

## An iterative technique for bounding derivatives of solutions of Stein equations

Christian Döbler\*    Robert E. Gaunt†    Sebastian J. Vollmer‡

### Abstract

We introduce a simple iterative technique for bounding derivatives of solutions of Stein equations  $Lf = h - \mathbb{E}h(Z)$ , where  $L$  is a linear differential operator and  $Z$  is the limit random variable. Given bounds on just the solutions or certain lower order derivatives of the solution, the technique allows one to deduce bounds for derivatives of any order, in terms of supremum norms of derivatives of the test function  $h$ . This approach can be readily applied to many Stein equations from the literature. We consider a number of applications; in particular, we derive new bounds for derivatives of any order of the solution of the general variance-gamma Stein equation. Finally, we present a connection between Stein equations and Poisson equations, from which we first recognised the importance of the iterative technique to Stein’s method.

**Keywords:** Stein’s method; Stein equation; differential equation; Poisson equation; variance-gamma distribution.

**AMS MSC 2010:** 60F05; 35A24.

Submitted to EJP on October 10, 2016, final version accepted on October 19, 2017.

## 1 Introduction

Stein’s method is a powerful technique for assessing the distributional distance between a random variable of interest  $W$  and a limit random variable  $Z$ . Stein [52] originally developed the method for normal approximation, and the method was adapted to Poisson approximation by Chen [7].

### 1.1 Outline of Stein’s method

For a general target random variable  $Z$ , Stein’s method involves three steps; see [49]. In the first step, a suitable characterisation of the target distribution is obtained; namely, a Stein operator  $L$  such that the random variable  $X$  is equal in law to  $Z$  if and only if

$$\mathbb{E}[Lf(X)] = 0 \tag{1.1}$$

---

\*Université du Luxembourg, Unité de Recherche en Mathématiques, Maison du Nombre, 4, Avenue de la Fonte, L-4364 Esch-sur-Alzette E-mail: christian.doebler@uni.lu

†School of Mathematics, The University of Manchester, Manchester M13 9PL, UK E-mail: robert.gaunt@manchester.ac.uk

‡Mathematics Institute, Zeeman Building, University of Warwick, Coventry, CV4 7AL, UK E-mail: svollmer@turing.ac.uk

for all functions  $f$  belonging to some measure determining class. For continuous random variables,  $L$  is a differential operator; for discrete random variables,  $L$  is a difference operator. Such characterisations have often been derived via Stein's density approach [53], [54] or the generator approach of Barbour and Götze [3], [32]. The scope of the density approach has recently been extended by [38], and other techniques for obtaining Stein characterisations are discussed in that work.

The characterisation (1.1) leads to the so-called Stein equation:

$$Lf(x) = h(x) - \mathbb{E}h(Z), \tag{1.2}$$

where the test function  $h$  is real-valued. The second step of Stein's method, which will be the focus of this paper, concerns the problem of obtaining a solution  $f$  to the Stein equation (1.2) and then establishing estimates for  $f$  and some of its lower order derivatives (for continuous distributions).

Evaluating (1.2) at a random variable  $W$  and taking expectations gives

$$\mathbb{E}h(W) - \mathbb{E}h(Z) = \mathbb{E}[Lf(W)], \tag{1.3}$$

and thus the problem of bounding the quantity  $\mathbb{E}h(W) - \mathbb{E}h(Z)$  reduces to solving (1.2) and bounding the right-hand side of (1.3). The third step of Stein's method concerns the problem of bounding the expectation  $\mathbb{E}[Lf(W)]$ . For continuous limit distributions, such bounds are usually obtained via Taylor expansions and coupling techniques.

For many classical distributions, the problem of obtaining the first two necessary ingredients is relatively tractable. As a result, over the years, Stein's method has been adapted to many standard distributions, including the beta [11], [30], gamma [28], [39], exponential [5], [19], [45], Laplace [48] and, more generally, the class of variance-gamma distributions [21]. The method has also been adapted to distributions arising from specific problems, such as preferential attachment graphs [46], the Curie-Weiss model [6] and statistical mechanics [13], [14]. For a comprehensive overview of the literature, see [38].

## 1.2 Estimates for solutions of Stein equations

The Stein equations of many classical distributions, such as the normal, beta and gamma, are linear first order ODEs with simple coefficients. As a result, the problem of solving the Stein equation and bounding the derivatives of the solution is reasonably tractable.

In fact, for Stein characterisations arising from the generator approach, the Stein operator  $L$  associated with the distribution  $\mu$  can be viewed as the generator of a Markov process with stationary distribution  $\mu$ . For example, the classical standard normal Stein operator  $Lf(x) = f'(x) - xf(x)$  is the generator of an Ornstein-Uhlenbeck process (if we take  $f = g'$ ). In such cases, by standard theory of Markov processes ([17], Proposition 1.5), the solution to the Stein equation is  $f = g'$ , where

$$g(x) = - \int_0^\infty \{P_t h(x) - \mathbb{E}h(Z)\} dt, \tag{1.4}$$

if the integral exists. Here,  $P_t$ , defined by  $P_t f(x) = \mathbb{E}_x f(X_t)$ , is the transition semigroup operator. If it is permissible to differentiate under the integral sign, then

$$f^{(k-1)}(x) = - \int_0^\infty (P_t h)^{(k)}(x) dt, \quad k \geq 1.$$

In cases where a formula is available for  $(P_t h)^{(k)}(x)$ , it is often possible to easily deduce uniform bounds for the derivatives of  $f$  in terms of supremum norms of derivatives of  $h$ . This approach has yielded uniform bounds for derivatives of any order of the solutions of the multivariate normal [3], [31] and gamma [39] Stein equations.

However, for many Stein equations in the literature, this approach cannot be applied or is not tractable. This can either be because the Stein operator cannot be viewed as the generator of a Markov diffusion process (for example, the Stein operator for the product of three or more independent standard normals [25]) or because a simple formula is not available for  $(P_t h)^{(k)}(x)$  (for example, the variance-gamma Stein operator; see Section 3.3.2 of [20]). In such cases, estimates for the solution and its derivatives are usually obtained via analytic methods, by which formulas for the solution and its derivatives are derived, and these quantities are then bounded through technical calculations. This is often a tedious and lengthy process, which usually only yields bounds on lower order derivatives.

In recent years, Stein's method has been adapted to a number of distributions for which the Stein operator is a higher order linear differential operator. Second order operators have been obtained for the Laplace [48], variance-gamma [21] and PRR (Peköz-Röllin-Ross) distributions [46], whilst operators of order greater than two have been for the distribution of mixed products of independent normal, beta and gamma random variables [25], [26], [27], and for the distribution of a general linear combination of independent centered gamma random variables [2].

These higher order Stein operators have simple coefficients, so bounding the quantity  $E[Lf(W)]$  via various coupling techniques is often quite tractable. However, solving the Stein equation and obtaining estimates for the derivatives of the solution is often a challenging problem. Indeed, to date, there do not exist uniform bounds for the derivatives of the solution of the general variance-gamma Stein equation.

This paper is to some extent motivated by this challenge. In this paper, we introduce a simple iterative technique for bounding derivatives of solutions of Stein equations. We introduce the technique in Section 2 and consider applications of it in Section 3. Given a bound on just the solution or some of its lower order derivatives, the iterative procedure allows one to deduce bounds for derivatives of any order in terms of supremum norms of the derivatives of  $h$ . The technique allows us to obtain bounds for derivatives of any order of the solutions of many of the Stein equations from the literature. Such bounds have previously only been obtained for the solutions of the Stein equations for the normal and multivariate normal [8], [23], [31], [37], gamma [28], [39] and beta distributions [11]. In this paper, we restrict our attention to Stein equations of continuous distributions, but it should be noted that our iterative technique can be applied analogously to solutions of Stein equations for discrete distributions. However, one can obtain bounds on higher order forward differences of the solution trivially via the triangle inequality given either a bound on the solution or its first forward difference, without needing to apply our iterative technique.

The power of the technique is particularly well demonstrated if we consider the variance-gamma Stein equation. For this Stein equation, our iterative procedure reduces the problem of bounding derivatives of any order of the solution to just bounding the solution and its first derivative. This is a tractable problem, and we obtain uniform bounds for these quantities in Section 3. On the other hand, bounding even the second and third derivatives via analytic calculations may require very involved calculations; see Section 3 and Appendix D of [20].

When the Stein operator  $L$  is a second order (possibly degenerate) elliptic differential operator, then the Stein equation coincides with the Poisson equation. It was in the context of the Poisson equation that we first recognised the importance of the iterative technique to Stein's method. Indeed, in Section 4, we consider the connection between Stein equations and Poisson equations, and consider an application of the iterative procedure to Poisson equations in Section 4.2.

### 1.3 Outline of the paper

In Section 2, we introduce a general iterative technique for bounding derivatives of any order of solutions of Stein equations. In Lemmas 2.3–2.5, we present general bounds for solutions of certain classes of Stein equations. These lemmas cover the normal, gamma, beta, Student’s  $t$ , inverse-gamma, PRR and variance-gamma Stein equations. In Section 3, we apply these lemmas to obtain estimates for  $n$ -th order derivatives of the solutions to these Stein equations. Many of these bounds are new. The technique also applies equally well to many Stein equations not covered by Lemmas 2.3–2.5. We demonstrate this by considering the application of the technique to the Stein equations for the multivariate normal distribution, as well as to a distribution with density proportional to  $e^{-x^4/12}$ . In Section 4, we consider the connection between Stein equations and Poisson equations. We also consider the iterative technique in the context of Poisson equations in Section 4.2. For the sake of brevity, some formulas are given in Section 3 without proof. Their simple proofs are given in Appendix A.

## 2 An iterative approach to bounding solutions of Stein equations

Consider the ordinary differential equation

$$Lf(x) = h(x), \tag{2.1}$$

where  $L$  is a linear differential operator of the form

$$Lf(x) = \sum_{j=0}^m a_j(x)f^{(j)}(x).$$

In Stein’s method, one frequently encounters linear ODEs of the type (2.1). Indeed, usually one is interested in some distribution  $\mu$ , which is characterised by  $L$ , meaning that

$$\mathbb{E}[Lf(X)] = 0,$$

for all sufficiently smooth functions  $f$ , if and only if  $X$  has distribution  $\mu$ . One is then interested in solving (2.1) for  $f$  in terms of  $h$ , which will be supposed to be centred with respect to  $\mu$ , and in bounding  $f$  and derivatives of  $f$  in terms of  $h$  and its derivatives. In this paper, we address the latter problem within the following general set-up: Suppose that we are given a finite or infinite sequence  $(\mu_k, L_k)$  of distributions  $\mu_k$  and linear differential operators  $L_k$  of the form

$$L_k f(x) = \sum_{j=0}^m a_{k,j}(x)f^{(j)}(x) \tag{2.2}$$

such that  $L_k$  is characterizing for  $\mu_k$  in the above mentioned sense. We shall assume that  $a_{k+1,m}(x) = a_{k,m}(x) = \dots = a_{0,m}(x)$ . This mild assumption (which holds for all Stein equations for univariate distributions in the literature) allows us to make the following important assumption. We assume that the operators  $L_k$  are linked by the relation

$$\frac{d}{dx}L_k f(x) = L_{k+1}f'(x) - T_k f(x), \tag{2.3}$$

where  $T_k$  is some other linear differential operator. We will assume that all operators  $L_k$  have the same order  $m$  and, then,  $T_k$  will have the form

$$T_k f(x) = \sum_{j=0}^m b_{k,j}(x)f^{(j)}(x), \tag{2.4}$$

for some coefficients  $b_{k,j}(x)$  (which are in fact linear combinations of derivatives of the coefficients  $a_{k,j}(x)$  of  $L_k$ ). That  $T_k$  has order  $m$  can be seen because the lead-

ing terms of  $\frac{d}{dx}L_k f(x)$  and  $L_{k+1}f'(x)$  are given, respectively, by  $a_{k,m}(x)f^{(m+1)}(x)$  and  $a_{k+1,m}f^{(m+1)}(x)$ , which are equal because we assume that  $a_{k,m}(x) = a_{k+1,m}(x)$ . Thus, if  $h_k$  is such that

$$\mu_k(h_k) = 0, \tag{2.5}$$

then, writing  $f_k = f_{k,h_k}$  for a particular solution to

$$L_k f_k(x) = h_k(x), \tag{2.6}$$

we find that  $f_{k+1} = f'_k$  solves

$$L_{k+1}f_{k+1}(x) = h'_k(x) + T_k f_k(x). \tag{2.7}$$

Often, given  $h_k$  with (2.5), there is only one solution  $f_{k,h_k}$  to (2.6) which has nice boundedness properties. This is due to the fact that every other solution is obtained by adding to  $f_{k,h_k}$  a solution to the corresponding homogeneous equation

$$L_k g(x) = 0,$$

whose non-zero solutions (or one of their lower-order derivatives) often explode in a certain sense. In this general framework, we will always assume that there is one such distinguished solution  $f_{k,h_k}$  to (2.6). The curious fact is that then, if the operators  $L_m$  and  $T_m$  are well-chosen, usually, the right hand side

$$h_{k+1}(x) := h'_k(x) + T_k f_{k,h_k}(x)$$

of (2.7) is centred with respect to the measure  $\mu_{k+1}$  and that  $f'_{k,h_k} = f_{k+1,h_{k+1}}$ . Given this, we can apply bounds holding for the solution  $f_{k+1,h_{k+1}}$  in order to bound  $f'_{k,h_k}$ . Iterating this observation already gives the theoretical essence of our method. For ease of reference, we now state the technical conditions.

**Assumption 2.1.** Assume that, for some  $N \in \mathbb{N} \cup \{\infty\}$  and  $m \in \mathbb{N}$ , we are given a sequence  $(\mu_k, L_k, T_k)$ ,  $0 \leq k < N + 1$ , of probability measures  $\mu_k$  on  $(\mathbb{R}, \mathcal{B})$ , linear differential operators  $L_k$  and  $T_k$  of the form (2.2) and (2.4), respectively, such that  $L_k$  is characterizing for  $\mu_k$  and (2.3) holds. Also, suppose that for each  $0 \leq k < N + 1$  and each function  $h_k$  such that  $\mu_k(h_k) = 0$ , there is a distinguished solution  $f_{k,h_k}$  to the Stein equation (2.6) and that the function  $h_{k+1} := h'_k + T_k f_{k,h_k}$  is integrable and centred with respect to  $\mu_{k+1}$ . Then, we know from (2.3) that  $f'_{k,h_k}$  solves  $L_{k+1}f = h_{k+1}$  and we further assume that  $f'_{k,h_k} = f_{k+1,h_{k+1}}$  is the distinguished solution of this equation.

**Remark 2.2.** Note that we do not state in general, what the attribute distinguished in this context precisely means. Indeed, for the general theory, it only has to be guaranteed that the choice of  $f_{k,h_k}$  is unambiguous. In all our examples in the sequel except for the variance-gamma Stein equation when  $0 < r < 1$  (see Section 3.1.7 for the p.d.f.), the distinguished solution will be the unique bounded solution. In case of the variance-gamma Stein equation when  $0 < r < 1$ , there are infinitely many bounded solutions to the Stein equation, for general bounded test function  $h$ . However, if  $h$  is bounded, then for all parameter values there is a unique solution  $f$  to the variance-gamma Stein equation such that  $f$  and  $f'$  are bounded (see [20], Lemma 3.14). This then becomes the distinguished solution.

Now, we present three general lemmas, which provide general useful bounds under different additional conditions, if the operators  $T_k$  have a special and simple form. As we shall see in Section 3, these lemmas cover many of the Stein equations from the literature. In the following lemmas, we set the empty product  $\prod_{i=0}^{-1} a_i$  to be equal to 1.

**Lemma 2.3.** Suppose that Assumption 2.1 holds and that there are constants  $\alpha_j$ ,  $0 \leq j < N + 1$ , such that  $T_j f = \alpha_j f$  holds for all  $0 \leq j < N + 1$ , where  $f$  is distinguished solution

according to Assumption 2.1. Also, assume that for some  $0 \leq n < N + 1$  a function  $h \in C_b^n(\mathbb{R})$  with  $\mu_0(h) = 0$  is given and write  $f$  for  $f_{0,h}$ . Then, writing  $a_j := |\sum_{i=0}^j \alpha_i|$ ,  $0 \leq j < N + 1$ :

- (i) If there exist constants  $C_l$ ,  $0 \leq l < N + 1$ , such that  $\|f_{l,h_l}\| \leq C_l \|h_l\|$  holds for each  $0 \leq l < N + 1$  and  $h_l$  with  $\mu_l(h_l) = 0$ , then we have

$$\|f^{(n)}\| \leq C_n \sum_{j=0}^n \left( \prod_{i=j}^{n-1} C_i a_i \right) \|h^{(j)}\|.$$

- (ii) If there exist constants  $D_l$ ,  $0 \leq l < N + 1$ , such that  $\|f'_{l,h_l}\| \leq D_l \|h_l\|$  holds for each  $0 \leq l < N + 1$  and  $h_l$  with  $\mu_l(h_l) = 0$ , then we have, for  $n = 2k + 1$ ,

$$\|f^{(2k+1)}\| \leq D_{2k} \sum_{j=0}^k \left( \prod_{i=j}^{k-1} D_{2i} a_{2i+1} \right) \|h^{(2j)}\|$$

and, for  $n = 2k$ ,

$$\|f^{(2k)}\| \leq D_{2k-1} \sum_{j=1}^k \left( \prod_{i=j}^{k-1} D_{2i-1} a_{2i} \right) \|h^{(2j-1)}\| + \|f\| \prod_{i=1}^k D_{2i-1} a_{2i-2}.$$

- (iii) If  $1 \leq n < N + 1$  and if there exist constants  $E_l$ ,  $0 \leq l < N + 1$ , such that  $\|f'_{l,h_l}\| \leq E_l \|h'_l\|$  holds for each  $0 \leq l < N + 1$  and Lipschitz-continuous  $h_l$  with  $\mu_l(h_l) = 0$ , then we have

$$\|f^{(n)}\| \leq E_{n-1} \sum_{j=1}^n \left( \prod_{i=j}^{n-1} E_{i-1} a_{i-1} \right) \|h^{(j)}\|.$$

*Proof.* From the hypothesis of the lemma we obtain inductively that

$$L_{j+1} f^{(j+1)} = h_{j+1} = h'_j + \alpha_j f^{(j)} = \dots = h^{(j+1)} + \left( \sum_{i=0}^j \alpha_i \right) f^{(j)}$$

holds for each  $0 \leq j < N$ . This, in particular implies that

$$\|h_{j+1}\| \leq \|h^{(j+1)}\| + \left| \sum_{i=0}^j \alpha_i \right| \|f^{(j)}\|$$

for each  $0 \leq j < N$ .

