Applied Probability Trust (24 February 2023)

SUPPLEMENTARY MATERIAL: STEIN'S METHOD AND APPROXIMATING THE MULTIDIMENSIONAL QUANTUM HARMONIC OSCILLATOR

IAN W. MCKEAGUE,* *Columbia University*

YVIK SWAN,** *Université libre de Bruxelles*

1. Preliminary remarks on Stein operators

Consider a probability distribution $\mathbb{P}$ with cdf P and pdf p w.r.t. Lebesgue measure on $\mathbb{R}$. Suppose that p itself is absolutely continuous, with a.e. derivative $p'(x)$. Let $L^1(p)$ be the collection of Lebesgue measurable functions $h \colon \mathbb{R} \to \mathbb{R}$ such that $\int_{-\infty}^{\infty} |h(x)| p(x) \mathrm{d}x < \infty$ and write $P(h) = \mathbb{E}_p h = \int_{-\infty}^{\infty} h(x)p(x)\mathrm{d}x$. We also denote $\mathcal{F}^{(0)}(p)$ the collection of all mean 0 functions under p. Following [2], to p we associate the Stein operators

$$\mathcal{T}_p f(x) = \frac{(f(x)p(x))'}{p(x)} \tag{1}$$

$$\mathcal{L}_p h(x) = \frac{1}{p(x)} \int_{-\infty}^{x} (h(u) - P(h))p(u)\mathrm{d}u \tag{2}$$

with the convention that $\mathcal{T}_p f(x) = \mathcal{L}_p h(x) = 0$ for all x such that $p(x) = 0$. In the sequel we denote $\mathcal{S}(p) = \{x \mid p(x) > 0\}$, $a = \inf \mathcal{S}(p)$ and $b = \sup \mathcal{S}(p)$; we assume that $\mathcal{S}(p)$ is the union of a finite number of intervals.

Of course (1) is only defined for functions f such that fp is absolutely continuous. We denote $\mathcal{F}(p)$ the collection of functions f such that not only is fp absolutely continuous, but also $(f(x)p(x))' \in L^1(p)$ and $\lim_{x\to a} f(x)p(x) = \lim_{x\to b} f(x)p(x) = 0$. This class of functions is important because $\mathcal{T}_p f \in \mathcal{F}^{(0)}(p)$

Received 24 August 2021; revision received 21 August 2022.

* Postal address: Columbia University, New York, U.S.A. Email: im2131@cumc.columbia.edu

** Postal address: Université libre de Bruxelles, Bruxelles, Belgium. Email: yvik.swan@ulb.be

for all $f \in \mathcal{F}(p)$; this is crucial for Stein's method as it gives rise to many "Stein identities" which can be used for a variety of purposes. Similarly, (2) is only defined for functions $h \in L^1(p)$, in which case $\mathcal{L}_p h \in \mathcal{F}(p)$ for all $h \in L^1(p)$.

As described in [5], it is interesting to "standardize" the operator (1) by fixing some $c \in \mathcal{F}(p)$ and considering the family of "standardized Stein operators" $\mathcal{A}_c f(x) = \mathcal{T}_p(cf)(x) = c(x)f'(x) + \Big(c'(x) + p'(x)/p(x)c(x)\Big)f(x)$ acting on some class $\mathcal{F}(\mathcal{A}_c)$ made of all functions for which $cf \in \mathcal{F}(p)$. Note that, by definition, we have $\mathbb{E}[\mathcal{A}_c f(X)] = 0$ for all $f \in \mathcal{F}(\mathcal{A}_c)$. It is important that c be well-chosen to ensure that $\mathcal{A}_c$ has a manageable expression; as is now well known, there are many instances of densities p (even intractable densities) for which this turns out to be possible, leading to many powerful handles on p which can then serve for a variety of purposes including but not limited to distributional approximation.

