.2.1. Definitions

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

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

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

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

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

### B.2.2. Basic properties

For  $\nu \in \mathbb{R}$ , the modified Bessel function of the first kind  $I_\nu(x)$  and the modified Bessel function of the second kind  $K_\nu(x)$  are regular functions of  $x$ . For  $n \in \mathbb{Z}$ ,  $I_{2n}(x)$  is a real-valued function for all  $x \in \mathbb{R}$ , with  $I_{2n}(-x) = I_{2n}(x)$ . The modified Bessel function  $I_{2n+1}(x)$  is a real-valued for all  $x \in \mathbb{R}$ , with  $I_{2n+1}(-x) = -I_{2n+1}(x)$ . For  $\nu \geq 0$  and  $x > 0$  we have  $I_\nu(x) > 0$  and  $K_\nu(x) > 0$ . For all  $\nu \in \mathbb{R}$ , the modified Bessel function  $K_\nu(x)$  is complex-valued in the region  $x < 0$ .

### B.2.3. Representation in terms of the Meijer G-function

$$I_0(x) = G_{0,2}^{1,0} \left( -\frac{x^2}{4} \middle| 0, 0 \right), \quad x \in \mathbb{R}, \quad (43)$$

$$K_0(|x|) = \frac{1}{2} G_{0,2}^{2,0} \left( \frac{x^2}{4} \middle| 0, 0 \right), \quad x \in \mathbb{R}. \quad (44)$$

### B.2.4. Asymptotic expansions

$$K_0(x) \sim -\log x, \quad x \downarrow 0, \quad (45)$$

$$K_\nu(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x}, \quad x \rightarrow \infty, \quad (46)$$

$$I_\nu(x) \sim \frac{e^x}{\sqrt{2\pi x}}, \quad x \rightarrow \infty. \quad (47)$$

## B.2.5. Differentiation

$$I_0'(x) = I_1(x), \quad (48)$$

$$K_0'(x) = -K_1(x), \quad (49)$$

$$I_0''(x) = I_0(x) - \frac{I_1(x)}{x}, \quad (50)$$

$$K_0''(x) = K_0(x) + \frac{K_1(x)}{x}, \quad (51)$$

$$I_0^{(3)}(x) = -I_0(x) + \left(1 + \frac{2}{x^2}\right)I_1(x), \quad (52)$$

$$K_0^{(3)}(x) = -K_0(x) - \left(1 + \frac{2}{x^2}\right)K_1(x). \quad (53)$$

## B.2.6. Differential equation

The modified Bessel differential equation is

$$x^2 f''(x) + x f'(x) - (x^2 + \nu^2) f(x) = 0. \quad (54)$$

The general solution is  $f(x) = A I_\nu(x) + B K_\nu(x)$ .

## Appendix C: Bounds for expressions involving derivatives and integrals of modified Bessel functions

The following bounds, which can be found in Gaunt [15,20], are used to bound the derivatives of the solution to the PN(2,  $\sigma^2$ ) Stein equation (9). For  $x \geq 0$ ,

$$\begin{aligned} |I_{(0,0,n)}(x) K_0^{(n)}(x)| &\leq 2^{n-1}, & n = 1, 2, 3, \dots, \\ |x I_{(0,0,n)}(x) K_0^{(n)}(x)| &\leq 2^{n-1}, & n = 1, 2, 3, \dots, \\ |I_{(0,0,1)}(x) K_0(x)| &< 1, \\ |x I_{(0,0,1)}(x) K_0(x)| &< 1, \\ \left| I_0(x) \int_x^\infty K_0(y) dy \right| &\leq \frac{\pi}{2}, \\ \left| x I_0(x) \int_x^\infty K_0(y) dy \right| &< 0.615, \\ \left| I_0^{(2n)}(x) \int_x^\infty K_0(y) dy \right| &< \frac{1}{4} + \frac{\sqrt{\pi}}{2}, & n = 0, 1, 2, \dots, \end{aligned}$$