- (i) The result follows from a simple induction on  $n$ . Suppose the result is true for all  $l \leq n$ . Then

$$\begin{aligned} \|f^{(n+1)}\| &\leq C_{n+1} \|h^{(n+1)}\| + \left( \sum_{i=0}^n \alpha_i \right) \|f^{(n)}\| \\ &\leq C_{n+1} \|h^{(n+1)}\| + C_{n+1} a_n \|f^{(n)}\| \\ &\leq C_{n+1} \|h^{(n+1)}\| + C_{n+1} a_n \cdot C_n \sum_{j=0}^n \left( \prod_{i=j}^{n-1} C_i a_i \right) \|h^{(j)}\| \\ &= C_{n+1} \sum_{j=0}^{n+1} \left( \prod_{i=j}^n C_i a_i \right) \|h^{(j)}\|, \end{aligned}$$

and so the result has been proved by induction on  $n$ .

- (ii) We prove the result by induction by considering the cases of odd and even  $n$  separately. Firstly, suppose the result holds for all odd  $l \leq 2k + 1$ . Then

$$\begin{aligned} \|f^{(2k+3)}\| &\leq D_{2k+2}\|h^{(2k+2)}\| + \left(\sum_{i=0}^{2k+1} \alpha_i\right) f^{(2k+1)}\| \\ &\leq D_{2k+2}\|h^{(2k+2)}\| + D_{2k+2}a_{2k+1}\|f^{(2k+1)}\| \\ &\leq D_{2k+2}\|h^{(2k+2)}\| + D_{2k+2}a_{2k+1} \cdot D_{2k} \sum_{j=0}^k \left(\prod_{i=j}^{k-1} D_{2i}a_{2i+1}\right) \|h^{(2j)}\| \\ &= D_{2k+2} \sum_{j=0}^{k+1} \left(\prod_{i=j}^k D_{2i}a_{2i+1}\right) \|h^{(2j)}\|, \end{aligned}$$

as required. Now, we suppose the result holds for all even  $l \leq 2k$ . Then

$$\begin{aligned} \|f^{(2k+2)}\| &\leq D_{2k+1}\|h^{(2k+1)}\| + \left(\sum_{i=0}^{2k} \alpha_i\right) f^{(2k)}\| \\ &\leq D_{2k+1}\|h^{(2k+1)}\| + D_{2k+1}a_{2k}\|f^{(2k)}\| \\ &\leq D_{2k+1}\|h^{(2k+1)}\| + D_{2k+1}a_{2k} \cdot D_{2k-1} \sum_{j=1}^k \left(\prod_{i=j}^{k-1} D_{2i-1}a_{2i}\right) \|h^{(2j-1)}\| \\ &\quad + D_{2k+1}a_{2k-1}\|f\| \prod_{i=1}^{k-1} D_{2i-1}a_{2i-2} \\ &= D_{2k+1} \sum_{j=1}^{k+1} \left(\prod_{i=j}^k D_{2i-1}a_{2i}\right) \|h^{(2j-1)}\| + \|f\| \prod_{i=1}^k D_{2i-1}a_{2i-2}, \end{aligned}$$

which completes the proof. The proof of (iii) is very similar to that of (i) and is therefore omitted. □

In the following lemma we suppose now that  $T_j f = \alpha_j f'$  instead of  $T_j f = \alpha_j f$ . The proof is analogous to that of Lemma 2.3 and is omitted.

**Lemma 2.4.** *Suppose that Assumption 2.1 holds and that there are constants  $\alpha_j, 0 \leq j < N + 1$ , such that  $T_j f = \alpha_j f'$  holds for all  $0 \leq j < N + 1$ . Also, assume that for some  $0 \leq n < N + 1$  a function  $h \in C_b^n(\mathbb{R})$  with  $\mu_0(h) = 0$  is given and write  $f$  for  $f_{0,h}$ . Then, writing  $a_j := |\sum_{i=0}^j \alpha_i|, 0 \leq j < N + 1$ :*

- (i) *If there exist constants  $C_l, 0 \leq l < N + 1$ , such that  $\|f'_{l,h_l}\| \leq C_l \|h_l\|$  holds for each  $0 \leq l < N + 1$  and  $h_l$  with  $\mu_l(h_l) = 0$ , then we have*

$$\|f^{(n)}\| \leq C_{n-1} \sum_{j=0}^{n-1} \left(\prod_{i=j}^{n-2} C_i a_i\right) \|h^{(j)}\|.$$

- (ii) *If there exists constants  $D_l, 0 \leq l < N + 1$ , such that  $\|f''_{l,h_l}\| \leq D_l \|h_l\|$  holds for each  $0 \leq l < N + 1$  and  $h_l$  with  $\mu_l(h_l) = 0$ , then we have, for  $1 \leq n = 2k + 1$ ,*

$$\|f^{(2k+1)}\| \leq D_{2k-1} \sum_{j=1}^k \left(\prod_{i=j}^{k-1} D_{2i-1}a_{2i}\right) \|h^{(2j-1)}\| + \|f'\| \prod_{i=1}^k D_{2i-1}a_{2i-2}$$

and, for  $2 \leq n = 2k$ ,

$$\|f^{(2k)}\| \leq D_{2k-2} \sum_{j=1}^{k-1} \left(\prod_{i=j}^{k-2} D_{2i}a_{2i+1}\right) \|h^{(2j)}\|.$$

In the following lemma we generalise the settings of Lemmas 2.3 and 2.4 to suppose that  $T_j f = \alpha_{j+1} f + \beta_{j+1} f'$ .

**Lemma 2.5.** *Suppose that Assumption 2.1 holds and that there are constants  $\alpha_j$  and  $\beta_j$ ,  $1 \leq j < N + 2$ , such that  $T_j f = \alpha_{j+1} f + \beta_{j+1} f'$  holds for all  $0 \leq j < N + 1$ . Also, assume that for some  $1 \leq m < N + 1$  a function  $h \in C_b^m(\mathbb{R})$  with  $\mu_0(h) = 0$  is given and write  $f$  for  $f_{0,h}$ .*

*If there exist constants  $K_j$  and  $D_j$ ,  $0 \leq j < N + 1$ , such that  $\|f_{j,h_j}\| \leq K_j \|h_j\|$  and  $\|f'_{j,h_j}\| \leq D_j \|h_j\|$  hold for each  $0 \leq j < N + 1$  and  $h_k$  with  $\mu_k(h_k) = 0$ , then, with  $a_j := |\sum_{i=0}^j \alpha_i|$  and  $b_j := |\sum_{i=0}^j \beta_i|$ , we have that*

$$\|f^{(m+1)}\| \leq \sum_{j=0}^{m-1} D_{m-j} \left( \sum_{l \in I_j} A_{j,l}^{(m)} \right) \|h^{(m-j)}\| + \left( D_0 \sum_{l \in I_m} A_{m,l}^{(m)} + a_1 K_0 D_1 \sum_{l \in I_{m-1}} A_{m-1,l}^{(m)} \right) \|h\|.$$

Here, for  $j \in \{0, 1, \dots\}$ , we let

$$I_j := \left\{ j + 1 - \left\lfloor \frac{j}{2} \right\rfloor, \dots, j + 1 \right\} \quad \text{such that} \quad |I_j| = 1 + \left\lfloor \frac{j}{2} \right\rfloor$$

and, for  $l \in I_j$ , we define the constants

$$A_{j,l}^{(m)} := \sum_{M \in S_{j,l}^{(m)}} \prod_{i \in M \setminus \{m-j\}} (b_i 1_{\{i-1 \in M\}} + a_i 1_{\{i-1 \notin M\}}) D_i,$$

where  $S_{j,l}^{(m)}$  is the collection of those size  $l$  subsets  $M$  of  $M_j^{(m)} := \{m-j, m-j+1, \dots, m\}$  such that  $\{m, m-j\} \subseteq M$  and  $M \cap \{i, i+1\} \neq \emptyset$  for all  $m-j \leq i \leq m-1$ .

*Proof.* By induction one can see that

$$L_{j+1} f^{(j+1)} = h^{(j+1)} + \left( \sum_{i=1}^{j+1} \alpha_i \right) f^{(j)} + \left( \sum_{i=1}^{j+1} \beta_i \right) f^{(j+1)}$$

holds for all  $0 \leq j < N$ . We will occasionally suppress the index  $m$  from the notation. Note that for each fixed  $1 \leq k \leq m$  we have

$$\|f^{(k+1)}\| \leq D_k \|h^{(k)}\| + b_k D_k \|f^{(k)}\| + a_k D_k \|f^{(k-1)}\|.$$

From this, for each  $0 \leq k \leq m - 1$ , we can obtain a formula of the form

$$\|f^{(m+1)}\| \leq \sum_{j=0}^k C_1^{(m-j)} \|h^{(m-j)}\| + C_2^{(m-k)} \|f^{(m-k)}\| + C_3^{(m-k)} \|f^{(m-k-1)}\| \quad (2.8)$$

with sequences  $C_1^{(j)}, C_2^{(j)}, C_3^{(j)}$ ,  $j = 1, \dots, m$ , which satisfy the following recursive relations:

$$C_1^{(k-1)} = D_{k-1} C_2^{(k)} \quad (2.9)$$

$$C_2^{(k-1)} = C_3^{(k)} + b_{k-1} D_{k-1} C_2^{(k)} \quad \text{and} \quad (2.10)$$

$$C_3^{(k-1)} = a_{k-1} D_{k-1} C_2^{(k)}. \quad (2.11)$$

The initial values are given by

$$C_1^{(m)} = D_m, \quad C_2^{(m)} = b_m D_m \quad \text{and} \quad C_3^{(m)} = a_m D_m. \quad (2.12)$$

Choosing  $k = m - 1$  in (2.8), we obtain

$$\|f^{(m+1)}\| \leq \sum_{j=0}^{m-1} C_1^{(m-j)} \|h^{(m-j)}\| + C_2^{(1)} \|f'\| + C_3^{(1)} \|f\|$$

$$\begin{aligned} &\leq \sum_{j=0}^{m-1} C_1^{(m-j)} \|h^{(m-j)}\| + D_0 C_2^{(1)} \|h\| + K_0 C_3^{(1)} \|h\| \\ &= \sum_{j=0}^{m-1} C_1^{(m-j)} \|h^{(m-j)}\| + C_1^{(0)} \|h\|, \end{aligned} \tag{2.13}$$

where

$$C_1^{(0)} := D_0 C_2^{(1)} + K_0 C_3^{(1)}.$$

Thus, it suffices to find explicit expressions for the sequences  $C_1^{(m-j)}, C_2^{(m-j)}, C_3^{(m-j)}$ ,  $j = 0, 1, \dots, m - 1$ . We claim that these are given by

$$\begin{aligned} C_1^{(m-j)} &= \sum_{l \in I_j} D_{m-j} A_{j,l} = \sum_{l \in I_j} \sum_{M \in \mathcal{S}_{j,l}} D_{m-j} \prod_{i \in M \setminus \{m-j\}} (b_i 1_{\{i-1 \in M\}} + a_i 1_{\{i-1 \notin M\}}) D_i, \\ C_2^{(m-j)} &= \sum_{l \in I_{j+1}} A_{j+1,l} = \sum_{l \in I_{j+1}} \sum_{M \in \mathcal{S}_{j+1,l}} \prod_{i \in M \setminus \{m-j-1\}} (b_i 1_{\{i-1 \in M\}} + a_i 1_{\{i-1 \notin M\}}) D_i, \\ C_3^{(m-j)} &= \sum_{l \in I_j} a_{m-j} D_{m-j} A_{j,l} = \sum_{l \in I_j} \sum_{M \in \mathcal{S}_{j,l}} a_{m-j} D_{m-j} \\ &\quad \times \prod_{i \in M \setminus \{m-j\}} (b_i 1_{\{i-1 \in M\}} + a_i 1_{\{i-1 \notin M\}}) D_i. \end{aligned}$$

This will be shown by induction on  $j = 0, 1, \dots, m - 1$ . If  $j = 0$ , then the claim is true by virtue of (2.12). Suppose that it holds for some  $0 \leq j \leq m - 2$ . Then, from (2.9) and the induction hypothesis for  $C_2^{(m-j)}$  we obtain

$$C_1^{(m-j-1)} = D_{m-j-1} C_2^{(m-j)} = \sum_{l \in I_{j+1}} D_{m-j-1} A_{j+1,l},$$

as claimed. Similarly, we obtain the claim for  $C_3^{(m-j-1)}$ . As to  $C_2^{(m-j-1)}$ , from (2.10) and the induction hypothesis for  $C_2^{(m-j)}$  and for  $C_3^{(m-j)}$  we obtain that

$$\begin{aligned} C_2^{(m-j-1)} &= C_3^{(m-j)} + b_{m-j-1} D_{m-j-1} C_2^{(m-j)} \\ &= \sum_{l \in I_j} a_{m-j} D_{m-j} A_{j,l} + \sum_{l \in I_{j+1}} b_{m-j-1} D_{m-j-1} A_{j+1,l}. \end{aligned} \tag{2.14}$$

Suppose that  $l \in I_j$ . Then, if  $M \in \mathcal{S}_{j,l}$ , we have  $M' := M \cup \{m - j - 2\} \in \mathcal{S}_{j+2,l+1}$  and  $m - j - 1 \notin M'$ . This implies that

$$\begin{aligned} a_{m-j} D_{m-j} A_{j,l} &= \sum_{M \in \mathcal{S}_{j,l}} a_{m-j} D_{m-j} \prod_{i \in M \setminus \{m-j\}} D_i (b_i 1_{\{i-1 \in M\}} + a_i 1_{\{i-1 \notin M\}}) \\ &= \sum_{\substack{M' \in \mathcal{S}_{j+2,l+1}: \\ m-j-1 \notin M'}} \prod_{i \in M' \setminus \{m-j-2\}} D_i (b_i 1_{\{i-1 \in M\}} + a_i 1_{\{i-1 \notin M\}}), \end{aligned}$$

because, for  $M' \in \mathcal{S}_{j+2,l+1}$ , the condition that  $m - j - 1 \notin M'$  necessitates that  $m - j \in M'$ . Furthermore, if  $j$  is even, then it is easy to see that

$$a_{m-j} D_{m-j} A_{j,j/2+1} = A_{j+2,j/2+2},$$

because  $M' \in \mathcal{S}_{j+2,j/2+2}$  necessarily implies that  $m - j - 1 \notin M'$ .

Similarly, if  $l \in I_{j+1}$  and  $N \in \mathcal{S}_{j+1,l}$ , then  $N' := N \cup \{m - j - 2\} \in \mathcal{S}_{j+2,l+1}$  and  $m - j - 1 \in N \subseteq N'$  and, thus,

$$b_{m-j-1} D_{m-j-1} A_{j+1,l} = \sum_{N \in \mathcal{S}_{j+1,l}} b_{m-j-1} D_{m-j-1} \prod_{i \in N \setminus \{m-j-1\}} D_i (b_i 1_{\{i-1 \in N\}} + a_i 1_{\{i-1 \notin N\}})$$

$$= \sum_{\substack{N' \in \mathcal{S}_{j+2, l+1}: \\ m-j-1 \in N'}} \prod_{i \in N' \setminus \{m-j-2\}} D_i (b_i 1_{\{i-1 \in N'\}} + a_i 1_{\{i-1 \notin N'\}}).$$

Also, if  $j$  happens to be odd, then it is easily checked that

$$b_{m-j-1} D_{m-j-1} A_{j+1, j+2} = A_{j+2, j+3}, \tag{2.15}$$

because for each  $j$  the only size  $j + 1$  subset of  $M_j^{(m)}$  is  $M_j^{(m)}$  itself. Thus, from (2.14)-(2.15) we conclude that, indeed,

$$C_2^{(m-j-1)} = \sum_{l \in I_{j+2}} A_{j+2, l},$$

as claimed. This finishes the induction step and (2.13) now yields

$$\begin{aligned} \|f^{(m+1)}\| &\leq \sum_{j=0}^{m-1} C_1^{(m-j)} \|h^{(m-j)}\| + (D_0 C_2^{(1)} + K_0 C_3^{(1)}) \|h\| \\ &= \sum_{j=0}^{m-1} D_{m-j} \left( \sum_{l \in I_j} A_{j, l}^{(m)} \right) \|h^{(m-j)}\| + \left( D_0 \sum_{l \in I_m} A_{m, l}^{(m)} + a_1 K_0 D_1 \sum_{l \in I_{m-1}} A_{m-1, l}^{(m)} \right) \|h\|, \end{aligned}$$

completing the proof. □

### 3 Bounds for derivatives of solutions of Stein equations via the iterative technique

In this section, we apply the iterative technique, that was introduced in Section 2, to obtain bounds for derivatives of any order for the solutions of a number of Stein equations from the literature. The iterative method presented in this paper is particularly suitable for linear Stein equations, whose coefficients are polynomials of order at most one. In particular, we can directly apply Lemmas 2.3–2.5 to obtain estimates for the solutions of the normal, gamma, beta, Student’s  $t$ , PRR and variance-gamma distributions.

However, the iterative approach also applies equally well to many linear Stein equations not covered by these lemmas. We illustrate this with an application to a distribution with density proportional to  $e^{-x^4/12}$ , which appears as the distributional limit of the magnetization in a Curie-Weiss model. The iterative method can also be used for Stein equations that are linear PDEs, and we demonstrate this with an application of the iterative technique to the multivariate normal Stein equation.