Given an operator $\mathcal{A}_c$, classical instantiations of Stein's method begin with a "Stein equation", i.e. a differential equation of the form

$$\mathcal{A}_c f(x) = h(x) - P(h) \tag{3}$$

for h some function belonging to a class $\mathcal{H}$ of test functions. Typically, Stein's method practitionners work with one of the following classes: (i) $h \in \mathcal{H} := \mathrm{Kol} = \{\mathbb{I}(-\infty, z], z \in \mathbb{R}\}$ the indicators of a lower half line; (ii) $h \in \mathcal{H} := \mathrm{TV}$ the collection of functions such that $\|h\| \le 1$; (iii) $h \in \mathcal{H} := \mathrm{Wass}$ the collection of Lipschitz functions such that $\|h'\| \le 1$. In the sequel, we restrict our attention to $\mathcal{H} = \mathrm{Wass}$, and we assume that for each $h \in \mathcal{H}$ there exists a unique function $f \in \mathcal{F}(\mathcal{A}_c)$ for which (3) holds for all $x \in \mathcal{S}(p)$. Under "reasonable assumptions on p" (to be verified on a case-by-case basis) we can write $\mathcal{T}_p\mathcal{L}_p h = h - \mathbb{E}_p h$ for all $h \in L^1(p)$ and in particular $\mathcal{T}_p\mathcal{L}_p = \mathrm{Id}$ over $\mathcal{F}^{(0)}(p)$ (Id is the identity function). Similarly $\mathcal{L}_p\mathcal{T}_p = \mathrm{Id}$ over $\mathcal{F}(p)$. In other words, under "reasonable assumptions on p", the solution to (3) is $f_h(x) = \mathcal{L}_p h(x)/c(x)$ at all $x \in \mathcal{S}(p)$ for which $c(x) \neq 0$. Then, since the Wasserstein distance between two probability measures $\mathbb{P}$ and $\mathbb{Q}$ can be written as $d_{\mathrm{W}}(\mathbb{P}, \mathbb{Q}) = \sup_{h \in \mathrm{Wass}} |\mathbb{E}h(X) - \mathbb{E}h(Y)|$

where $X \sim \mathbb{P}$ and $Y \sim \mathbb{Q}$, it holds that

$$d_{\mathrm{W}}(\mathbb{P}, \mathbb{Q}) = \sup_{h \in \mathrm{Wass}} |\mathbb{E}\left[\mathcal{A}_c f_h(Y)\right]|\,; \tag{4}$$

Stein's method in Wasserstein distance consists in exploiting this last identity for the purpose of estimating the Wasserstein distance between the laws $\mathbb{P}$ and $\mathbb{Q}$.

In order to be able to use (4) successfully, it is crucial to control solutions f_h and their derivatives. In [3] the following representations for (2) are provided (recall that h is Lipschitz with a.e. derivative h'):

$$\begin{aligned}
-\mathcal{L}_p h(x) &= -\mathbb{E}\left[(h(X) - P(h))\frac{\mathbb{I}[X \le x]}{p(x)}\right] \\
&= -\mathbb{E}\left[(h(X) - P(h))\frac{\mathbb{I}[X \le x] - P(x)}{p(x)}\right] \\
&= \mathbb{E}\left[(h(X_2) - h(X_1))\frac{\mathbb{I}[X_1 \le x \le X_2]}{p(x)}\right] && (5) \\
&= \mathbb{E}\left[h'(X)\frac{P(x \wedge X)(1 - P(x \vee X))}{p(x)p(X)}\right] && (6)
\end{aligned}$$

where, in (5), the random variables X_1, X_2 are independent copies of X. A simpler way to write (6) is

$$-\mathcal{L}_p h(x) = \int_{-\infty}^{\infty} h'(y)\frac{P_\infty(y \wedge x)\bar{P}_\infty(y \vee x)}{p_\infty(x)}\mathrm{d}y$$

It is also shown that

$$\bar{h}(x) := h(x) - P(h) = \mathbb{E}\left[h'(X)\frac{P(X) - \mathbb{I}[x \le X]}{p(X)}\right]$$

for all $x \in \mathcal{S}(p)$. All these representations will be used in the next section to control the solutions to the Stein equations; this in turn will lead to the distributional approximation results.

2. Stein's method for radial distributions

2.1. Notations and background

Before specializing to radial densities, it is enlightening to first widen the scope somewhat and consider targets F_∞ with density of the form $p_\infty(x) =$

$b(x)\gamma(x)$, for γ some "basis density" and b some positive γ-integrable "tilting" function. This theory may also be of independent interest.