$$\begin{aligned}
\left| I_0^{(2n+1)}(x) \int_x^\infty K_0(y) dy \right| &< \frac{1}{2}, \quad n = 0, 1, 2, \dots, \\
\left| x I_0'(x) \int_x^\infty K_0(y) dy \right| &\leq \frac{1}{2}, \\
\left| \frac{1}{x} - I_{(0,0,1)}(x) K_0''(x) \right| &< 3, \\
\left| \frac{1}{x} + I_{(0,0,2)}(x) K_0^{(3)}(x) \right| &< 5, \\
\left| \frac{1}{x} - I_{(0,0,3)}(x) K_0^{(4)}(x) \right| &< 9, \\
\left| \frac{2}{x^2} + I_{(0,0,1)}(x) K_0^{(3)}(x) \right| &< 4.39, \\
\left| \frac{3}{x^2} - I_{(0,0,2)}(x) K_0^{(4)}(x) \right| &< 6.81, \\
\left| \frac{6}{x^3} + \frac{1}{x} - I_{(0,0,1)}(x) K_0^{(4)}(x) \right| &< 14.61.
\end{aligned}$$

## Acknowledgements

During the course of this research the author was supported by an EPSRC DPhil Studentship and an EPSRC Doctoral Prize, and is currently supported by EPSRC grant EP/K032402/1. The author would like to thank Gesine Reinert for the valuable guidance she provided on this project. The author would also like to thank Larry Goldstein for a helpful discussion, which lead to section of this paper concerning the connection between the zero bias transformation of order  $n$  and the square bias transformation. Finally, the author would like to thank the anonymous referees and an associate editor for some highly constructive comments and suggestions. The author is particularly grateful to one referee, who gave some interesting insights regarding the zero bias transformation of order  $n$ , which lead to a substantial improvement in the presentation of this paper.

## References

- [1] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. (2016). Stein's method on the second Wiener chaos: 2-Wasserstein distance. Preprint. Available at [arXiv:1601.03301](https://arxiv.org/abs/1601.03301).
- [2] Azmoodeh, E., Peccati, G. and Poly, G. (2015). Convergence towards linear combinations of chi-squared random variables: A Malliavin-based approach. In *In Memoriam Marc Yor – Séminaire de Probabilités XLVII. Lecture Notes in Math.* **2137** 339–367. Cham: Springer. MR3444306

- [3] Baldi, P., Rinott, Y. and Stein, C. (1989). A normal approximation for the number of local maxima of a random function on a graph. In *Probability, Statistics, and Mathematics* (T.W. Anderson, K.B. Athreya and D.L. Iglehart, eds.) 59–81. Academic Press, Boston, MA. [MR1031278](#)
- [4] Chatterjee, S., Fulman, J. and Röllin, A. (2011). Exponential approximation by Stein’s method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** 197–223. [MR2802856](#)
- [5] Chen, L.H.Y. (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3** 534–545. [MR0428387](#)
- [6] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein’s Method*. Heidelberg: Springer. [MR2732624](#)
- [7] Collins, P.J. (2006). *Differential and Integral Equations*. London: Oxford Univ. Press.
- [8] Döbler, C. (2015). Stein’s method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.* **20** 1–34. [MR3418541](#)
- [9] Döbler, C. (2016). Distributional transformations without orthogonality relations. *J. Theoret. Probab.* To appear.
- [10] Döbler, C., Gaunt, R.E. and Vollmer, S.J. (2015). An iterative technique for bounding derivatives of solutions of Stein equations. Preprint. Available at [arXiv:1510.02623](#).
- [11] Eden, R. and Viens, F. (2013). General upper and lower tail estimates using Malliavin calculus and Stein’s equations. In *Seminar on Stochastic Analysis, Random Fields and Applications VII* (R.C. Dalang, M. Dozzi and F. Russo, eds.) *Progress in Probability* **67** 55–84. Birkhäuser/Springer, Basel. [MR3380095](#)
- [12] Eden, R. and Víquez, J. (2015). Nourdin–Peccati analysis on Wiener and Wiener–Poisson space for general distributions. *Stochastic Process. Appl.* **125** 182–216. [MR3274696](#)
- [13] Eichelsbacher, P. and Thäle, C. (2015). Malliavin–Stein method for variance-gamma approximation on Wiener space. *Electron. J. Probab.* **20** 1–28. [MR3425543](#)
- [14] Folland, G.B. (1984). *Real Analysis: Modern Techniques and Their Applications*. New York: Wiley. [MR0767633](#)
- [15] Gaunt, R.E. (2013). Rates of convergence of variance-gamma approximations via Stein’s method. Ph.D. thesis, Univ. Oxford.
- [16] Gaunt, R.E. (2014). Variance-gamma approximation via Stein’s method. *Electron. J. Probab.* **19** 1–33. [MR3194737](#)
- [17] Gaunt, R.E. (2015). Products of normal, beta and gamma random variables: Stein operators and distributional theory. Preprint. Available at [arXiv:1507.07696](#).
- [18] Gaunt, R.E. (2015). Stein’s method for functions of multivariate normal random variables. Preprint. Available at [arXiv:1507.08688](#).
- [19] Gaunt, R.E. (2016). A Stein characterisation of the generalized hyperbolic distribution. Preprint. Available at [arXiv:1603.05675](#).
- [20] Gaunt, R.E. (2016). Uniform bounds for expressions involving modified Bessel functions. *Math. Inequal. Appl.* To appear.
- [21] Gaunt, R.E., Pickett, A. and Reinert, G. (2015). Chi-square approximation by Stein’s method with application to Pearson’s statistic. Preprint. Available at [arXiv:1507.01707](#).
- [22] Goldstein, L. and Reinert, G. (1997). Stein’s method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. [MR1484792](#)
- [23] Goldstein, L. and Reinert, G. (2005). Zero biasing in one and higher dimensions, and applications. In *Stein’s Method and Applications. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** 1–18. Singapore: Singapore Univ. Press. [MR2201883](#)
- [24] Goldstein, L. and Reinert, G. (2005). Distributional transformations, orthogonal polynomials, and Stein characterizations. *J. Theoret. Probab.* **18** 237–260. [MR2132278](#)