#### 3.1 Estimates for solutions of Stein equations

Here, we consider the application of the iterative method to a number of Stein equations from the literature. For each of the distributions considered, we state the Stein equation and its solution, as well as the best existing bounds for the solution and its derivatives. For each distribution, the iterative technique can be applied to bound derivatives of any order given bounds on just the solution or some of its lower order derivatives. The iterative technique does not lead to improvements in estimates for derivatives of the solution of the normal and gamma Stein equations, although it does allow for simpler derivations at the expense of slightly larger constants (see Sections 3.2.1 and 3.2.2). However, we do obtain new bounds for Student’s  $t$  and the inverse-gamma distributions, the distribution with density proportional to  $e^{-x^4/12}$ , as well as new bounds for higher order derivatives of the PRR and variance-gamma distributions. The results are stated without further comment; proofs and further discussions are given

in Section 3.2. It should be understood that the bounds stated below are new, unless a citation is given.

*Notation.* The double factorial of a positive integer  $n$  is given by

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

and we define  $(-1)!! = 0!! = 1$  ([1], p. 547). The Pochhammer symbol is defined by  $(x)_n = x(x+1)\cdots(x+n-1)$  and the Pochhammer  $k$ -symbol is defined as  $(x)_{n,k} = x(x+k)(x+2k)\cdots(x+(n-1)k)$ . As in Section 2, we set the empty product  $\prod_{i=0}^{-1} a_i$  to be equal to 1. We say that the function  $h : S \subseteq \mathbb{R} \rightarrow \mathbb{R}$  belongs to the class  $C_b^n(S)$  if  $h^{(n-1)}$  exists and is absolutely continuous, with derivatives up to  $n$ -th order bounded. It shall also be convenient to let  $\tilde{h}(x) = h(x) - \mathbb{E}h(Z)$  and  $h^{(0)} \equiv h$ . For a function  $f : S \subseteq \mathbb{R} \rightarrow \mathbb{R}$ , we shall write  $\|f\| = \|f\|_\infty = \sup_{x \in S} |f(x)|$ . Finally, we shall let  $D$  denote the usual differential operator and  $I$  denote the identity operator

### 3.1.1 Normal distribution

1. Let  $Z \sim N(0, 1)$  with density  $p(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ ,  $x \in \mathbb{R}$ .
2. The Stein equation is (see [52], [53])

$$Lf(x) = f'(x) - xf(x) = \tilde{h}(x) \tag{3.1}$$

where  $\tilde{h}(x) = h(x) - \mathbb{E}h(Z)$ , and, for  $k \geq 1$ ,

$$Lf^{(k)}(x) = h^{(k)}(x) + kf^{(k-1)}(x). \tag{3.2}$$

We can verify (3.2) by induction on  $k$ . Differentiating both sides of (3.2) gives

$$\frac{d}{dx}Lf^{(k)}(x) = f^{(k+2)}(x) - xf^{(k+1)}(x) - f^{(k)}(x) = h^{(k+1)}(x) + kf^{(k)}(x), \tag{3.3}$$

which on rearranging yields

$$Lf^{(k+1)}(x) = h^{(k+1)}(x) + (k+1)f^{(k)}(x),$$

as required. From (3.3) we can read off that  $(L_k, T_k) = (L, I)$  for  $k \geq 0$ . We can therefore bound the derivatives of the solution via Lemma 2.3, with  $a_k = k + 1$  for  $k \geq 0$ .

3. The solution to (3.1) is

$$f(x) = e^{x^2/2} \int_{-\infty}^x \tilde{h}(t)e^{-t^2/2} dt. \tag{3.4}$$

4. The following bounds for the solution were obtained by [53]:

$$\|f\| \leq \sqrt{\frac{\pi}{2}}\|\tilde{h}\|, \quad \|f'\| \leq 2\|\tilde{h}\|, \quad \|f''\| \leq 2\|h'\|. \tag{3.5}$$

These bounds have since been extended to higher order derivatives. For  $h \in C_b^n(\mathbb{R})$ ,  $n \geq 0$ , the solution satisfies (see [31], [23], [8] respectively)

$$\|f^{(n)}\| \leq \frac{\|h^{(n+1)}\|}{n+1}, \tag{3.6}$$

$$\|f^{(n)}\| \leq \frac{\Gamma(\frac{n+1}{2})\|h^{(n+1)}\|}{\sqrt{2}\Gamma(\frac{n}{2}+1)}, \tag{3.7}$$

$$\|f^{(n+2)}\| \leq 2\|h^{(n+1)}\|. \tag{3.8}$$

**3.1.2 Gamma distribution**

1. Let  $Z \sim \Gamma(r, \lambda)$ ,  $r, \lambda > 0$ , with density  $p(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$ ,  $x > 0$ .
2. The Stein equation is (see [9])

$$L_{r,\lambda}f(x) = xf'(x) + (r - \lambda x)f(x) = \tilde{h}(x) \tag{3.9}$$

and, for  $k \geq 1$ ,

$$L_{r+k,\lambda}f^{(k)}(x) = h^{(k)}(x) + k\lambda f^{(k-1)}(x). \tag{3.10}$$

As was the case for the normal distribution, we can verify (3.10) by induction. This is the case for all the subsequent distributions, and for the sake of brevity we omit the details for all subsequent distributions. Differentiating both sides of (3.9) gives

$$\begin{aligned} \frac{d}{dx}L_{r+k,\lambda}f^{(k)}(x) &= xf^{(k+2)}(x) + (r + k - \lambda x)f^{(k+1)}(x) + f^{(k+1)}(x) - \lambda f^{(k)}(x) \\ &= h^{(k+1)}(x) + k\lambda f^{(k)}(x), \end{aligned} \tag{3.11}$$

which on rearranging gives

$$L_{r+k+1,\lambda}f^{(k+1)}(x) = h^{(k+1)}(x) + (k + 1)\lambda f^{(k)}(x),$$

as required. From (3.11), we can read off that  $(L_k, T_k) = (L_{r+k,\lambda}, \lambda I)$  for  $k \geq 0$ . We can therefore bound the derivatives of the solution via Lemma 2.3, with  $a_k = \lambda(k+1)$  for  $k \geq 0$ .

3. The solution to (3.9) is

$$f(x) = \frac{e^{\lambda x}}{x^r} \int_0^x t^{r-1} e^{-\lambda t} \tilde{h}(t) dt. \tag{3.12}$$

4. For  $h \in C_b^n(\mathbb{R}_+)$ ,  $n \geq 0$ , the solution satisfies (see [39], [20], [28] respectively)

$$\|f^{(n-1)}\| \leq \frac{1}{n\lambda} \|h^{(n)}\|, \tag{3.13}$$

$$\|f^{(n)}\| \leq \left\{ \frac{\sqrt{2\pi} + e^{-1}}{\sqrt{r+n}} + \frac{2}{r+n} \right\} \|h^{(n)}\|, \tag{3.14}$$

$$\|f^{(n)}\| \leq \frac{2}{r+n} \{3\|h^{(n)}\| + 2\lambda\|h^{(n-1)}\|\}. \tag{3.15}$$

A straightforward application of Lemma 1 of [50] yields

$$\|f\| \leq \frac{e^r \Gamma(r)}{r^r} \|\tilde{h}\|, \tag{3.16}$$

which has a smaller constant than (3.14) when  $n = 0$ .

5. Consider the  $\text{Exp}(\lambda) \stackrel{D}{=} \Gamma(1, \lambda)$  distribution. For the case  $r = 1$ , some of the estimates for the solution (3.9) can be improved. For a continuously differentiable test function  $h$  with a Lipschitz derivative  $h'$ , [11] obtained the bounds

$$\|f\| \leq \frac{1}{\lambda} \|h'\|, \quad \|f'\| \leq \|h'\|, \quad \|f''\| \leq \frac{2\lambda}{3} \|h'\| + \frac{2}{3} \|h''\|,$$

which have better constants than (3.14) and (3.15).

**3.1.3 Beta distribution**

1. Let  $Z \sim \mathcal{B}(\alpha, \beta)$ ,  $\alpha, \beta > 0$ , with density  $p(x) = \frac{1}{B(\alpha, \beta)}x^{\alpha-1}(1-x)^{\beta-1}$ ,  $0 < x < 1$ , where  $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$  is the beta function.
2. The Stein equation is (see [51])

$$L_{\alpha, \beta}f(x) = x(1-x)f'(x) + (\alpha - (\alpha + \beta)x)f(x) = \tilde{h}(x). \tag{3.17}$$

and, for  $k \geq 1$ ,

$$L_{\alpha+k, \beta+k}f^{(k)}(x) = h^{(k)}(x) + k(\alpha + \beta + k - 1)f^{(k-1)}(x).$$

We have that  $(L_k, T_k) = (L_{\alpha+k, \beta+k}, (\alpha + \beta + 2k)I)$  for  $k \geq 0$ . We can therefore bound the derivatives of the solution via Lemma 2.3, with  $a_k = \sum_{i=0}^k (\alpha + \beta + 2i) = k(\alpha + \beta + k - 1)$ .

3. The solution to (3.17) is

$$f(x) = \frac{1}{x^\alpha(1-x)^\beta} \int_0^x t^{\alpha-1}(1-t)^{\beta-1}\tilde{h}(t) dt.$$

4. It was shown by [11] that, if  $h$  is bounded,

$$\|f\| \leq \frac{\|\tilde{h}\|}{2m(1-m)p(m)},$$

where  $m$  is the median of the  $\mathcal{B}(\alpha, \beta)$  distribution. Also, for Lipschitz  $h$  with Lipschitz constant  $\|h'\|$ ,

$$\|f\| \leq \frac{\|h'\|}{\alpha + \beta} \quad \text{and} \quad \|f'\| \leq C(\alpha, \beta)\|h'\|,$$

where

$$C(\alpha, \beta) = 2(\alpha + \beta) \begin{cases} B(\alpha, \beta), & \alpha \leq 1, \beta \leq 1 \\ \alpha^{-1}, & \alpha \leq 1, \beta > 1 \\ \beta^{-1}, & \alpha > 1, \beta \leq 1 \\ \alpha^{-1}\beta^{-1}B(\alpha, \beta)^{-1}, & \alpha > 1, \beta > 1 \end{cases} \quad \text{if } \alpha \neq \beta \text{ and}$$

$$C(\alpha, \alpha) = \begin{cases} 4, & 0 < \alpha < 1 \\ \frac{2\alpha\sqrt{\pi}\Gamma(\alpha)}{\Gamma(\alpha+1/2)}, & \alpha \geq 1. \end{cases}$$

A bound of the form  $\|f'\| \leq K(\alpha, \beta)\|h'\|$  was also obtained by [30], in which the constant  $K(\alpha, \beta)$ , for various parameter values, is sometimes better and sometimes worse than  $C(\alpha, \beta)$ . An iterative approach with base case  $\|f'\| \leq C(\alpha, \beta)\|h'\|$  was used by [11] to obtain the bound, for  $h \in C_b^n((0, 1))$ ,  $n \geq 1$ ,

$$\|f^{(n)}\| \leq \sum_{j=1}^n j!(\alpha + \beta + j - 1)_{n-j} \left( \prod_{l=1}^n C(\alpha + l - 1, \beta + l - 1) \right) \|h^{(j)}\|.$$

5. For the special case of the  $\mathcal{B}(1/2, 1/2)$  distribution, also known as the arcsine distribution, [10] showed that

$$\|f'\| \leq \min \left( \|h'\|, \frac{\|\tilde{h}\|}{\pi} \right), \quad \|f'\| \leq 4\|h'\|, \quad \|f''\| \leq 6\pi\|h'\| + \frac{3}{2}\pi\|h''\|.$$

**3.1.4 Student's  $t$  distribution**

1. For  $\delta > 0$  and  $d > 0$ , the non-standardised Student's  $t$ -distribution has p.d.f.

$$p(x) = \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\pi}\delta^2\Gamma(\frac{d}{2})} \left(1 + \frac{x^2}{\delta^2}\right)^{-\frac{1}{2}(d+1)}, \quad x \in \mathbb{R}. \tag{3.18}$$

In the case  $d = \nu$ ,  $\delta = \sqrt{\nu}$  the density (3.18) is that of Student's  $t$ -distribution with  $\nu$  degrees of freedom.

2. The Stein equation is (see [20], Section 7)

$$L_{d,\delta}f(x) = (\delta^2 + x^2)f'(x) - (d-1)xf(x) = \tilde{h}(x) \tag{3.19}$$

and, for  $k \geq 1$ ,

$$L_{d-2k,\delta}f^{(k)}(x) = h^{(k)}(x) + k(d-k)f^{(k-1)}(x). \tag{3.20}$$

We have that  $(L_k, T_k) = (L_{d-2k,\delta}, (d-2k-1)I)$  for  $k \geq 0$ . We can therefore bound the derivatives of the solution via Lemma 2.3, with  $a_k = \sum_{i=0}^k (d-2i-1) = (k+1)(d-k-1)$ .

3. The solution to (3.19) is

$$f(x) = (\delta^2 + x^2)^{(d-1)/2} \int_{-\infty}^x (\delta^2 + t^2)^{-(d+1)/2} \tilde{h}(t) dt.$$

4. The density  $p$  satisfies  $(s(x)p(x))' = \tau(x)p(x)$  with  $s(x) = \delta^2 + x^2$  and  $\tau(x) = -(d-1)x$ , and therefore from Lemmas 1 and 3 of [50] we have the bounds

$$\|f\| \leq \frac{\sqrt{\pi}\Gamma(\frac{d}{2})}{2\delta\Gamma(\frac{d+1}{2})} \|\tilde{h}\| \tag{3.21}$$

and

$$\|f'\| \leq 2\|(\delta^2 + x^2)^{-1}\| \|\tilde{h}\| = \frac{2}{\delta^2} \|\tilde{h}\|. \tag{3.22}$$

5. We can use Lemma 2.3 with (3.21) and (3.22) as base cases to bound higher order derivatives of the solution.

Suppose that  $h \in C_b^n(\mathbb{R})$  and  $d - 2n > 0$ . Then

$$\|f^{(n)}\| \leq \sum_{j=0}^n \left(\frac{\sqrt{\pi}}{2\delta}\right)^{n+1-j} \frac{n!}{j!} (d-n)_{n-j} A_j \|\tilde{h}^{(j)}\|,$$

where  $A_j = \prod_{i=j}^n \frac{\Gamma(\frac{d}{2}-i)}{\Gamma(\frac{d+1}{2}-i)}$ .

Suppose now that  $h \in C_b^{n-1}(\mathbb{R})$  and  $d - 2(n-1) > 0$ . Then, for  $n = 2k + 1$ ,

$$\|f^{(2k+1)}\| \leq \sum_{j=0}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \frac{(2k)!!}{(2j)!!} (d-2k)_{k-j,2} \|\tilde{h}^{(2j)}\|$$

and, for  $n = 2k$ ,

$$\begin{aligned} \|f^{(2k)}\| &\leq \sum_{j=1}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \frac{(2k-1)!!}{(2j-1)!!} (d-2k+1)_{k-j,2} \|h^{(2j-1)}\| \\ &\quad + \frac{\sqrt{\pi}\Gamma(\frac{d}{2})}{2\delta\Gamma(\frac{d+1}{2})} \left(\frac{2}{\delta^2}\right)^k (2k-1)!! (d-2k+1)_{k,2} \|\tilde{h}\|. \end{aligned}$$

### 3.1.5 Inverse-gamma distribution

1. Consider an inverse-gamma random variable  $Z$  with parameters  $\alpha, \beta > 0$  and density  $p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}$ ,  $x > 0$ .
2. The Stein equation is (see [35])

$$L_{\alpha,\beta} f(x) = x^2 f'(x) + (\beta - (\alpha - 1)x) f(x) = \tilde{h}(x) \tag{3.23}$$

and, for  $k \geq 1$ ,

$$L_{\alpha-2k,\beta} f^{(k)}(x) = h^{(k)}(x) + k(\alpha - k) f^{(k-1)}(x).$$

We have that  $(L_k, T_k) = (L_{\alpha-2k,\beta}, (\alpha - 2k - 1)I)$  for  $k \geq 0$ . We can therefore bound the derivatives of the solution via Lemma 2.3, with  $a_k = \sum_{i=0}^k (\alpha - 2i - 1) = (k + 1)(\alpha - k - 1)$ .

3. The solution to (3.23) is

$$f(x) = x^{\alpha-1} e^{\beta/x} \int_0^x \tilde{h}(t) t^{-\alpha-1} e^{-\beta/t} dt.$$

4. The density  $p$  satisfies  $(s(x)p(x))' = \tau(x)p(x)$  with  $s(x) = x^2$  and  $\tau(x) = \beta - (\alpha - 1)x$ . Therefore using Lemma 1 of [50] together with the fact that  $\mathbb{E}Z = \frac{\beta}{\alpha-1}$  for  $\alpha > 1$ , we have the bound, for  $\alpha > 1$ ,

$$\|f\| \leq \frac{\Gamma(\alpha)}{\beta} \left( \frac{e}{\alpha - 1} \right)^{\alpha-1} \|\tilde{h}\|. \tag{3.24}$$

5. We can use Lemma 2.3 with base case (3.24) to bound higher order derivatives of the solution.

Suppose that  $h \in C_b^n(\mathbb{R}_+)$  and  $\alpha > 2n + 1$ . Then

$$\|f^{(n)}\| \leq \sum_{j=0}^n \frac{n!}{j!} (\alpha - n)_{n-j} B_j \|\tilde{h}^{(j)}\|,$$

where  $B_j = \prod_{i=j}^n \frac{\Gamma(\alpha-2i)}{\beta} \left( \frac{e}{\alpha-2i-1} \right)^{\alpha-2i-1}$ .

### 3.1.6 PRR (Pékoz-Röllin-Ross) distribution

1. Considered the PRR distribution (introduced by [46]) with density

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

where  $U(a, b, x)$  denotes the confluent hypergeometric function of the second kind ([43], Chapter 13).