First note that, in order for p_∞ to be a density, it is necessary that $b \geq 0 \in L^1(\gamma)$ and $\mathbb{E}[b(Z)] = 1$, where here and throughout we denote $Z \sim \gamma$. We further impose the following assumptions on p_∞. First, we require that γ is a differentiable probability density function with support the full real line, such that moreover $\gamma' \in L^1(\mathrm{d}x)$ has exactly one sign change (which, for simplicity, we fix at 0) and $\int \gamma'(x)\mathrm{d}x = 0$. Second, we let $B(x)$ be an absolutely continuous nondecreasing function with continuous derivative b, we denote $\mathcal{S}(b) = \{x \in \mathbb{R},$ such that $b(x) > 0\}$ and suppose that $\mathcal{S}(b)$ is the union of a finite number of intervals. Following [5], we also introduce $\mathcal{F}(\gamma)$ the Stein class of γ; this is the class of functions $f : \mathbb{R} \to \mathbb{R}$ such that $(f\gamma)' \in L^1(\mathrm{d}x)$ and $\int_{\mathbb{R}}(f(x)\gamma(x))'\mathrm{d}x = 0$. We assume that $B \in \mathcal{F}(\gamma)$; since $b \in L^1(\gamma)$, this assures us that $\int b\gamma = -\int B\gamma'$ so that integration by parts holds without a remainder term. Finally, letting $F_\infty \sim p_\infty$, we impose that $\mathbb{E}F_\infty (= \mathbb{E}[Zb(Z)]) = 0$.

With these assumptions we are now ready to provide a Stein's method theory for $p_\infty = b\gamma$; the backbone of our approach comes from [7].

Definition 1. (*Generalized (b,γ)-bias transformation.*) Suppose that F is such that $P(F \in \mathcal{S}(b)) = 1$ and define

$$\sigma_B^2(F) = \mathbb{E}\left[-\frac{\gamma'(F)}{\gamma(F)}\frac{B(F)}{b(F)}\right].$$

The random variable $F^\star$ satisfying

$$\sigma_B^2(F)\mathbb{E}\left[\frac{f'(F^\star)}{b(F^\star)}\right] = \mathbb{E}\left[-\frac{\gamma'(F)}{\gamma(F)}\frac{f(F)}{b(F)}\right]$$

for all f such that both integrals exist is said to have the generalized (b,γ)-bias distribution. The random variable $F^\star$ is the generalized (b,γ)-bias transform of F.

By construction, we always have

$$\sigma_B^2(F_\infty) = \mathbb{E}\left[-\frac{\gamma'(F_\infty)}{\gamma(F_\infty)}\frac{B(F_\infty)}{b(F_\infty)}\right] = -\int_{\mathcal{S}(b)} \gamma'(x)B(x)\mathrm{d}x$$
$$= \int_{\mathcal{S}(b)} b(x)\gamma(x)\mathrm{d}x = \mathbb{E}b(Z) = 1.$$

Moreover, for any sufficiently regular function f:

$$\mathbb{E}\left[-\frac{\gamma'(F_\infty)}{\gamma(F_\infty)}\frac{f(F_\infty)}{b(F_\infty)}\right] = \mathbb{E}\left[\frac{f'(F_\infty)}{b(F_\infty)}\right].$$

Therefore $F_\infty = F^\star$, i.e. p_∞ is a fixed point of the generalized (b,γ)-bias transform. More generally, the following holds true.

Lemma 1. *If F is a random variable such that $P(F \in \mathcal{S}(b)) = 1$, $\mathbb{E}\left[\frac{1}{b(F)}\frac{\gamma'(F)}{\gamma(F)}\right] = 0$ and $\sigma_B^2(F) \in (0,\infty)$ then its generalized (b,γ)-bias transform $F^\star$ exists and is absolutely continuous with density*

$$p^\star(x) = -\frac{b(x)}{\sigma_B^2(F)}\mathbb{E}\left[\frac{1}{b(F)}\frac{\gamma'(F)}{\gamma(F)}\mathbb{I}[F \geq x]\right].$$

Moreover F_∞ is the unique fixed point of this tranformation, in the sense that if $F \stackrel{\mathcal{D}}{=} F^\star$ then $F \stackrel{\mathcal{D}}{=} F_\infty$ (equality in distribution).

Proof. All points follow from arguments nearly identical to those in [1, Proposition 2.1]. □