- [25] Goldstein, L. and Reinert, G. (2013). Stein's method for the beta distribution and the Pólya–Eggenberger urn. *J. Appl. Probab.* **50** 1187–1205. [MR3161381](#)
- [26] Goldstein, L. and Rinott, Y. (1996). Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** 1–17. [MR1371949](#)
- [27] Kreyszig, E. (1999). *Advanced Engineering Mathematics*, 8th ed. New York: Wiley. [MR1665766](#)
- [28] Ley, C., Reinert, G. and Swan, Y. (2014). Approximate computation of expectations: A canonical Stein operator. Preprint. Available at [arXiv:1408.2998](#).
- [29] Luk, H.M. (1994). *Stein's Method for the Gamma Distribution and Related Statistical Applications*. Ph.D. thesis, University of Southern California. [MR2693204](#)
- [30] Luke, Y.L. (1969). *The Special Functions and Their Approximations*, Vol. 1. New York: Academic Press.
- [31] Nourdin, I. and Peccati, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. [MR2520122](#)
- [32] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. [MR2962301](#)
- [33] Olver, F.W.J., Lozier, D.W., Boisvert, R.F. and Clark, C.W., eds. (2010). *NIST Handbook of Mathematical Functions*. Cambridge: Cambridge Univ. Press. [MR2723248](#)
- [34] Peköz, E.A. and Röllin, A. (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** 587–608. [MR2789507](#)
- [35] Peköz, E.A., Röllin, A. and Ross, N. (2013). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23** 1188–1218. [MR3076682](#)
- [36] Pike, J. and Ren, H. (2014). Stein's method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** 571–587. [MR3283586](#)
- [37] Reinert, G. (2005). Three general approaches to Stein's method. In *An Introduction to Stein's Method* (A.D. Barbour and L.H.Y. Chen, eds.) *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **4** 183–221. Singapore: Singapore Univ. Press. [MR2235451](#)
- [38] Springer, M.D. and Thompson, W.E. (1970). The distribution of products of beta, gamma and Gaussian random variables. *SIAM J. Appl. Math.* **18** 721–737. [MR0264733](#)
- [39] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, Vol. II: Probability Theory 583–602. Univ. California Press, Berkeley, Calif. [MR0402873](#)
- [40] Stein, C. (1986). *Approximate Computation of Expectations*. Hayward, CA: IMS. [MR0882007](#)
- [41] Stein, C., Diaconis, P., Holmes, S. and Reinert, G. (2004). Use of exchangeable pairs in the analysis of simulations. In *Stein's Method: Expository Lectures and Applications* (P. Diaconis and S. Holmes, eds.) *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **46** 1–26. Beachwood, OH: IMS. [MR2118600](#)
- [42] Stuart, A. and Ord, J.K. (1987). *Kendall's Advanced Theory of Statistics: Vol. I*, 5th ed. London: Charles Griffin.

Received June 2014 and revised March 2016