2. The Stein equation is (see [46])

$$L_s f(x) = s f''(x) - x f'(x) - 2(s - 1) f(x) = \tilde{h}(x) \tag{3.25}$$

and, for  $k \geq 1$ ,

$$L_s f^{(k)}(x) = h^{(k)}(x) + k f^{(k)}(x).$$

We have that  $(L_k, T_k) = (L_s, D)$  for  $k \geq 0$ . We can therefore bound the derivatives of the solution via Lemma 2.4, with  $a_k = k + 1$ .

3. The solution to (3.25) is

$$f(x) = \frac{1}{s} V_s(x) \int_0^x \frac{1}{V_s(y) \kappa_s(y)} \int_0^y \tilde{h}(t) \kappa_s(t) dt dy,$$

where  $V_s(x) = U\left(s - 1, \frac{1}{2}, \frac{x^2}{2s}\right)$ .

4. The following bounds were obtained by [46]. For  $s \geq 1$ ,

$$\|f'\| \leq \sqrt{2\pi} \|h\|, \quad \|f''\| \leq 2\left(\pi\sqrt{s} + \frac{1}{s}\right) \|h\|, \quad \|f^{(3)}\| \leq 8\left(s + \frac{1}{4} + \sqrt{\frac{\pi}{2}}\right) \|h'\|$$

and, for  $s = 1/2$ ,

$$\|f''\| \leq 4 \|h\|, \quad \|f^{(3)}\| \leq 4 \|h'\|.$$

5. We can apply Lemma 2.4 with  $\|f'\| \leq \sqrt{2\pi} \|h\|$  and  $\|f''\| \leq 4 \|h\|$  as base cases to bound higher order derivatives of the solution.

Suppose  $h \in C_b^{n-1}(\mathbb{R}_+)$ . Then, for  $s \geq 1$ ,

$$\|f^{(n)}\| \leq (n-1)! \sum_{j=0}^{n-1} \frac{(2\pi)^{(n-j)/2}}{j!} \|h^{(j)}\|.$$

Suppose now that  $h \in C_b^{n-2}(\mathbb{R}_+)$ . Then, for  $n = 2k + 1$ ,

$$\|f^{(2k+1)}\| \leq \sqrt{2\pi} C^k (2k-1)!! \|h\| + \sum_{j=1}^{k-1} C^{k-j+1} \frac{(2k-1)!!}{(2j-1)!!} \|h^{(2j-1)}\|$$

and, for  $n = 2k$ ,

$$\|f^{(2k)}\| \leq \sum_{j=0}^{k-1} C^{k-j+1} \frac{(2k)!!}{(2j)!!} \|h^{(2j)}\|,$$

where  $C = 4$  if  $s = 1/2$  and  $C = 2(\sqrt{\pi}s + s^{-1})$  if  $s \geq 1$ .

### 3.1.7 Variance-gamma distribution

1. The  $\text{VG}(r, \theta, \sigma, \mu)$  distribution with parameters  $r > 0, \theta \in \mathbb{R}, \sigma > 0, \mu \in \mathbb{R}$  has p.d.f.

$$p(x) = \frac{1}{\sigma\sqrt{\pi}\Gamma(\frac{r}{2})} e^{\frac{\theta}{\sigma^2}(x-\mu)} \left(\frac{|x-\mu|}{2\sqrt{\theta^2 + \sigma^2}}\right)^{\frac{r-1}{2}} K_{\frac{r-1}{2}}\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|x-\mu|\right), \quad (3.26)$$

where  $x \in \mathbb{R}$  and the modified Bessel function of the second kind is given, for  $x > 0$ , by  $K_\nu(x) = \int_0^\infty e^{-x \cosh(t)} \cosh(\nu t) dt$  (see [43]). The support of the variance-gamma distributions is  $\mathbb{R}$  when  $\sigma > 0$ , but in the limit  $\sigma \rightarrow 0$  the support is the region  $(\mu, \infty)$  if  $\theta > 0$ , and is  $(-\infty, \mu)$  if  $\theta < 0$ . The class of variance-gamma distributions contains many classical distributions as special cases (see [21] for a list of these cases); in particular,

$$\begin{aligned} N(\mu, \sigma^2) &\stackrel{\mathcal{D}}{=} \lim_{r \rightarrow \infty} \text{VG}(r, 0, \sigma/\sqrt{r}, \mu), \\ \Gamma(r, \lambda) &\stackrel{\mathcal{D}}{=} \lim_{\sigma \downarrow 0} \text{VG}(2r, (2\lambda)^{-1}, \sigma, 0), \\ \text{Laplace}(\mu, \sigma) &\stackrel{\mathcal{D}}{=} \text{VG}(2, 0, \sigma, \mu), \\ \text{PN}(2, \sigma^2) &\stackrel{\mathcal{D}}{=} \text{VG}(1, 0, \sigma, 0), \end{aligned}$$

where  $\text{Laplace}(\mu, \sigma)$  denotes a Laplace distribution with density  $\frac{1}{2\sigma} \exp\left(-\frac{|x-\mu|}{\sigma}\right)$ ,  $x \in \mathbb{R}$ , and  $\text{PN}(2, \sigma_X^2, \sigma_Y^2)$  is the distribution of the product of the independent random variables  $X \sim N(0, \sigma_X^2)$  and  $Y \sim N(0, \sigma_Y^2)$ .

2. For simplicity, consider the case  $\mu = 0$ . The Stein equation is (see [21])

$$L_{r,\theta,\sigma}f(x) = \sigma^2 x f''(x) + (\sigma^2 r + 2\theta x) f'(x) - (r\theta - x) f(x) = \tilde{h}(x). \tag{3.27}$$

Note that the  $VG(r, \theta, \sigma, 0)$  Stein equation reduces to the standard normal (3.1) and gamma (3.9) Stein equations in the appropriate limits.

For  $k \geq 1$ , we have

$$L_{r+k,\theta,\sigma}f^{(k)}(x) = h^{(k)}(x) + k f^{(k-1)}(x) + k\theta f^{(k)}(x).$$

We have that  $(L_k, T_k) = (L_{r+k,\theta,\sigma}f^{(k)}, I + \theta D)$  for  $k \geq 0$ . When  $\theta = 0$  we can bound the derivatives of solution using Lemma 2.3, with  $a_k = k + 1$ . Otherwise, we can obtain estimates via Lemma 2.5, with  $a_k = k + 1$  and  $b_k = \theta(k + 1)$ .

3. The solution to (3.27) is

$$f(x) = -\frac{e^{-\theta x/\sigma^2}}{\sigma^2|x|^\nu} K_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|x|\right) \int_0^x e^{\theta t/\sigma^2} |t|^\nu I_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|t|\right) \tilde{h}(t) dt - \frac{e^{-\theta x/\sigma^2}}{\sigma^2|x|^\nu} I_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|x|\right) \int_x^\infty e^{\theta t/\sigma^2} |t|^\nu K_\nu\left(\frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}|t|\right) \tilde{h}(t) dt, \tag{3.28}$$

where  $\nu = \frac{r-1}{2}$  and  $I_\nu(x) = \sum_{k=0}^\infty \frac{1}{k! \Gamma(\nu+k+1)} \left(\frac{x}{2}\right)^{\nu+2k}$  is a modified Bessel function of the first kind (see [43]).

4. For the case  $\theta = 0$ , [21] obtained the estimates

$$\|f\| \leq \frac{1}{\sigma} \left( \frac{1}{r} + \frac{\pi \Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})} \right) \|\tilde{h}\|, \tag{3.29}$$

$$\|f'\| \leq \frac{2}{\sigma^2 r} \|\tilde{h}\|. \tag{3.30}$$

as well as uniform bounds for  $f''$ ,  $f^{(3)}$  and  $f^{(4)}$ .

5. An application of Lemma 2.3 with (3.29) and (3.30) as base cases allows us to obtain bounds for higher order derivatives.

Suppose  $h \in C_b^{n-1}(\mathbb{R})$ . Then, for  $n = 2k + 1$ ,

$$\|f^{(2k+1)}\| \leq \sum_{j=0}^k \frac{(2k)!!}{(2j)!!} \frac{1}{(r + 2j + 2)_{k-j+1,2}} \left(\frac{2}{\sigma^2}\right)^{k-j+1} \|\tilde{h}^{(2j)}\|$$

and, for  $n = 2k$ ,

$$\|f^{(2k)}\| \leq \sum_{j=1}^k \frac{(2k-1)!!}{(2j-1)!!} \frac{1}{(r + 2j + 1)_{k-j+1,2}} \left(\frac{2}{\sigma^2}\right)^{k-j+1} \|h^{(2j-1)}\| + \frac{(2k-1)!! 2^k}{(r+1)_{k,2} \sigma^{2k+1}} \left(\frac{1}{r} + \frac{\pi \Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})}\right) \|\tilde{h}\|.$$

6. Now, consider the case of general  $\theta$ . In Section 3.2.4, we obtain the estimates

$$\|f\| \leq \frac{\|\tilde{h}\|}{\sqrt{\theta^2 + \sigma^2}} \left(\frac{2}{r} + C(r, \theta, \sigma)\right), \tag{3.31}$$

$$\|f'\| \leq \frac{\|\tilde{h}\|}{\sigma^2} \left(\frac{2}{r} + C(r, \theta, \sigma)\right), \tag{3.32}$$

where

$$C(r, \theta, \sigma) = \begin{cases} \frac{\sqrt{\pi} \Gamma(\frac{r}{2})}{\Gamma(\frac{r+1}{2})} \left(1 + \frac{\theta^2}{\sigma^2}\right)^{r/2}, & r \geq 2, \\ 6\Gamma\left(\frac{r}{2}\right) \left(1 - \frac{1}{(1 + \sigma^2/\theta^2)^{1/2}}\right)^{-r/2}, & 0 < r < 2. \end{cases}$$

7. We can use Lemma 2.5 with (3.31) and (3.32) as base cases to obtain bounds for higher order derivatives.

Let  $B_j = \frac{2}{r+j} + C(r+j, \theta, \sigma)$ . Then, for  $h \in C_b^{n-1}(\mathbb{R})$ ,

$$\begin{aligned} \|f^{(n)}\| &\leq \frac{1}{\sigma^2} \sum_{j=0}^{n-1} \left( B_{n-j} \sum_{l \in I_j} A_{j,l}^{(n)} \right) \|h^{(n-j)}\| \\ &\quad + \frac{B_0}{\sigma^2} \left( \sum_{l \in I_n} A_{n,l}^{(n)} + \frac{B_1}{\sqrt{\theta^2 + \sigma^2}} \sum_{l \in I_{n-1}} A_{n-1,l}^{(n)} \right) \|\tilde{h}\|. \end{aligned}$$

where

$$A_{j,l}^{(n)} := \sum_{M \in \mathcal{S}_{j,l}^{(n)}} \prod_{i \in M \setminus \{n-j\}} (|\theta| 1_{\{i-1 \in M\}} + 1_{\{i-1 \notin M\}}) \frac{i B_i}{\sigma^2},$$

and the sets  $I_j$  and  $\mathcal{S}_{j,l}^{(n)}$  are defined as in Lemma 2.5.

In particular, applying the iterative technique with (3.31) and (3.32) as base cases yields the following simple bounds:

$$\begin{aligned} \|f''\| &\leq \frac{B_1}{\sigma^2} \|h'\| + \frac{B_0 B_1}{\sigma^2} \left[ \frac{|\theta|}{\sigma^2} + \frac{1}{\sqrt{\theta^2 + \sigma^2}} \right] \|\tilde{h}\|, \\ \|f^{(3)}\| &\leq \frac{B_2}{\sigma^2} \|h''\| + \frac{2|\theta| B_1 B_2}{\sigma^4} \|h'\| + \frac{2B_0 B_2}{\sigma^4} \left[ 1 + \frac{|\theta| B_1}{\sqrt{\theta^2 + \sigma^2}} + \frac{\theta^2 B_1}{\sigma^2} \right] \|\tilde{h}\|, \\ \|f^{(4)}\| &\leq \frac{B_3}{\sigma^2} \|h^{(3)}\| + \frac{3|\theta| B_2 B_{r+3}}{\sigma^4} \|h''\| + \frac{3B_1 B_3}{\sigma^4} \left[ 1 + \frac{2\theta^2 B_2}{\sigma^2} \right] \|h'\| \\ &\quad + \frac{3B_3}{\sigma^4} \left[ \frac{|\theta| B_0}{\sigma^2} \left( B_1 + 2\theta B_2 + \frac{2\theta^2 B_1 B_2}{\sigma^2} \right) + \frac{B_1}{\sqrt{\theta^2 + \sigma^2}} \left( 1 + \frac{2\theta^2 B_2}{\sigma^2} \right) \right] \|\tilde{h}\|. \end{aligned}$$

### 3.1.8 Density proportional to $e^{-x^4/12}$

1. Consider a probability distribution with density proportional to  $e^{-x^4/12}$ . We quote from [6] that the exact form of the density  $p$  of this distribution is

$$p(x) = \frac{\sqrt{2}}{3^{1/4} \Gamma(1/4)} e^{-x^4/12}, \quad x \in \mathbb{R}.$$

The limiting distribution of the properly scaled total magnetization in the Curie-Weiss model at the critical inverse temperature  $\beta = 1$  was shown to have this density in [16]. In [13] and [6] a version of Stein's method of exchangeable pairs was used to prove that also a Berry-Essén rate of order  $n^{-1/2}$  holds for this limit theorem. As in neither of the works [13] and [6] a constant for this rate is aimed at, no numerical bounds on the solution to the Stein equation have been stated, yet. However, in the appendix section of [13], actually some numerical bounds are proved.

We will firstly use these and the general bounds from [11] to provide concrete bounds on the solution to the Stein equation and, then, use these as starting points for our iterative method to derive bounds for derivatives of all orders.

2. The Stein equation is (see [13] and [6])

$$Lf(x) = f'(x) - \frac{x^3}{3} f(x) = \tilde{h}(x). \tag{3.33}$$

We have

$$Lf'(x) = h'(x) + x^2 f(x) =: h_1(x), \tag{3.34}$$

$$Lf''(x) = h''(x) + 2x f(x) + 2x^2 f'(x) =: h_2(x), \tag{3.35}$$

and, for  $k \geq 3$ ,

$$\begin{aligned} Lf^{(k)}(x) &= h^{(k)}(x) + \frac{1}{3}k(k-1)(k-2)f^{(k-3)}(x) + k(k-1)xf^{(k-2)}(x) + kx^2 f^{(k-1)}(x) \\ &=: h_k(x). \end{aligned}$$

We are unable to apply either of Lemmas 2.3–2.5 to bound the derivatives of the solution. However, we can still apply the iterative technique to bound the derivatives; the details are given in Section 3.2.5.

3. The solution to (3.33) is

$$f(x) = e^{x^4/12} \int_{-\infty}^x \tilde{h}(t)e^{-t^4/12} dt.$$

4. If  $h$  is bounded and measurable, then it holds that

$$|f(x)| \leq \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) \|\tilde{h}\|, \quad \|f'\| \leq 2\|\tilde{h}\|. \tag{3.36}$$

These bounds are not stated explicitly, but are actually proved in the appendix section of [13]. If  $h$  is Lipschitz-continuous, then we can prove that

$$\|f\| \leq \frac{\sqrt{3\pi}}{2} \|h'\|, \quad \|f'\| \leq 2^{1/2}3^{1/4}\Gamma(1/4)\|h'\|, \quad \|f''\| \leq 4\|h'\|. \tag{3.37}$$

More precisely, we can even show that for each  $x \in \mathbb{R}$  we have

$$|f(x)| \leq \min\left(\frac{\sqrt{3\pi}}{2}, \frac{3}{x^2}\right) \|h'\|, \quad |f'(x)| \leq \min\left(2^{1/2}3^{1/4}\Gamma(1/4), \frac{12}{|x|^3}\right) \|h'\|. \tag{3.38}$$

We would like to stress that the numerical bounds for Lipschitz test functions  $h$  in (3.37) are original and the first ones obtained for this distribution. The bound on  $f''$  in (3.37) will be proved by standard, but pretty involved calculations. It should thus not be left unmentioned that, using the fairly easy second bound in (3.36) and our iterative procedure, we can easily obtain the bound

$$\|f''\| \leq 8\|h'\|, \tag{3.39}$$

which has a slightly worse constant. This is completely analogous to the case of the normal distribution (see Section 3.2.1).

For bounds on derivatives of general order let  $h \in C_b^n(\mathbb{R})$ . We obtain the following generalization of the first bound in (3.36):

$$|f^{(n)}(x)| \leq \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) \sum_{k=0}^n a_{n-k}^{(n)} \|\tilde{h}^{(n-k)}\|, \quad x \in \mathbb{R}. \tag{3.40}$$

Here, the nonnegative real numbers  $a_j^{(n)}$ ,  $n \in \mathbb{N}_0$ ,  $0 \leq j \leq n$  are given as follows: For each  $n$  we have that

$$a_n^{(n)} = 1, \quad a_{n-1}^{(n)} = \frac{3n}{(6c_1)^{1/3}}, \quad a_{n-2}^{(n)} = \frac{12n(n-1)}{(6c_1)^{2/3}} \tag{3.41}$$

whenever these quantities are defined, and, for  $3 \leq k \leq n$ ,

$$a_{n-k}^{(n)} = \frac{3n}{(6c_1)^{1/3}} a_{n-k}^{(n-1)} + \frac{3n(n-1)}{(6c_1)^{2/3}} a_{n-k}^{(n-2)} + \frac{n(n-1)(n-2)}{6c_1} a_{n-k}^{(n-3)}, \quad (3.42)$$

where we write

$$c_1 = \frac{\sqrt{2}}{3^{1/4}\Gamma(1/4)}$$

for the normalizing constant. Since we can also show that

$$\|h_n\| \leq \sum_{k=0}^n a_{n-k}^{(n)} \|\tilde{h}^{(n-k)}\|, \quad (3.43)$$

from the second bound in (3.36) we obtain the following generalised version of this bound:

$$\|f^{(n+1)}\| \leq 2 \sum_{k=0}^n a_{n-k}^{(n)} \|\tilde{h}^{(n-k)}\|. \quad (3.44)$$

### 3.1.9 Multivariate normal distribution

1. Let  $Z \sim \text{MVN}(0, \Sigma)$  with density  $p(x) = \frac{1}{(2\pi)^{d/2} \sqrt{\det(\Sigma)}} \exp(-\frac{1}{2}x^T \Sigma^{-1}x)$ ,  $x \in \mathbb{R}^d$ . The covariance matrix  $\Sigma \in \mathbb{R}^{d \times d}$  is nonnegative-definite.
2. The Stein equation is (see [31])

$$Lf(x) = \sum_{i=1}^d \sum_{j \neq i}^d \sigma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) - \sum_{i=1}^d x_i \frac{\partial f}{\partial x_i}(x) = \tilde{h}(x), \quad (3.45)$$

where  $\sigma_{ij} = (\Sigma)_{ij}$ , and, for  $k \geq 1$ ,

$$L\left(\frac{\partial^k f(x)}{\prod_{j=1}^k \partial x_{i_j}}\right) = \frac{\partial^k h(x)}{\prod_{j=1}^k \partial x_{i_j}} + k \frac{\partial^k f(x)}{\prod_{j=1}^k \partial x_{i_j}}.$$

We are unable to apply either of Lemmas 2.3–2.5 to bound the derivatives of the solution. However, we can still apply the iterative technique to bound the derivatives; the details are given in Section 3.2.6.