Now consider the function

$$f(x) = \frac{1}{\gamma(x)}\int_{-\infty}^x \big(h(u) - \mathbb{E}h(F_\infty)\big)b(u)\gamma(u)\mathrm{d}u =: \frac{1}{\gamma(x)}\int_{-\infty}^x \bar{h}(u)b(u)\gamma(u)\mathrm{d}u$$

which is solution to the differential equation

$$(\bar{h}(x) :=) h(x) - \mathbb{E}h(F_\infty) = \frac{f'(x) + \frac{\gamma'(x)}{\gamma(x)}f(x)}{b(x)}$$

for all $x \in \mathcal{S}(b)$. Let F be a random variable such that $P(F \in \mathcal{S}(b)) = 1$ and $\sigma_B^2(F) = 1$. We have

$$\mathbb{E}h(F) - \mathbb{E}h(F_\infty) = \mathbb{E}\left[\frac{f'(F)}{b(F)} + \frac{\gamma'(F)}{\gamma(F)}\frac{f(F)}{b(F)}\right] = \mathbb{E}\left[\frac{f'(F)}{b(F)} - \frac{f'(F^\star)}{b(F^\star)}\right] \quad (7)$$

and it remains to express the right hand side of (7) in terms of manageable quantities, such as moments of $F, F^\star$ and $F - F^\star$. We cannot work directly with the function $x \mapsto f'(x)/b(x)$ because the latter is unbounded at $x = 0$. To bypass this difficulty, we introduce the notation

$$\mathcal{L}_\infty h(x) = \frac{1}{p_\infty(x)} \int_{-\infty}^x \bar{h}(u) p_\infty(u) \mathrm{d}u$$

(we stress that $\mathcal{L}_\infty \neq \mathcal{L}_\gamma$) and follow [7] by introducing the function $g = g_{\eta,h}$ given by

$$g(x) = \frac{\mathcal{L}_\infty h(x)}{\mathcal{L}_\infty \eta(x)} = \frac{\int_{-\infty}^x (h(u) - \mathbb{E}h(F_\infty)) b(u) \gamma(u) \mathrm{d}u}{\int_{-\infty}^x (\eta(u) - \mathbb{E}\eta(F_\infty)) b(u) \gamma(u) \mathrm{d}u} \tag{8}$$

at all $x \in \mathcal{S}(b)$ where h is fixed by the left hand side of (7) but η is kept unspecified, to be tuned to our needs at a later stage. Obviously, the above relations are only defined at x such that $p_\infty(x) \neq 0$; we suppose this to be the case here and in the sequel. The function g from (8) is then solution to the Stein equation

$$\left(\mathcal{L}_\infty \eta(x)\right) g'(x) + \bar{\eta}(x) g(x) = h(x) - \mathbb{E}h(F_\infty)$$

at all x inside the support of p_∞. It will be useful to note that the functions g, h and η satisfy the relations

$$\begin{aligned}
(\mathcal{L}_\infty \eta) g &= \mathcal{L}_\infty h \\
(\mathcal{L}_\infty \eta) g' &= \overline{h} - \overline{\eta} g = \overline{h} - \overline{\eta} \frac{\mathcal{L}_\infty h}{\mathcal{L}_\infty \eta} \\
(\mathcal{L}_\infty \eta) g'' &= h' - (\overline{\eta} + (\mathcal{L}_\infty \eta)') g' - \eta' g \\
&= \left(\overline{\eta} \frac{\overline{\eta} + (\mathcal{L}_\infty \eta)'}{\mathcal{L}_\infty \eta} - \eta' \right) g - \left(\overline{h} \frac{\overline{\eta} + (\mathcal{L}_\infty \eta)'}{\mathcal{L}_\infty \eta} - h' \right).
\end{aligned} \tag{9}$$

Straightforward manipulations of the definitions also lead to

$$\mathcal{L}_\infty h(x) = \frac{\mathcal{L}_\gamma (hb)(x)}{b(x)} - \mathbb{E}[h(Z) b(Z)] \frac{\mathcal{L}_\gamma b(x)}{b(x)}.$$

Finally note that g and f are related through $f(x) = \mathcal{L}_\infty \eta(x) g(x) b(x)$, so

$$\begin{aligned}\frac{f'(x)}{b(x)} &= \mathcal{L}_\infty \eta(x) g'(x) + \left(\overline{\eta}(x) - \frac{\gamma'(x)}{\gamma(x)} \mathcal{L}_\infty \eta(x)\right) g(x) \\ &=: \mathcal{L}_\infty \eta(x) g'(x) + \Psi_\infty \eta(x) g(x).\end{aligned}$$