3. The solution to (3.45) is

$$f(x) = - \int_0^\infty [\mathbb{E}h(e^{-s}x + \sqrt{1 - e^{-2s}}Z) - \mathbb{E}h(Z)] ds.$$

4. It was shown by [31] that, if  $\Sigma$  is nonnegative-definite and  $h \in C_b^n(\mathbb{R}^d)$ , then the solution satisfies

$$\left\| \frac{\partial^n f}{\prod_{j=1}^n \partial x_{i_j}} \right\| \leq \frac{1}{n} \left\| \frac{\partial^n h}{\prod_{j=1}^n \partial x_{i_j}} \right\|, \quad n \geq 1. \quad (3.46)$$

Suppose now that  $\Sigma$  is positive-definite matrix and that  $h$  is bounded. Then, the following estimate was obtained by [23]:

$$\left\| \frac{\partial f}{\partial x_i} \right\| \leq \sqrt{\frac{\pi}{2}} \left[ \sum_{j=1}^d \sigma_{ij}^2 \right]^{1/2} \|\tilde{h}\|, \quad (3.47)$$

If  $h \in C_b^{n-1}(\mathbb{R}^d)$ ,  $n \geq 2$ , then

$$\left\| \frac{\partial^n f}{\prod_{j=1}^n \partial x_{i_j}} \right\| \leq \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\Gamma(\frac{n+1}{2})} \min_{1 \leq l \leq n} \left\{ \left[ \sum_{j=1}^d \sigma_{ij}^2 \right]^{1/2} \left\| \frac{\partial^{n-1} h}{\prod_{\substack{1 \leq j \leq n \\ j \neq l}} \partial x_{i_j}} \right\| \right\}. \quad (3.48)$$

Analogous bounds to (3.46) and (3.48) were obtained for the  $n$ -th derivatives of the solution  $f$ , viewed as  $n$ -linear forms, by [41] and [23] respectively.

### 3.2 Further comments and proofs

In this section, we provide additional comments to complement the results of Section 3.1, and we also provide insights into the application of the iterative technique in practice.

#### 3.2.1 Normal distribution

Using elementary calculations, [53] obtained the bounds (3.5) for the solution (3.4) of the standard normal Stein equation. However, this approach quickly becomes tedious for bounding higher order derivatives. The first progress with regard to bounding higher order derivatives came from the observation by [3] that the standard normal Stein operator is the generator of an Ornstein-Uhlenbeck process (if we take  $f = g'$ ). Applying standard theory of Markov processes leads to a solution of the form (1.4):

$$g(x) = - \int_0^\infty [\mathbb{E}h(e^{-s}x + \sqrt{1 - e^{-2s}}Z) - \mathbb{E}h(Z)] ds, \quad (3.49)$$

If  $h \in C_b^k(\mathbb{R})$  then, by the dominated convergence theorem,

$$f^{(k-1)}(x) = g^{(k)}(x) = - \int_0^\infty e^{-ks} \mathbb{E}[h^{(k)}(e^{-s}x + \sqrt{1 - e^{-2s}}Z)] ds, \quad (3.50)$$

and the bound (3.6) now follows easily. The bound (3.7) results from an integration by parts on (3.50). However, one has to work significantly harder to obtain (3.8) and quite involved calculations were required to achieve the bound.

Let us now consider the application of our iterative method to the standard normal Stein equation. Using part (i) of Lemma 2.3 with  $\|f\| \leq \sqrt{\frac{\pi}{2}}\|\tilde{h}\|$  as the base case, we have that, for  $C_b^{(n)}(\mathbb{R})$ ,

$$\|f^{(n)}\| \leq \sum_{j=0}^n \left(\frac{\pi}{2}\right)^{(n-j+1)/2} \frac{n!}{j!} \|\tilde{h}^{(j)}\|,$$

and an application of part (ii) of Lemma 2.3 with base case  $\|f'\| \leq 2\|\tilde{h}\|$  leads to a bound on  $f^{(n)}$  involving one fewer derivative of  $h$ . These bounds are much worse than the best bounds from the literature, as given in Section 3.1.1. However, the iterative technique does allow us to deduce bounds for derivatives of any order of the solution given just the easily derived bounds  $\|f\| \leq \sqrt{\frac{\pi}{2}}\|\tilde{h}\|$  and  $\|f'\| \leq 2\|\tilde{h}\|$ .

An interesting application of the iterative method is that it allows us to easily deduce a bound of the form  $\|f^{(n+1)}\| \leq C\|h^{(n)}\|$ , where  $C$  is a universal constant, given just the bounds  $\|f'\| \leq 2\|\tilde{h}\|$  and  $\|f^{(n-1)}\| \leq \frac{1}{n}\|h^{(n)}\|$ . Recall that, for  $k \geq 1$ ,

$$Lf^{(k)}(x) = h^{(k)}(x) + kf^{(k-1)}(x).$$

Then, an application of the iterative procedure with these bounds yields

$$\|f^{(n+1)}\| \leq 2\|h^{(n)}(x) + nf^{(n-1)}(x)\| \leq 2(\|h^{(n)}\| + n\|f^{(n-1)}\|) \leq 4\|h^{(n)}\|.$$

The constant 4 is worse than the constant 2 in the bound (3.8) obtained by [8], but our derivation is considerably simpler. This example also illustrates how the iterative technique can be used to deduce bounds of the form  $\|f^{(n+1)}\| \leq C\|h^{(n)}\|$  from bounds of the form  $\|f^{(n)}\| \leq K\|h^{(n)}\|$ . This is of course useful in practice, because bounds involving fewer derivatives of the test function  $h$  are often desirable.

### 3.2.2 Gamma distribution

As is the case with the normal distribution, bounding derivatives of the solution of the gamma Stein equation using the representation (3.12) of the solution is tedious. The bound (3.16) for the solution can be obtained by this approach, but bounds for derivatives of the solution have not been established through this representation of the solution. However, [39] used a generator approach, similar to that described for the normal distribution, to obtain a formula for the  $n$ -th derivative of the solution of the Stein equation, from which the bound (3.13) follows. Through a technical calculation, [20] obtained (3.14) (this improved an earlier bound of [47]). The bound (3.15) was obtained by an iterative approach similar to that described in this paper.

As was the case for the normal distribution, a direct application of part (i) of Lemma 2.3 with (3.16) as the base case would yield bounds on the higher order derivatives of the solution that are much worse than the best bounds from the literature. However, the difficulty of bounding higher order derivatives is again greatly reduced, because one only requires the bound (3.16) to achieve such bounds.

For the gamma distribution, we can again use the iterative technique to obtain bounds of the form  $\|f^{(n)}\| \leq C\|h^{(n)}\|$  from bounds of the form  $\|f^{(n)}\| \leq K\|h^{(n+1)}\|$ . Using an iterative approach with inequalities (3.16) and (3.13) gives

$$\begin{aligned} \|f^{(n)}\| &\leq \frac{e^{r+n}\Gamma(r+n)}{(r+n)^{r+n}} \|h^{(n)}(x) + n\lambda f^{(n-1)}(x)\| \\ &\leq \frac{e^{r+n}\Gamma(r+n)}{(r+n)^{r+n}} (\|h^{(n)}\| + n\lambda\|f^{(n-1)}\|) \\ &\leq \frac{2e^{r+n}\Gamma(r+n)}{(r+n)^{r+n}} \|h^{(n)}\|. \end{aligned} \tag{3.51}$$

Thus, the iterative method has yielded a bound of similar form to (3.14), but with a much simpler proof. Using Stirling's approximation, we see that (3.51) is of order  $(r+n)^{-1/2}$  as  $r+n \rightarrow \infty$ , which is the same order as (3.14). However, the constant is better in (3.14) than (3.51) for all  $n \geq 1$ , whilst the constant in (3.51) is better than bound (3.14) in the case  $n = 0$ .

### 3.2.3 Beta, Student's $t$ , inverse-gamma and the PRR distribution

Unlike the normal and gamma distributions, there are currently no simple formulas in the literature for general  $n$ -th order derivatives of the solutions of the beta, Student's  $t$ , inverse-gamma and PRR Stein equations. However, for each of these distributions uniform bounds are available for the solution and some of its lower order derivatives (beta and PRR) or it is straightforward to derive such bounds (Student's  $t$  and inverse-gamma). Therefore, the iterative method of this paper is particularly useful for these distributions.

From a straightforward application of Lemmas 2.3 and 2.4 we arrive at our bounds for  $n$ -th order derivatives of the solutions of the non-standardised Student's  $t$ , inverse-gamma and PRR Stein equations respectively. We could also apply Lemma 2.3 to bound the higher order derivatives of the solution of the beta Stein equation; however, this was already done by [11]. The bound of [11] is slightly different from the one that would arise from an application of Lemma 2.3, as in that work the base case was  $\|f'\| \leq C(\alpha, \beta)\|h'\|$ , rather than a base case of the form  $\|f\| \leq K(\alpha, \beta)\|\tilde{h}\|$ .

The non-standardised Student's  $t$  example illustrates a pertinent point regarding the application of the iterative method. Suppose that we are interested in bounding the derivatives of solution of the classical Student's  $t$  Stein equation ( $d = \nu$ ,  $\delta = \sqrt{\nu}$ ). By (3.20), we have that, for  $k \geq 1$ , the solution satisfies

$$L_{\nu-2k, \sqrt{\nu}} f^{(k)}(x) = h^{(k)}(x) + k(\nu - k) f^{(k-1)}(x).$$

It is therefore clear that in order to use the iterative approach to bound higher order derivatives of the solution of the Student's  $t$  Stein equation, we must work with the Stein equation for the more general non-standardised Student's  $t$  distribution. It was also observed by [11] that in order to apply the iterative method to the exponential Stein equation, one must work with the more general gamma Stein equation.

### 3.2.4 Variance-gamma distribution

For the case  $\theta = 0$ , we were able to bound  $n$ -th order derivatives of the solution of the variance-gamma Stein equation via Lemma 2.3 with inequalities (3.29) and (3.30) as base cases. Previously, only uniform bounds existed for the solution and its first four derivatives (see [21]). Our bounds for the  $n$ -th order derivatives decay at a rate of order  $(r + n)^{-1}$  as  $r + n \rightarrow \infty$ ; better than the order  $r^{-1/2}$  rate obtained by [21], and our bounds also have smaller constants. Consequently, our bounds allows for an improvement of the  $O(r^{1/2}(m^{-1} + n^{-1}))$  bound of Theorem 4.9 of [21] to a  $O(m^{-1} + n^{-1})$  bound.

For the case of general  $\theta$ , only modest progress had been made on the problem of bounding the derivatives of the solution of the Stein equation. It was shown by [20] and [21] that the solution and its first derivative are bounded if the test function  $h$  is bounded, although explicit constants were not given. The complicated form of the solution (3.28) means that bounding higher order derivatives directly is a challenging problem; see [20], Section 3 and Appendix D.

This application clearly demonstrates the power of our method: bounding just the solution and its first derivative, and then applying Lemma 2.5 yields bounds on derivatives of the solution of any order. This makes the problem much more tractable: the problem is essentially reduced to bounding the solution and its first derivative. We end this section by obtaining such bounds; the bounds of Section 3.1.6 then follow immediately from Lemma 2.5 with inequalities (3.31) and (3.32) as base cases.

By an application of Stirling's inequality, we see that our bounds for  $n$ -th order derivatives of the solution of  $\text{VG}(r, \theta, \sigma, 0)$  Stein equation are of order  $(r + n)^{-1/2}$  as  $r + n \rightarrow \infty$ . This is slower than the  $O(r + n)^{-1}$  rate we obtained for the  $\theta = 0$  case. It is an open problem to improve the bounds for the  $\theta \neq 0$  case to this rate.

Finally, we note an application of our estimates for the solution of the  $\text{VG}(r, \theta, \sigma, 0)$  Stein equation to the work of [15] on variance-gamma approximation on Wiener space. In that work, it is shown that a sequence of distributions of random variables in the second Wiener chaos converge to a variance-gamma distribution if and only if their second to sixth order moments converge to those of a variance-gamma random variable. Bounds on the rate of convergence were also derived using the variance-gamma Stein equation. These bounds were given in terms of an absolute constant, because explicit bounds were not available for derivatives of the solution of the Stein equation. With our explicit bounds on the derivatives, it is possible to give explicit constants in the bounds of [15].

*Proof of inequalities (3.31) and (3.32).* In order to simplify the calculations, we make the following change of parameters

$$\nu = \frac{r - 1}{2}, \quad \alpha = \frac{\sqrt{\theta^2 + \sigma^2}}{\sigma^2}, \quad \beta = \frac{\theta}{\sigma^2}. \tag{3.52}$$

With these parameters, the solution (3.28) can be written as

$$f(x) = -\frac{e^{-\beta x} K_\nu(\alpha|x|)}{\sigma^2|x|^\nu} \int_0^x e^{\beta y} |y|^\nu I_\nu(\alpha|y|) \tilde{h}(y) dy$$

$$- \frac{e^{-\beta x} I_\nu(\alpha|x|)}{\sigma^2|x|^\nu} \int_x^\infty e^{\beta y} |y|^\nu K_\nu(\alpha|y|) \tilde{h}(y) dy \tag{3.53}$$

$$= - \frac{e^{-\beta x} K_\nu(\alpha|x|)}{\sigma^2|x|^\nu} \int_0^x e^{\beta y} |y|^\nu I_\nu(\alpha|y|) \tilde{h}(y) dy + \frac{e^{-\beta x} I_\nu(\alpha|x|)}{\sigma^2|x|^\nu} \int_{-\infty}^x e^{\beta y} |y|^\nu K_\nu(\alpha|y|) \tilde{h}(y) dy. \tag{3.54}$$

The equality between (3.53) and (3.54) means that in order to obtain uniform bounds for all  $x \in \mathbb{R}$  it suffices to bound the solution and its derivative in the region  $x \geq 0$  as long as we consider the case of both negative and positive  $\beta$ . That is, the uniform bound we derive for  $x \geq 0$  also holds for  $x \leq 0$ , and thus all  $x \in \mathbb{R}$ . Straightforward calculations show that, for  $x \geq 0$ ,

$$|f(x)| \leq \frac{\|\tilde{h}\|}{\sigma^2} \left| \frac{e^{-\beta x} K_\nu(\alpha x)}{x^\nu} \int_0^x e^{\beta y} y^\nu I_\nu(\alpha y) dy \right| + \frac{\|\tilde{h}\|}{\sigma^2} \left| \frac{e^{-\beta x} I_\nu(\alpha x)}{x^\nu} \int_x^\infty e^{\beta y} y^\nu K_\nu(\alpha y) dy \right|,$$

$$|f'(x)| \leq \frac{\|\tilde{h}\|}{\sigma^2} \left| \frac{d}{dx} \left( \frac{e^{-\beta x} K_\nu(\alpha x)}{x^\nu} \right) \int_0^x e^{\beta y} y^\nu I_\nu(\alpha y) dy \right| + \frac{\|\tilde{h}\|}{\sigma^2} \left| \frac{d}{dx} \left( \frac{e^{-\beta x} I_\nu(\alpha x)}{x^\nu} \right) \int_x^\infty e^{\beta y} y^\nu K_\nu(\alpha y) dy \right|.$$

We now note the following bounds, which follow almost immediately from Theorems 3.1 and 3.2 of [24]. Fix  $\nu > -\frac{1}{2}$  and  $0 < |\beta| < \alpha$ . Then, for all  $x > 0$ ,

$$\left| \frac{e^{-\beta x} K_\nu(\alpha x)}{x^\nu} \int_0^x e^{\beta y} y^\nu I_\nu(\alpha y) dy \right| < \frac{2}{\alpha(2\nu + 1)},$$

$$\left| \frac{e^{-\beta x} I_\nu(\alpha x)}{x^\nu} \int_x^\infty e^{\beta y} y^\nu K_\nu(\alpha y) dy \right| < \frac{K(\nu, \gamma)}{\alpha},$$

$$\left| \frac{d}{dx} \left( \frac{e^{-\beta x} K_\nu(\alpha x)}{x^\nu} \right) \int_0^x e^{\beta y} y^\nu I_\nu(\alpha y) dy \right| < \frac{2(\gamma + 1)}{2\nu + 1},$$

$$\left| \frac{d}{dx} \left( \frac{e^{-\beta x} I_\nu(\alpha x)}{x^\nu} \right) \int_x^\infty e^{\beta y} y^\nu K_\nu(\alpha y) dy \right| < K(\nu, \gamma),$$

where  $\gamma = \beta/\alpha$  and

$$K(\nu, \gamma) = \begin{cases} \frac{\sqrt{\pi}}{(1 - \gamma^2)^{\nu+1/2}} \frac{\Gamma(\nu + \frac{1}{2})}{\Gamma(\nu + 1)}, & \nu \geq \frac{1}{2}, \\ \frac{6\Gamma(\nu + \frac{1}{2})}{1 - |\gamma|}, & |\nu| < \frac{1}{2}. \end{cases}$$

Applying these inequalities to bound the above expressions for  $|f(x)|$  and  $|f'(x)|$ , and recalling that it suffices to bound the solution for  $x \geq 0$  leads to the following bounds, which hold for all  $x \in \mathbb{R}$ ,

$$|f(x)| \leq \frac{\|\tilde{h}\|}{\sigma^2 \alpha} \left( \frac{2}{2\nu + 1} + K(\nu, \gamma) \right),$$

$$|f'(x)| \leq \frac{\|\tilde{h}\|}{\sigma^2} \left( \frac{2(\gamma + 1)}{2\nu + 1} + K(\nu, \gamma) \right).$$

Using (3.52) to return to the original parameters yields (3.31) and (3.32), as required.  $\square$

### 3.2.5 Density proportional to $e^{-x^4/12}$

We first prove the bounds (3.37) and (3.38). Note that the density  $p$  and its Stein equation (3.33) satisfy the general Conditions 3.1 and 3.2 from the paper [11]. Hence,

by Proposition 3.13 (a) of that paper, we conclude that

$$|f(x)| \leq \frac{\int_x^\infty tp(t) dt}{p(x)} \|h'\| \tag{3.55}$$

holds for each  $x \in \mathbb{R}$ . Writing  $\Phi$  for the standard normal distribution function and  $\varphi = \Phi'$  for its density, it is an easy matter of partial integration to check that

$$p(x) = c_1 \sqrt{2\pi} \varphi\left(\frac{x^2}{\sqrt{6}}\right) \quad \text{and} \quad \int_x^\infty tp(t) dt = c_1 \sqrt{3\pi} \left(1 - \Phi\left(\frac{x^2}{\sqrt{6}}\right)\right), \tag{3.56}$$

where we denote by

$$c_1 := \frac{\sqrt{2}}{3^{1/4}\Gamma(1/4)}$$

the normalizing constant of  $p$ . Hence, from (3.55) we obtain the bound

$$|f(x)| \leq \frac{\sqrt{6}}{2} \frac{1 - \Phi\left(\frac{x^2}{\sqrt{6}}\right)}{\varphi\left(\frac{x^2}{\sqrt{6}}\right)} \|h'\|. \tag{3.57}$$

The Gaussian Mill's ratio inequality states that, for each  $t > 0$ ,

$$\frac{t}{1+t^2} \leq \frac{1 - \Phi(t)}{\varphi(t)} \leq \min\left(\sqrt{\frac{\pi}{2}}, \frac{1}{t}\right). \tag{3.58}$$

Hence, from (3.57) and (3.58) we conclude that

$$|f(x)| \leq \min\left(\frac{\sqrt{3\pi}}{2}, \frac{3}{x^2}\right) \|h'\|, \tag{3.59}$$

which already gives the first bound in (3.37) and (3.38), respectively. For the bounds on  $f'$  we use the once iterated Stein equation (3.34) together with the first bound in (3.36) and (3.59) to obtain

$$\begin{aligned} |f'(x)| &\leq \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) |h'(x) + x^2 f(x)| \\ &\leq \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) \|h'\| + \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) |x^2 f(x)| \\ &\leq \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) \|h'\| + 3 \min\left(\frac{3^{1/4}\Gamma(1/4)}{2\sqrt{2}}, \frac{3}{|x|^3}\right) \|h'\| \\ &= \min\left(2^{1/2}3^{1/4}\Gamma(1/4), \frac{12}{|x|^3}\right) \|h'\|, \end{aligned} \tag{3.60}$$

as claimed. In order to prove the bound on  $f''$ , we rely on the following representation of the second derivative of the Stein equation within the so-called density approach of Stein's method:

$$f''(x) = h'(x) + \frac{U(x)}{p(x)} \int_{-\infty}^x F(s)h'(s) ds + \frac{V(x)}{p(x)} \int_x^\infty (1 - F(s))h'(s) ds, \tag{3.61}$$

where  $F$  is the distribution function corresponding to  $p$  and

$$\begin{aligned} U(x) &= -(1 - F(x))\left(x^2 + \frac{x^6}{9}\right) + \frac{x^3}{3}p(x), \\ V(x) &= -F(x)\left(x^2 + \frac{x^6}{9}\right) - \frac{x^3}{3}p(x). \end{aligned}$$

Identity (3.61) is fairly standard in Stein’s method. A proof can, for instance, be found in the arXiv version of the paper [12]. With the help of the Mill’s ratio type inequalities (5.51)–(5.54) for the distribution corresponding to  $p$  from [13], one can easily see that both  $U$  and  $V$  are non-positive functions. Using the fact that

$$\int_{-\infty}^x F(s) \, ds = xF(x) - \int_{-\infty}^x sp(s) \, ds = xF(x) + \int_x^{\infty} sp(s) \, ds$$

and

$$\int_x^{\infty} (1 - F(s)) \, ds = -x(1 - F(x)) + \int_x^{\infty} sp(s) \, ds$$

we conclude that

$$\begin{aligned} |f''(x)| &\leq \|h'\| + \|h'\| \left( \frac{-U(x)}{p(x)} \int_{-\infty}^x xF(s) \, ds + \frac{-V(x)}{p(x)} \int_x^{\infty} (1 - F(s)) \, ds \right) \\ &= \|h'\| + \|h'\| \left( \left( -\frac{x^3}{3} + \frac{(1 - F(x))(x^2 + \frac{x^6}{9})}{p(x)} \right) \left( xF(x) + \int_x^{\infty} sp(s) \, ds \right) \right. \\ &\quad \left. + \left( \frac{x^3}{3} + \frac{F(x)(x^2 + \frac{x^6}{9})}{p(x)} \right) \left( -x(1 - F(x)) + \int_x^{\infty} sp(s) \, ds \right) \right) \\ &= \|h'\| + \|h'\| \left( -\frac{x^4}{3} + \left( x^2 + \frac{x^6}{9} \right) \frac{\int_x^{\infty} sp(s) \, ds}{p(x)} \right) \\ &\leq \|h'\| + \|h'\| \left( -\frac{x^4}{3} + \left( x^2 + \frac{x^6}{9} \right) \frac{3}{x^2} \right) \\ &= 4\|h'\|, \end{aligned}$$

where we used (3.56) and (3.58), again, to obtain the final inequality. Using the second bound in (3.36) for the iterated Stein equation (3.34) as well as (3.59), we easily obtain

$$|f''(x)| \leq 2|h'(x) + x^2 f(x)| \leq 2\|h'\| + 2 \cdot 3\|h'\| = 8\|h'\|,$$

which is a very simple derivation as compared to that of the previous bound.

Next, we prove (3.40), (3.41) and (3.42) by induction on  $n$ . We omit the cases  $n = 0, 1, 2$  and assume that  $n \geq 3$  and that the claim holds for all  $0 \leq k < n$ . It is easily checked that

$$|x| \min \left( \frac{1}{2c_1}, \frac{3}{|x|^3} \right) \leq \frac{3}{(6c_1)^{2/3}} \quad \text{and} \quad x^2 \min \left( \frac{1}{2c_1}, \frac{3}{|x|^3} \right) \leq \frac{3}{(6c_1)^{1/3}}. \quad (3.62)$$

Hence, from (3.36) applied to  $f_n$ , from the definition of  $h_n$  and the induction hypothesis we conclude that

$$\begin{aligned} |f^{(n)}(x)| &\leq \min \left( \frac{1}{2c_1}, \frac{3}{|x|^3} \right) \|h_n\| \\ &\leq \min \left( \frac{1}{2c_1}, \frac{3}{|x|^3} \right) \left( \|h^{(n)}\| + \frac{n(n-1)(n-2)}{6c_1} \sum_{j=0}^{n-3} a_{n-3-j}^{(n-3)} \|h^{(n-3-j)}\| \right. \\ &\quad \left. + n(n-1) \frac{3}{(6c_1)^{2/3}} \sum_{j=0}^{n-2} a_{n-2-j}^{(n-2)} \|h^{(n-2-j)}\| + \frac{3n}{(6c_1)^{1/3}} \sum_{j=0}^{n-1} a_{n-1-j}^{(n-1)} \|h^{(n-1-j)}\| \right). \end{aligned}$$

Now, collecting terms and using the induction hypothesis on (3.41) we obtain

$$|f^{(n)}(x)| \leq \min \left( \frac{1}{2c_1}, \frac{3}{|x|^3} \right) \left( \|h^{(n)}\| + \frac{3n}{(6c_1)^{1/3}} \|h^{(n-1)}\| \right)$$

$$\begin{aligned}
 & + \left( \frac{3n}{(6c_1)^{1/3}} a_{n-2}^{(n-1)} + \frac{3n(n-1)}{(6c_1)^{2/3}} a_{n-2}^{(n-2)} \right) \|h^{(n-2)}\| \\
 & + \sum_{k=3}^n \left( \frac{3n}{(6c_1)^{1/3}} a_{n-k}^{(n-1)} + \frac{3n(n-1)}{(6c_1)^{2/3}} a_{n-k}^{(n-2)} + \frac{n(n-1)(n-2)}{6c_1} a_{n-k}^{(n-3)} \right) \|h^{(n-k)}\| \\
 = & \min \left( \frac{1}{2c_1}, \frac{3}{|x|^3} \right) \left( \|h^{(n)}\| + \frac{3n}{(6c_1)^{1/3}} \|h^{(n-1)}\| + \frac{12n(n-1)}{(6c_1)^{2/3}} \|h^{(n-2)}\| \right. \\
 & \left. + \sum_{k=3}^n \left( \frac{3n}{(6c_1)^{1/3}} a_{n-k}^{(n-1)} + \frac{3n(n-1)}{(6c_1)^{2/3}} a_{n-k}^{(n-2)} + \frac{n(n-1)(n-2)}{6c_1} a_{n-k}^{(n-3)} \right) \|h^{(n-k)}\| \right).
 \end{aligned}$$

Note that, in passing, we have also proved (3.43) and, now, (3.36) yields (3.44).

### 3.2.6 Multivariate normal distribution

The iterative method, which has been presented mostly for linear ODEs in this paper, applies equally to linear PDEs. Indeed, we can obtain bounds for  $n$ -th order partial derivatives of the solution of the multivariate normal Stein equation if we are given just the bound (3.47). For simplicity, we consider the standard multivariate normal distribution with covariance matrix  $\Sigma = I_d$ , the  $d \times d$  identity matrix. Recall that, for  $k \geq 1$ ,

$$L \left( \frac{\partial^k f(x)}{\prod_{j=1}^k \partial x_{i_j}} \right) = \frac{\partial^k h(x)}{\prod_{j=1}^k \partial x_{i_j}} + k \frac{\partial^k f(x)}{\prod_{j=1}^k \partial x_{i_j}}.$$

Then applying the iterative technique with the bound (3.47) gives

$$\begin{aligned}
 \left\| \frac{\partial^2 f}{\partial x_i \partial x_j} \right\| & \leq \sqrt{\frac{\pi}{2}} \min_{l \in \{i,j\}} \left\| \frac{\partial h(x)}{\partial x_l} + \frac{\partial h(x)}{\partial x_l} \right\| \leq \sqrt{\frac{\pi}{2}} \min_{l \in \{i,j\}} \left\{ \left\| \frac{\partial h}{\partial x_l} \right\| + \left\| \frac{\partial f}{\partial x_l} \right\| \right\} \\
 & \leq \sqrt{\frac{\pi}{2}} \min_{l \in \{i,j\}} \left\| \frac{\partial h}{\partial x_l} \right\| + \frac{\pi}{2} \|\tilde{h}\|.
 \end{aligned}$$

Repeating this procedure allows us to bound partial derivatives of  $f$  of any order. However, the resulting bound would be much worse than the best bounds from the literature.

## 4 Connection between Stein equations and Poisson equations

If  $L$  in equation (1.2) takes the form of a second order (possibly degenerate) elliptic differential operator, then the Stein equation coincides with the Poisson equation

$$Lf = h - \mu(h). \tag{4.1}$$

There is a large literature on existence and regularity results for elliptic differential equations, see [18] and [29] as well as references thereof and therein. Most of these results focus on bounded domains  $D \subsetneq \mathbb{R}^d$ . However, [44] establishes regularity results for  $D = \mathbb{R}^d$ . We concentrate on presenting the formal relation before summarising and exploiting the results of [44] in Section 4.1.

These regularity results are achieved by studying the SDE

$$dX_t = b(X_t) dt + \sigma(X_t) dB_t, \tag{4.2}$$

where  $B_t$  is  $d_1$ -dimensional Brownian motion, that has

$$L\phi(x) = \sum_{i=1}^d b_i(x) \frac{\partial \phi}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2 \phi}{\partial x_i \partial x_j}, \quad a(x) = \sigma \sigma^*(x)$$

as its infinitesimal generator. In the Stein equation literature this is known as generator approach which was introduced by Barbour and Götze [3], [32], but instead of looking at a case by case basis we establish a general connection that has to best of our knowledge not been exploited in the literature. However, this connection has been hinted at in Section 6.2 of [40] without making the connection with regularity results of [44]. The iteration of the regularity results of [44] similar to results presented below has been used in [55] to analyse MCMC algorithms with subsampling. Recently, in [33] and in [4], the generator approach of Stein’s method for certain classes of diffusion processes has been used for applications related to queuing theory. In fact, [33] seems to re-develop this technique, being unaware of Barbour’s and Götze’s previous insights.

Stein’s method can be viewed as a quantitative version of the Echeverría-Weiss theorem (more details about this theorem can be found in Sections 4.1-4.2 in [36] and in [17]).

**Theorem 4.1.** (Echeverría-Weiss theorem) *Suppose there is a probability density  $\mu$  on  $\mathbb{R}^d$  such that, for all  $\varphi \in C_c^2$ ,*

$$\int_{\mathbb{R}^d} L\varphi(x)\mu(x) dx = 0.$$

*Then  $\mu$  is an invariant distribution for the SDE in equation (4.2).*

This theorem illustrates the relation between the invariant/stationary distribution of the associated stochastic process and the Stein differential operator. Moreover, sufficiently fast convergence of the stochastic process to its invariant distribution guarantees the existence of a solution to the Poisson equation (4.1).

#### 4.1 Solution of the Poisson equation

Theorem 1 and 2 in [44] characterise the smoothness and growth of the solution to the Poisson equation associated with equation (4.2). We now state a simplified version of these theorems, which we make use of in Section 4.2. Before doing so, we state some assumptions. We suppose that  $b$  is a locally bounded Borel vector function of dimension  $d_1 \geq d$ , that  $\sigma$  is a  $d \times d_1$  matrix-valued uniformly continuous function, and that  $\sigma\sigma^*$  is bounded and positive-definite. We also suppose that

$$\langle b(x), x/|x| \rangle \leq -r|x|, \quad |x| \geq M_0, \tag{4.3}$$

with  $M_0 \geq 0$  and  $r > 0$ . This condition prevents the solution of the SDE (4.2) from blowing up, so that the process  $\{X_t\}$  is well-defined for all  $t > 0$ .

**Theorem 4.2.** [44] *Suppose that the above assumptions hold, and that there also exist  $C > 0$  and  $\beta \geq 0$  such that*

$$|h(x) - \mu(h)| \leq C(1 + |x|)^\beta. \tag{4.4}$$

*Then there exists a solution  $f$ , belonging to the Sobolev class  $W_{p,\text{loc}}^2$  for any  $p > 1$ , to the Poisson equation (4.1).*

1. *If  $\beta < 0$ , then  $f$  is bounded. Moreover,*

$$\sup_x |f(x)| \leq C'(\beta) \sup_x |(h(x) - \mu(h))(1 + |x|)^{-\beta}|,$$

*and there exists a constant  $C''(\beta)$  such that*

$$|\nabla f(x)| \leq C''(\beta)(1 + |x|^\beta).$$

2. Suppose now that  $\beta > 0$ . Then there exists  $C'(\beta)$  such that

$$|f(x)| \leq C'(\beta) (1 + |x|)^\beta.$$

Finally, there exists  $C''(\beta)$  such that

$$|\nabla f(x)| \leq C''(\beta) (1 + |x|)^\beta.$$

In both cases  $C'(\beta), C''(\beta) \rightarrow \infty$  as  $|\beta| \rightarrow 0$ .

Recall that  $P_t f(x) = \mathbb{E}_x f(X_t)$  is the semigroup associated with  $L$ , that can also be expressed by  $\int f(y) \mu_t^x(dy)$  using the distribution  $\mu_t^x$  of  $X_t$  started at  $x$ . On that basis we see that

$$f(x) = - \int_0^\infty (P_t(h - \mu(h)))(x) dt \tag{4.5}$$

is a formal solution to Poisson equation (4.1) by calculating

$$\begin{aligned} Lf(x) &= - \int_0^\infty L(P_t(h - \mu(h)))(x) dt = - \int_0^\infty \frac{d}{dt} (P_t(h - \mu(h)))(x) dt \\ &= h(x) - \mu(h) - \lim_{t \rightarrow \infty} (P_t(h - \mu(h)))(x), \end{aligned}$$

where we use that the generator satisfies  $LP_t = \frac{d}{dt} P_t$ . Thus, the existence of  $f$  can be shown by establishing that  $\lim_{s \rightarrow \infty} - \int_0^s (P_t(h - \mu(h)))(x) dt$  exists in an appropriate functions space and solves the Poisson equation (4.1) in a weak sense.