Identity (7) becomes

$$\begin{aligned}\mathbb{E}h(F) - \mathbb{E}h(F_\infty) = \mathbb{E}\Big[&\mathcal{L}_\infty \eta(F) g'(F) - \mathcal{L}_\infty \eta(F^\star) g'(F^\star)\Big] \\ &+ \mathbb{E}\Big[\Psi_\infty \eta(F) g(F) - \Psi_\infty \eta(F^\star) g(F^\star)\Big] \qquad (10)\end{aligned}$$

which is close to what is required. This is however not exactly what we need because, although we shall see that for reasonable choices of η, the function g from (8) and its derivative g' are bounded, the second derivative g'' is often not. In order to cater for this, we introduce some further degrees of liberty in the expressions and rewrite (10) as

$$\begin{aligned}\mathbb{E}h(F) - \mathbb{E}h(F_\infty) &= \mathbb{E}\Big[\Big(r_1(F)\mathcal{L}_\infty \eta(F) - r_1(F^\star)\mathcal{L}_\infty \eta(F^\star)\Big) \frac{g'(F^\star)}{r_1(F^\star)}\Big] \\ &+ \mathbb{E}\Big[r_1(F)\mathcal{L}_\infty \eta(F) \Big(\frac{g'(F)}{r_1(F)} - \frac{g'(F^\star)}{r_1(F^\star)}\Big)\Big] \\ &+ \mathbb{E}\Big[\Big(r_2(F)\Psi_\infty \eta(F) - r_2(F^\star)\Psi_\infty \eta(F^\star)\Big) \frac{g(F^\star)}{r_2(F^\star)}\Big] \\ &+ \mathbb{E}\Big[r_2(F)\Psi_\infty \eta(F) \Big(\frac{g(F)}{r_2(F)} - \frac{g(F^\star)}{r_2(F^\star)}\Big)\Big],\end{aligned}$$

with r_1, r_2 two functions left to be determined. These considerations lead to the main result of the Section.

Proposition 1. *Let the previous notations and assumptions prevail. Then*

$$\begin{aligned}&|\mathbb{E}h(F) - \mathbb{E}h(F_\infty)| \\ &\leq \kappa_1 \mathbb{E}\Big[\Big|r_1(F)\mathcal{L}_\infty \eta(F) - r_1(F^\star)\mathcal{L}_\infty \eta(F^\star)\Big|\Big] + \kappa_2 \mathbb{E}\Big[|r_1(F)\mathcal{L}_\infty \eta(F)||F - F^\star|\Big] \\ &\quad + \kappa_3 \mathbb{E}\Big[\Big|r_2(F)\Psi_\infty \eta(F) - r_2(F^\star)\Psi_\infty \eta(F^\star)\Big|\Big] + \kappa_4 \mathbb{E}\Big[|r_2(F)\Psi_\infty \eta(F)||F - F^\star|\Big] \qquad (11)\end{aligned}$$

where $\kappa_j = \sup_x |\kappa_j(x)|$ *for* $j = 1, \ldots, 4$ *with*

$$\kappa_1(x) = \frac{g'(x)}{r_1(x)}, \kappa_2(x) = \left(\frac{g'(x)}{r_1(x)}\right)', \ \kappa_3(x) = \frac{g(x)}{r_2(x)} \ \textit{and} \ \kappa_4(x) = \left(\frac{g(x)}{r_2(x)}\right)'. \tag{12}$$

Remark 1. The functions r_1, r_2 and η can, for all intents and purposes, be chosen freely. A good choice of function η seems to be $\eta(x) = \eta_\gamma(x) = \gamma'(x)/\gamma(x)$, at least if $\gamma'(x)/\gamma(x)$ is non-increasing on $\mathbb{R}$ and $\mathbb{E}[b'(Z)] = 0$ when $Z \sim \gamma$. Indeed in this case:

$$\mathbb{E}[\eta_\gamma(F_\infty)] = -\mathbb{E}\Big[\frac{b'(F_\infty)}{b(F_\infty)}\Big] = -\mathbb{E}[b'(Z)] = 0 \text{ and } \overline{\eta}_\gamma(x) = \eta_\gamma(x)$$
$$\mathcal{L}_\infty \eta_\gamma(x) = 1 - \frac{\mathcal{L}_\gamma b'(x)}{b(x)} \text{ and } \overline{\eta}_\gamma(x) - \frac{\gamma'(x)}{\gamma(x)}\mathcal{L}_\infty \eta_\gamma(x) = \frac{\gamma'(x)}{\gamma(x)}\frac{\mathcal{L}_\gamma b'(x)}{b(x)}.$$

Another natural choice (which turns out to be equivalent to the previous one when γ is the Gaussian density) is $\eta(x) = -\mathrm{Id}(x) = -x$ for which

$$-\mathbb{E}[\mathrm{Id}(F_\infty)] = -\mathbb{E}\left[Zb(Z)\right] = \mathbb{E}[b'(Z)]$$
$$-\mathcal{L}_\infty \mathrm{Id}(x) = \tau_\infty(x) \quad \text{(the Stein kernel)}.$$

Other choices are possible, depending on the properties of the density γ; it may be worthwhile investigating this avenue, though we will not do it here.