The regularity results comes from classical regularity results for elliptic partial differential equation, more precisely from the following inequality

$$|\nabla f(x)| \leq C \left( |f|_{L^\infty(B_{x,2})} + |Lf|_{L^\infty(B_{x,2})} \right), \tag{4.6}$$

where  $B_{x,r}$  is a ball of radius  $r$  around  $x$ ; see p. 1070 of [44]. The second term on the RHS is just the RHS of the Poisson equation which is bounded by assumption. The problem lies in the term  $|f|_{L^\infty(B_{x,2})}$ . It is known that for typical classes of test functions  $h$ , like Lipschitz functions, the solution  $f$  is not uniformly bounded on  $\mathbb{R}^d$ . However, such uniform bounds on  $f$  would be necessary in order to apply the iterative procedure. Nevertheless, in the next section we obtain bounds of the form

$$|\partial^\alpha f| \leq C \left( 1 + |x|^{\phi(|\alpha|)} + \sum_{i=0}^{|\alpha|} |\partial^{\beta_i} (h - \mu(h))| \right),$$

where we used multi-index notation for derivatives and  $|\beta_i| = i$  and  $\beta_0 < \dots < \beta_{|\alpha|}$ . This discussion highlights that a bound on  $|\nabla f|$  involving  $|f|_{L^\infty(B_{x,2})}$  is too weak to obtain results like the classical plug-in theorems which are familiar from Stein's method for univariate distributions. However, even non-uniform bounds on  $f$  with a polynomial order of growth can sometimes be of use, if one has control of certain moments of the random variable under consideration. This has been exemplified recently in [33], [4] and in [22].

**Remark 4.3.** *Even though the use of (4.6) is fundamentally flawed, we describe how bounds on  $|f|_{L^\infty(B_{x,2})}$  are obtained in [44].*

**1. a priori using a direct calculation following page 1068 of [44]**

Using equation (4.5), bounds can be obtained as follows

$$|f(x)| \leq \int_0^\infty |P_t(h - \mu(h))(x)| dt$$

$$\begin{aligned}
 &= \int_0^\infty \|h\| \left| \int \frac{h(y)}{\|h\|} (\mu_t^x - \mu)(dy) \right| dt \\
 &\leq \|h\| \int_0^\infty |\mu_t^x - \mu|_{TV} dt.
 \end{aligned}$$

For most processes of interest the convergence rate of  $|\mu_t^x - \mu|_{TV}$  depends on  $x$ . In case of the Langevin equation, we study below, the bounds would be of the form

$$|P_t - \mu|_{TV} \leq CV(x) \exp(-\lambda t);$$

see Theorem 6.1 of [42]. Typically,  $1 \leq V(x)$  grows at least linearly in  $|x|$ . This is not a deficiency of the method because similar bounds are strict for the Ornstein-Uhlenbeck process. While the convergence to equilibrium thus plays a role in bounding  $|f|$  it is not so clear how the convergence relates to bounds on derivatives of  $f$ .

**2. a posteriori using a representation formula**

The result in Theorem 4.2 is proved using the representation formula

$$f(x) = \mathbb{E}_x f(X_{\tau^R}) + \mathbb{E}_x \int_0^{\tau^R} (h(X_t) - \mu(h)) dt$$

in combination with Itô's formula and (4.4). For more details see p. 1070 of [44].

**4.2 The over-damped Langevin SDE**

We consider the Langevin SDE in  $\mathbb{R}^d$ ,

$$dX_t = \frac{1}{2} \nabla \log \mu(X_t) dt + dW_t,$$

we see that  $\mu$  is the invariant distribution by verifying that it solves the corresponding Fokker-Planck equation ( $L^* \mu = 0$ , where  $L^*$  is the adjoint operator to  $L$ ). Matching the coefficients to equation (4.2), we see that

$$b(x) = \frac{1}{2} \nabla \log \mu(x) \quad \text{and} \quad \sigma = I.$$

In order for Stein's method to be applicable, we need bounds on higher derivatives as well. For this reason, we will iterate Theorem 4.2. In order to do so, we note that the derivatives of  $f$  can be expressed as solutions to Poisson equations with different RHSs:

$$Lf = h - \mu(h), \tag{4.7}$$

$$L\partial_i f = \partial_i h - \frac{1}{2} \nabla f \cdot \partial_i b, \tag{4.8}$$

$$L\partial_{ij} f = \partial_{ij} h - \frac{1}{2} \nabla \partial_j f \cdot \partial_i b - \frac{1}{2} \nabla f \cdot \partial_{ij} b - \frac{1}{2} \nabla f \cdot \partial_j b, \tag{4.9}$$

$$\begin{aligned}
 L\partial_{ijk} f &= \partial_{ijk} h - \frac{1}{2} \nabla \partial_{jk} f \cdot \partial_i b - \frac{1}{2} \nabla \partial_j f \cdot \partial_{ik} b - \frac{1}{2} \nabla \partial_k f \cdot \partial_{ij} b \\
 &\quad - \frac{1}{2} \nabla f \cdot \partial_{ijk} b - \frac{1}{2} \nabla \partial_k f \cdot \partial_{jkb} - \frac{1}{2} \nabla f \cdot \partial_{jkb} - \frac{1}{2} \nabla f \cdot \partial_{kb}.
 \end{aligned} \tag{4.10}$$

We will denote by  $\beta_{f,i}$  numbers that satisfy

$$\sup_{|\alpha|=i} |\partial^\alpha f(x)| \lesssim (1 + |x|^{\beta_{f,i}}), \tag{4.11}$$

where we used multi-index notation for derivatives. We use a similar notation for the derivatives of  $f$ , that is  $\beta_{f,i}$  and assume that these bounds are a priori given.

Using Theorem 4.2 we can obtain  $\beta_{f,i}$  to satisfy (4.11) in terms of the  $\beta$ 's, which we formulate as the following proposition.

**Proposition 4.4.** *Suppose that  $h$  and its derivatives up to third order are bounded, and that assumption (4.3) holds. Then, for any  $\epsilon > 0$ ,*

$$\beta_{f,0} = \epsilon, \quad \beta_{f,1} = \epsilon, \quad \beta_{f,2} = \epsilon \vee \frac{\beta_{b,1}}{2}, \quad \beta_{f,3} = \epsilon \vee \beta_{b,1} \vee \frac{\beta_{b,2}}{2}$$

and  $\beta_{f,4} = \epsilon \vee \frac{1}{2} (3\beta_{b,1} \vee (\beta_{b,1} + \beta_{b,2}) \vee \beta_{b,3})$

satisfy equation (4.11).

*Proof.* Applying Theorem 4.2 to equation (4.7) implies that  $\beta_{f,0} := \beta_{f,1} := \epsilon$  satisfies equation (4.11). Applying Theorem 4.2 to equation (4.8) yields that  $\beta_{f,2} \leq \epsilon \vee \beta_{b,1}$ , and then applying the theorem to equation (4.9) gives  $\beta_{f,3} \leq 2\beta_{b,1} \vee \beta_{b,2}$ . Finally, applying Theorem 4.2 to equation (4.10) yields that

$$\begin{aligned} \beta_{f,4} &\leq (\beta_{b,1} + (2\beta_{b,1} \vee \beta_{b,2})) \vee (\beta_{b,1} + \beta_{b,2}) \vee \beta_{b,3} \\ &\leq 3\beta_{b,1} \vee (\beta_{b,1} + \beta_{b,2}) \vee \beta_{b,3}, \end{aligned}$$

completing the proof. □

**Acknowledgments.** RG is supported by a Dame Kathleen Ollerenshaw Fellowship and acknowledges support from EPSRC grant EP/K032402/1. SV's research is partially supported by the grants EP/K009850/1 and EP/N000188/1. We would like to thank the referees for a very thorough reading of our paper, and for their helpful suggestions.

## References

- [1] Arfken, G. *Mathematical Methods for Physicists*, 3rd ed. Orlando, FL: Academic Press, 1985. MR-0205512
- [2] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. Stein's method on the second Wiener chaos: 2-Wasserstein distance. arXiv:1601:03301, 2016.
- [3] Barbour, A. D. Stein's method for diffusion approximations. *Probab. Theory Rel.* **84** (1990), pp. 297–322. MR-1035659
- [4] Braverman, A. and Dai, J. G. Stein's method for steady-state diffusion approximations of  $M/Ph/n + M$  systems. *Ann. Appl. Probab.* **27** (2016), pp. 550–581. MR-3619795
- [5] Chatterjee, S., Fulman, J. and Röllin, A. Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** (2011), pp. 197–223. MR-2802856
- [6] Chatterjee, S. and Shao, Q.-M. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.* **21** (2011), pp. 464–483. MR-2807964
- [7] Chen, L. H. Y. Poisson approximation for dependent trials. *Ann. Probab.* **3** (1975), pp. 534–545. MR-0428387
- [8] Daly, F. Upper bounds for Stein-type operators. *Electon. J. Probab.* **13** (2008), pp. 566–587. MR-2399291
- [9] Diaconis, P. and Zabell, S. Closed Form Summation for Classical Distributions: Variations on a Theme of De Moivre. *Statist. Sci.* **6** (1991), pp. 284–302. MR-1144242
- [10] Döbler, C. A rate of convergence for the arcsine law by Stein's method. arXiv:1207.2401, 2012.
- [11] Döbler, C. Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.* no. 109 (2015), pp. 1–34. MR-3418541
- [12] Döbler, C. Stein's method for the half-normal distribution with applications to limit theorems related to the simple symmetric random walk. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** (2015), pp. 171–191. MR-3343481

- [13] Eichelsbacher, P. and Löwe, M. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **15** (2010), pp. 962–988. MR-2659754
- [14] Eichelsbacher, P. and Martschink, B. Rates of Convergence in the Blume-Emery-Griffiths Model. *J. Stat. Phys.* **154** (2014), pp. 1483–1507. MR-3176416
- [15] Eichelsbacher, P. and Thäle, C. Malliavin-Stein method for Variance-Gamma approximation on Wiener space. *Electron. J. Probab.* **20** no. 123 (2015), pp. 1–28. MR-3425543
- [16] Ellis, R. S. and Newman, C. M. Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete* **44** (1978), pp. 117–139. MR-0503333
- [17] Ethier, S. N. and Kurtz, T. G. *Markov Processes – Characterisation and Convergence*. Wiley, 1986. MR-0838085
- [18] Evans, L. C. *Partial differential equations*. American Mathematical Society, Providence, RI, second edition, 2010. MR-2597943
- [19] Fulman, J. and Ross, N. Exponential approximation and Stein's method of exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** (2013), pp. 1–13. MR-3083916
- [20] Gaunt, R. E. *Rates of Convergence of Variance-Gamma Approximations via Stein's Method*. DPhil thesis, University of Oxford, 2013. MR-3194737
- [21] Gaunt, R. E. Variance-Gamma approximation via Stein's method. *Electron. J. Probab.* **19** no. 38 (2014), pp. 1–33. MR-3194737
- [22] Gaunt, R. E. Stein's method for functions of multivariate normal random variables. arXiv:1507.08688, 2015.
- [23] Gaunt, R. E. Rates of Convergence in Normal Approximation Under Moment Conditions Via New Bounds on Solutions of the Stein Equation. *J. Theor. Probab.* **29** (2016), pp. 231–247. MR-3463084
- [24] Gaunt, R. E. Uniform bounds for expressions involving modified Bessel functions. *Math. Inequal. Appl.* **19** (2016), pp. 1003–1012. MR-3535219
- [25] Gaunt, R. E. On Stein's method for products of normal random variables and zero bias couplings. *Bernoulli* **23** (2017), pp. 3311–3345. MR-3654808
- [26] Gaunt, R. E. Products of normal, beta and gamma random variables: Stein operators and distributional theory. To appear in *Braz. J. Probab. Stat.*, 2017+.
- [27] Gaunt, R. E., Mijoule, G. and Swan Y. Stein operators for product distributions, with applications. arXiv:1604:06819, 2016.
- [28] Gaunt, R. E., Pickett, A. M. and Reinert, G. Chi-square approximation by Stein's method with application to Pearson's statistic. *Ann. Appl. Probab.* **27** (2017), pp. 720–756. MR-3655852
- [29] Gilbarg, D. and Trudinger, N. S. *Elliptic partial differential equations of second order*. Springer-Verlag, Berlin, 2001. MR-1814364
- [30] Goldstein, L. and Reinert, G. Stein's method for the Beta distribution and the Pólya-Eggenberger Urn. *J. Appl. Probab.* **50** (2013), pp. 1187–1205. MR-3161381
- [31] Goldstein, L. and Rinott, Y. Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** (1996), pp. 1–17. MR-1371949
- [32] Götze, F. On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19** (1991), pp. 724–739. MR-1106283
- [33] Gurvich, I. Diffusion models and steady-state approximations for exponentially ergodic Markovian queues. *Ann. Appl. Probab.* **24** (2014), pp. 2527–2559. MR-3262510
- [34] Komorowski, T. and Walczuk, A. Central Limit Theorem for Markov Processes with Spectral Gap in the Wasserstein Metric. *Stoch. Proc. Appl.* **122** (2012), pp. 2155–2184. MR-2921976
- [35] Koudou, A. E. and Ley, C. Characterizations of GIG laws: a survey complemented with two new results. *Probab. Surv.* **11** (2014), pp. 161–176. MR-3264557
- [36] Lamberton, D. and Pages, G. Recursive computation of the invariant distribution of a diffusion. *Bernoulli* **8** (2002), pp. 367–405. MR-1913112
- [37] Lefèvre, C. and Utev, S. Exact norms of a Stein-type operator and associated stochastic orderings. *Probab. Theory Rel.* **127** (2003), pp. 353–366. MR-2018920

- [38] Ley, C., Reinert, G. and Swan, Y. Stein’s method for comparison of univariate distributions. *Probab. Surv.* **14** (2017), pp. 1–52. MR-3595350
- [39] Luk, H. *Stein’s Method for the Gamma Distribution and Related Statistical Applications*. PhD thesis, University of Southern California, 1994. MR-2693204
- [40] Mattingly, J. C., Stuart, A. M. and Tretyakov, M. V. Convergence of Numerical Time-Averaging and Stationary Measures via Poisson Equations. *SIAM J. Numer. Anal.* **48** (2010), pp. 552–577. MR-2669996
- [41] Meckes, E. On Stein’s method for multivariate normal approximation. *IMS Collect.* **5** (2009), pp. 153–178. MR-2797946
- [42] Meyn, S. P. and Tweedie, R. L. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.* **25** (1993), pp. 518–548. MR-1234295
- [43] Olver, F. W. J., Lozier, D. W., Boisvert, R. F. and Clark, C. W. *NIST Handbook of Mathematical Functions*. Cambridge University Press, 2010. MR-2655349
- [44] Pardoux, E. and Veretennikov, A. Y. On the Poisson Equation and Diffusion Approximation. 1. *Ann. Probab.* **29** (2001), pp. 1061–1085. MR-1872736
- [45] Peköz, E. and Röllin, A. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** (2011), pp. 587–608. MR-2789507
- [46] Peköz, E., Röllin, A. and Ross, N. Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23** (2013), pp. 1188–1218. MR-3076682
- [47] Pickett, A. *Rates of Convergence of  $\chi^2$  Approximations via Stein’s Method*. DPhil thesis, University of Oxford, 2004.
- [48] Pike, J. and Ren, H. Stein’s method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** (2014), pp. 571–587. MR-3283586
- [49] Reinert, G. Three general approaches to Stein’s method. In *An Introduction to Stein’s Method. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* Barbour, A. D. and Chen, L. H. Y. eds. Singapore Univ. Press, Singapore. **4** (2005), pp. 183–221. MR-2235451
- [50] Schoutens, W. Orthogonal Polynomials in Steins Method. EURANDOM Report 99-041, EURANDOM, 1999.
- [51] Schoutens, W. Orthogonal polynomials in Stein’s method. *J. Math. Anal. Appl.* **253** (2001), pp. 515–531. MR-1808151
- [52] Stein, C. A bound for the error in the normal approximation to the the distribution of a sum of dependent random variables. In *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* (1972), vol. 2, Univ. California Press, Berkeley, pp. 583–602. MR-0402873
- [53] Stein, C. *Approximate Computation of Expectations*. IMS, Hayward, California, 1986. MR-0882007
- [54] Stein, C., Diaconis, P., Holmes, S. and Reinert, G. Use of exchangeable pairs in the analysis of simulations. In *Stein’s method: expository lectures and applications*, volume 46 of IMS Lecture Notes Monogr. Ser., Inst. Math. Statist., Beachwood, OH (2004), pp. 1–26. MR-2118600
- [55] Teh, Y. W., Vollmer, S. J. and Zygalakis K. C. (Non-) asymptotic properties of stochastic gradient langevin dynamics. arXiv:1501.00438, 2015.

## A Further calculations

For the sake of brevity, the following formulas were given in Section 3 without proof. We provide their simple proofs here.

### A.1 Normal distribution

We now prove the following inequality that is found in Section 3.2:

$$\|f^{(n)}\| \leq \sum_{j=0}^n \left(\frac{\pi}{2}\right)^{(n-j+1)/2} \frac{n!}{j!} \|\tilde{h}^{(j)}\|.$$

*Proof.* Apply Lemma 2.3 with  $a_k = k + 1$  and  $C_k = \sqrt{\frac{\pi}{2}}$  to get

$$\begin{aligned} \|f^{(n)}\| &\leq \sqrt{\frac{\pi}{2}} \sum_{j=0}^n \left( \prod_{i=j}^{n-1} \sqrt{\frac{\pi}{2}}(i+1) \right) \|\tilde{h}^{(j)}\| \\ &= \sum_{j=0}^n \left(\frac{\pi}{2}\right)^{(n-1+j)/2} \left( \prod_{i=j+1}^n i \right) \|\tilde{h}^{(j)}\| \\ &= \sum_{j=0}^n \left(\frac{\pi}{2}\right)^{(n-j+1)/2} \frac{n!}{j!} \|\tilde{h}^{(j)}\|. \quad \square \end{aligned}$$

**A.2 Beta distribution**

$$L_{\alpha,\beta}f(x) = x(1-x)f'(x) + (\alpha - (\alpha + \beta)x)f(x) = \tilde{h}(x) \tag{A.1}$$

and, for  $k \geq 1$ ,

$$L_{\alpha+k,\beta+k}f^{(k)}(x) = h^{(k)}(x) + k(\alpha + \beta)f^{(k-1)}(x). \tag{A.2}$$

*Proof.* We prove the result by induction. Differentiating both sides of (A.2) gives

$$\begin{aligned} x(1-x)f^{(k+2)}(x) + (\alpha + k - (\alpha + \beta + 2k)x)f^{(k+1)}(x) + (1-2x)f^{(k+1)}(x) \\ - (\alpha + \beta + 2k)f^{(k)}(x) = h^{(k+1)}(x) + k(\alpha + \beta)f^{(k)}(x). \end{aligned}$$

Rearranging gives

$$\begin{aligned} x(1-x)f^{(k+2)}(x) + (\alpha + k + 1 - (\alpha\beta + 2k + 2)x)f^{(k+1)}(x) \\ = h^{(k)}(x) + [(\alpha + \beta + 2k) + k(\alpha + \beta + k - 1)]f^{(k-1)}(x), \end{aligned}$$

and so

$$\begin{aligned} x(1-x)f^{(k+2)}(x) + (\alpha + k + 1 - (\alpha\beta + 2k + 2)x)f^{(k+1)}(x) \\ = h^{(k)}(x) + (k + 1)(\alpha + \beta + k)f^{(k-1)}(x). \end{aligned}$$

We recognise the left-hand side as  $L_{\alpha+k+1,\beta+k+1}f^{(k+1)}(x)$ , and so the result has been proved by induction on  $k$ . □

**A.3 Student's  $t$  distribution**

$$L_{d,\delta}f(x) = (\delta^2 + x^2)f'(x) - (d-1)xf(x) = \tilde{h}(x) \tag{A.3}$$

and, for  $k \geq 1$ ,

$$L_{d-2k,\delta}f^{(k)}(x) = h^{(k)}(x) + k(d-k)f^{(k-1)}(x). \tag{A.4}$$

*Proof.* Differentiating both sides of (A.3) gives

$$\begin{aligned} (\delta^2 + x^2)f^{(k+2)}(x) - (d-2k-1)xf^{(k+1)}(x) + 2xf^{(k+1)}(x) - (d-2k-1)f^{(k)}(x) \\ = h^{(k+1)}(x) + k(d-k)f^{(k)}(x), \end{aligned}$$

which on rearranging gives

$$\begin{aligned} L_{d-2(k+1),\delta} = h^{(k+1)}(x) + [k(d-k) + (d-2k-1)]f^{(k)}(x) \\ = h^{(k+1)}(x) + (k+1)(d-k-1)f^{(k)}(x), \end{aligned}$$

as required. □

$$\|f^{(n)}\| \leq \sum_{j=0}^n \left(\frac{\sqrt{\pi}}{2\delta}\right)^{n+1-j} \frac{n!}{j!} (d-n)_{n-j} A_j \|\tilde{h}^{(j)}\|,$$

where  $A_j = \prod_{i=j}^n \frac{\Gamma(\frac{d}{2}-i)}{\Gamma(\frac{d+1}{2}-i)}$ .

*Proof.* Apply Lemma 2.3 with  $a_k = (k+1)(d-k-1)$  and  $C_k = \frac{\sqrt{\pi}\Gamma(\frac{d}{2}-k)}{2\delta\Gamma(\frac{d+1}{2}-k)}$  to get

$$\begin{aligned} \|f^{(n)}\| &\leq C_n \sum_{j=0}^n \left( \sum_{i=j}^{n-1} C_i (i+1)(d-i-1) \right) \|\tilde{h}^{(j)}\| \\ &= \sum_{j=0}^n \left(\frac{\sqrt{\pi}}{2\delta}\right)^{n+1-j} A_j \left( \prod_{i=j+1}^n \right) \|\tilde{h}^{(j)}\| \\ &= \sum_{j=0}^n \left(\frac{\sqrt{\pi}}{2\delta}\right)^{n+1-j} \frac{n!}{j!} (d-n)_{n-j} A_j \|\tilde{h}^{(j)}\|. \end{aligned} \quad \square$$

$$\|f^{(2k+1)}\| \leq \sum_{j=0}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \frac{(2k)!!}{(2j)!!} (d-2k)_{k-j,2} \|\tilde{h}^{(2j)}\|$$

and

$$\begin{aligned} \|f^{(2k)}\| &\leq \sum_{j=1}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \frac{(2k-1)!!}{(2j-1)!!} (d-2k+1)_{k-j,2} \|h^{(2j-1)}\| \\ &\quad + \frac{\sqrt{\pi}\Gamma(\frac{d}{2})}{2\delta\Gamma(\frac{d+1}{2})} \left(\frac{2}{\delta^2}\right)^k (2k-1)!! (d-2k+1)_{k,2} \|\tilde{h}\|. \end{aligned}$$

*Proof.* Apply Lemma 2.3 with  $a_k = (k+1)(d-k-1)$  and  $C_k = \frac{2}{\delta^2}$  to get

$$\begin{aligned} \|f^{(2k+1)}\| &\leq \frac{2}{\delta^2} \sum_{j=0}^k \left( \prod_{i=j}^{k-1} \frac{2}{\delta^2} (2i+2)(d-2i-2) \right) \|\tilde{h}^{(2j)}\| \\ &= \sum_{j=0}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \left( \prod_{i=j+1}^k (2i)(d-2i) \right) \|\tilde{h}^{(2j)}\| \\ &= \sum_{j=0}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \frac{(2k)!!}{(2j)!!} (d-2k)_{k-j,2} \|\tilde{h}^{(2j)}\| \end{aligned}$$

and

$$\begin{aligned} \|f^{(2k)}\| &\leq \frac{2}{\delta^2} \sum_{j=1}^k \left( \prod_{i=j}^{k-1} \frac{2}{\delta^2} (2i+1)(d-2i-1) \right) \|\tilde{h}^{(2j-1)}\| \\ &\quad + \frac{\sqrt{\pi}\Gamma(\frac{d}{2})}{2\delta\Gamma(\frac{d+1}{2})} \|\tilde{h}\| \cdot \prod_{i=1}^k \frac{2}{\delta^2} (2i-1)(d-2i+1) \\ &\leq \sum_{j=1}^k \left(\frac{2}{\delta^2}\right)^{k-j+1} \frac{(2k-1)!!}{(2j-1)!!} (d-2k+1)_{k-j,2} \|h^{(2j-1)}\| \\ &\quad + \frac{\sqrt{\pi}\Gamma(\frac{d}{2})}{2\delta\Gamma(\frac{d+1}{2})} \left(\frac{2}{\delta^2}\right)^k (2k-1)!! (d-2k+1)_{k,2} \|\tilde{h}\|. \end{aligned} \quad \square$$

**A.4 Inverse-gamma distribution**

$$L_{\alpha,\beta}f(x) = x^2 f'(x) + (\beta - (\alpha - 1)x)f(x) = \tilde{h}(x) \tag{A.5}$$

and, for  $k \geq 1$ ,

$$L_{\alpha-2k,\beta}f^{(k)}(x) = h^{(k)}(x) + k(\alpha - k)f^{(k-1)}(x).$$

*Proof.* Differentiating both sides of (A.5) gives

$$\begin{aligned} x^2 f^{(k+2)}(x) + (\beta - (\alpha - 2k - 1)x)f^{(k+1)}(x) + 2xf^{(k+1)}(x) - (\alpha - 2k - 1)f^{(k)}(x) \\ = h^{(k+1)}(x) + k(\alpha - k)f^{(k)}(x), \end{aligned}$$

which on rearranging gives

$$\begin{aligned} L_{\alpha-2(k+1),\beta}f^{(k+1)}(x) &= h^{(k+1)}(x) + [k(\alpha - k) + (\alpha - 2k - 1)]f^{(k)}(x) \\ &= h^{(k+1)}(x) + (k + 1)(\alpha - k - 1)f^{(k)}(x), \end{aligned}$$

as required. □

$$\|f^{(n)}\| \leq \sum_{j=0}^n \frac{n!}{j!} (\alpha - n)_{n-j} B_j \|\tilde{h}^{(j)}\|,$$

where  $B_j = \prod_{i=j}^n \frac{\Gamma(\alpha-2i)}{\beta} \left(\frac{e}{\alpha-2i-1}\right)^{\alpha-2i-1}$ .

*Proof.* Applying Lemma 2.3 with  $a_k = (k+1)(\alpha - k - 1)$  and  $C_k = \frac{\Gamma(\alpha-2k)}{\beta} \left(\frac{\alpha}{\alpha-2k-1}\right)^{\alpha-2k-1}$  gives

$$\begin{aligned} \|f^{(n)}\| &\leq C_k \sum_{j=0}^n \left( \prod_{i=j}^{n-1} C_i (i+1)(\alpha - i - 1) \right) \|\tilde{h}^{(j)}\| \\ &= \sum_{j=0}^n B_j \left( \prod_{i=j+1}^n i(\alpha - i) \right) \|\tilde{h}^{(j)}\| \\ &= \sum_{j=0}^n \frac{n!}{j!} (\alpha - n)_{n-j} B_j \|\tilde{h}^{(j)}\|. \end{aligned} \tag{□}$$

**A.5 PRR distribution**

$$L_s f(x) = s f''(x) - x f'(x) - 2(s - 1)f(x) = \tilde{h}(x) \tag{A.6}$$

and, for  $k \geq 1$ ,

$$L_s f^{(k)}(x) = h^{(k)}(x) + k f^{(k)}(x).$$

*Proof.* Differentiating both sides of (A.6) gives

$$\begin{aligned} L_s f^{(k+1)}(x) &= s f^{(k+3)}(x) - x f^{(k+2)}(x) - 2(s - 1)f^{(k+1)}(x) - f^{(k+1)}(x) \\ &= h^{(k+1)}(x) + k f^{(k+1)}(x), \end{aligned}$$

which on rearranging gives

$$L_s f^{(k+1)}(x) = h^{(k+1)}(x) + (k + 1)f^{(k+1)}(x),$$

as required. □

$$\|f^{(n)}\| \leq (n-1)! \sum_{j=0}^{n-1} \frac{(2\pi)^{(n-j)/2}}{j!} \|h^{(j)}\|.$$

*Proof.* Apply Lemma 2.4 with  $a_k = k + 1$  and  $C_k = \sqrt{2\pi}$  to get

$$\begin{aligned} \|f^{(n)}\| &\leq \sqrt{2\pi} \sum_{j=0}^{n-1} \left( \prod_{i=j}^{n-2} \sqrt{2\pi}(i+1) \right) \|h^{(j)}\| \\ &= (n-1)! \sum_{j=0}^{n-1} \frac{(2\pi)^{(n-j)/2}}{j!} \|h^{(j)}\|. \end{aligned} \quad \square$$

$$\|f^{(2k+1)}\| \leq \sqrt{2\pi} C^k (2k-1)!! \|h\| + \sum_{j=1}^{k-1} C^{k-j+1} \frac{(2k-1)!!}{(2j-1)!!} \|h^{(2j-1)}\|$$

and

$$\|f^{(2k)}\| \leq \sum_{j=0}^{k-1} C^{k-j+1} \frac{(2k)!!}{(2j)!!} \|h^{(2j)}\|,$$

where  $C = 4$  if  $s = 1/2$  and  $C = 2(\sqrt{\pi}s + s^{-1})$  if  $s \geq 1$ .

*Proof.* Apply Lemma 2.4 with  $a_k = k + 1$  and  $D_k = C$  to get

$$\begin{aligned} \|f^{(2k+1)}\| &\leq C \sum_{j=1}^k \left( \prod_{i=j}^{k-1} C(2i+1) \right) \|h^{(2j-1)}\| + \sqrt{2\pi} \|h\| \prod_{i=1}^k C(2i-1) \\ &= \sqrt{2\pi} C^k (2k-1)!! \|h\| + \sum_{j=1}^{k-1} C^{k-j+1} \frac{(2k-1)!!}{(2j-1)!!} \|h^{(2j-1)}\| \end{aligned}$$

and

$$\begin{aligned} \|f^{(2k)}\| &\leq C \sum_{j=1}^{k-1} \left( \prod_{i=j}^{k-2} C(2i+2) \right) \|h^{(2j)}\| \\ &= \sum_{j=0}^{k-1} C^{k-j+1} \frac{(2k)!!}{(2j)!!} \|h^{(2j)}\|. \end{aligned} \quad \square$$

### A.6 Variance-gamma distribution

$$L_{r,\theta,\sigma} f(x) = \sigma^2 x f''(x) + (\sigma^2 r + 2\theta x) f'(x) - (r\theta - x) f(x) = \tilde{h}(x) \tag{A.7}$$

and for  $k \geq 1$ , we have

$$L_{r+k,\theta,\sigma} f^{(k)}(x) = h^{(k)}(x) + k f^{(k-1)}(x) + k\theta f^{(k)}(x).$$

*Proof.* Differentiating both sides of (A.7) gives

$$\begin{aligned} &\sigma^2 x f^{(k+3)}(x) + (\sigma^2(r+k) + 2\theta x) f^{(k+2)}(x) + (r+k)\theta - x f^{(k+1)}(x) \\ &\quad + \sigma^2 f^{(k+2)}(x) + 2\theta f^{(k+1)}(x) - f^{(k)}(x) \\ &= h^{(k+1)}(x) + k f^{(k)}(x) + k\theta f^{(k+1)}(x), \end{aligned}$$

which on rearranging gives

$$L_{r+k+1,\theta,\sigma} f^{(k+1)}(x) = h^{(k+1)}(x) + (k+1) f^{(k)}(x) + (k\theta + \theta) f^{(k+1)}(x),$$

as required. □

$$\|f^{(2k+1)}\| \leq \sum_{j=0}^k \frac{(2k)!!}{(2j)!!} \frac{1}{(r+2j+2)_{k-j+1,2}} \left(\frac{2}{\sigma^2}\right)^{k-j+1} \|\tilde{h}^{(2j)}\|.$$

*Proof.* Apply Lemma 2.3 with  $a_k = k + 1$  and  $D_k = \frac{2}{\sigma^2(r+k)}$  to get

$$\begin{aligned} \|f^{(2k+1)}\| &\leq \frac{2}{\sigma^2(r+k)} \sum_{j=0}^k \left( \prod_{i=j}^{k-1} \frac{2}{\sigma^2(r+2i)} \cdot (2i+2) \right) \|\tilde{h}^{(2j)}\| \\ &= \sum_{j=0}^k \frac{(2k)!!}{(2j)!!} \frac{1}{(r+2j+2)_{k-j+1,2}} \left(\frac{2}{\sigma^2}\right)^{k-j+1} \|\tilde{h}^{(2j)}\|. \quad \square \end{aligned}$$

$$\begin{aligned} \|f^{(2k)}\| &\leq \sum_{j=1}^k \frac{(2k-1)!!}{(2j-1)!!} \frac{1}{(r+2j+1)_{k-j+1,2}} \left(\frac{2}{\sigma^2}\right)^{k-j+1} \|h^{(2j-1)}\| \\ &\quad + \frac{(2k-1)!! 2^k}{(r+1)_{k,2} \sigma^{2k+1}} \left( \frac{1}{r} + \frac{\pi \Gamma(\frac{r}{2})}{2\Gamma(\frac{r+1}{2})} \right) \|\tilde{h}\|. \end{aligned}$$

The proof is very similar to the one for  $\|f^{(2k+1)}\|$ .

### A.7 Density proportional to $e^{-x^4/12}$

$$Lf(x) = f'(x) - \frac{x^3}{3}f(x) = \tilde{h}(x). \tag{A.8}$$

We have

$$Lf'(x) = h'(x) + x^2f(x) =: h_1(x), \tag{A.9}$$

$$Lf''(x) = h''(x) + 2xf(x) + 2x^2f'(x) =: h_2(x), \tag{A.10}$$

and, for  $k \geq 3$ ,

$$\begin{aligned} Lf^{(k)}(x) &= h_k(x) := h^{(k)}(x) + \frac{1}{3}k(k-1)(k-2)f^{(k-3)}(x) + k(k-1)xf^{(k-2)}(x) \\ &\quad + kx^2f^{(k-1)}(x). \end{aligned} \tag{A.11}$$

*Proof.* Differentiating both sides of (A.8) gives

$$f''(x) - \frac{x^3}{3}f'(x) - x^2f(x) = h'(x),$$

which on rearranging yields (A.9). Differentiating (A.9) gives that

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

which on rearranging yields (A.10). We now use induction to prove (A.11) for  $k \geq 3$ . Differentiating both sides of (A.11) gives

$$\begin{aligned} Lf^{(k+1)} &= x^2f^{(k)}(x) + h^{(k+1)}(x) + 2kxf^{(k-1)}(x) + kx^2f^{(k)}(x) + k(k-1)f^{(k-2)}(x) \\ &\quad + k(k-1)xf^{(k-1)}(x) + \frac{1}{3}k(k-1)(k-2)f^{(k-2)} \\ &= h^{(k+1)}(x) + (k+1)x^2f^{(k)}(x) + [2k+k(k-1)]xf^{(k-1)}(x) + [k(k-1) \\ &\quad + \frac{1}{3}k(k-1)(k-2)]f^{(k-2)}(x) \end{aligned}$$

$$\begin{aligned} &= h^{(k+1)}(x) + (k+1)x^2 f^{(k)}(x) + k(k+1)xf^{(k-1)}(x) + [k(k-1) \\ &\quad + \frac{1}{3}(k+1)k(k-1)]f^{(k-2)}(x) = h_{k+1}(x), \end{aligned}$$

as required. □

---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS<sup>1</sup>)
- Easy interface (EJMS<sup>2</sup>)

### Economical model of EJP-ECP

- Non profit, sponsored by IMS<sup>3</sup>, BS<sup>4</sup>, ProjectEuclid<sup>5</sup>
- Purely electronic

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>2</sup>EJMS: Electronic Journal Management System <http://www.vtex.lt/en/ejms.html>

<sup>3</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>4</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>5</sup>Project Euclid: <https://projecteuclid.org/>

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