2.2. When the base distribution is standard Gaussian

We now specialize the previous construction to the case that $\gamma(x)$ is the standard Gaussian density. As before, we suppose that b is chosen in such a way that $\mathbb{E}[F_\infty] = 0$; note that if $Z \sim \gamma$ the standard normal then we also have $\mathbb{E}[F_\infty] = \mathbb{E}[Zb(Z)] = -\mathbb{E}[b'(Z)]$. If γ is the Gaussian density then many of the previous expressions simplify, because $\gamma'(x)/\gamma(x) = -\mathrm{Id}(x) := -x$. For instance $\sigma_B^2(F) = \mathbb{E}[FB(F)/b(F)]$ and taking $\eta = -\mathrm{Id}$ we get $\mathcal{L}_\gamma \eta = 1$. Also $\mathcal{L}_\infty \eta = \tau_\infty$ is now the so-called Stein kernel of p_∞; this function is well known to have very good properties for the analysis of p_∞, see e.g. [2] for an overview. At this stage it suffices to remark that $\tau_\infty(x) \geq 0$ for all x. We also have the nice identity $\Psi_\infty \eta(x) = x(\tau_\infty(x) - 1)$ so that (11) becomes

$$\begin{aligned}
&|\mathbb{E}h(F) - \mathbb{E}h(F_\infty)| \\
&\leq \kappa_1 \mathbb{E}\Big[\Big|r_1(F)\tau_\infty(F) - r_1(F^\star)\tau_\infty(F^\star)\Big|\Big] + \kappa_2 \mathbb{E}\Big[|r_1(F)|\tau_\infty(F)|F - F^\star|\Big] \\
&\quad + \kappa_3 \mathbb{E}\Big[\Big|Fr_2(F)(\tau_\infty(F) - 1) - F^\star r_2(F^\star)(\tau_\infty(F^\star) - 1)\Big|\Big] \\
&\quad + \kappa_4 \mathbb{E}\Big[|Fr_2(F)(\tau_\infty(F) - 1)||F - F^\star|\Big] \qquad (13)
\end{aligned}$$

with the coefficients $\kappa_j, j = 1, \ldots, 4$ defined just before (12). The following general result provides bounds on the functions in (12).

Lemma 2. *Let all above notations and assumptions prevail (in particular $\|h'\| \leq 1$). Then*

$$\begin{aligned}
\kappa_1(x) &\leq \frac{2}{|r_1(x)|} \frac{R_\infty(x)}{(\tau_\infty(x))^2} \qquad (14) \\
\kappa_2(x) &\leq \frac{2}{|r_1(x)|\tau_\infty(x)} \left(1 + \left|\frac{2x}{\tau_\infty(x)} - x + \frac{b'(x)}{b(x)} - \frac{r_1'(x)}{r_1(x)}\right| \frac{R_\infty(x)}{\tau_\infty(x)}\right) \\
\kappa_3(x) &\leq \frac{1}{|r_2(x)|} \qquad (15) \\
\kappa_4(x) &\leq \frac{1}{|r_2(x)|} \left(\frac{2R_\infty(x)}{(\tau_\infty(x))^2} + \left|\frac{r_2'(x)}{r_2(x)}\right|\right)
\end{aligned}$$

where $R_\infty(x) = \int_{-\infty}^x P_\infty(u)\mathrm{d}u \int_x^\infty \overline{P}_\infty(u)\mathrm{d}u / p_\infty(x)$.

Proof of Lemma 2. Let g be defined in (8) with $\eta = -\mathrm{Id}$. Suppose that $\mathbb{E}b'(Z) = 0$ and let h be absolutely continuous. We start with the fact that, from (9):

$$\tau_\infty(x)g''(x) = h'(x) - (-x + \tau_\infty'(x))g'(x) + g(x).$$

Using

$$\tau_\infty'(x) = \left(x - \frac{b'(x)}{b(x)}\right)\tau_\infty(x) - x$$

we get

$$g''(x) = \frac{g(x) + h'(x)}{\tau_\infty(x)} + \left(\frac{2x}{\tau_\infty(x)} - x + \frac{b'(x)}{b(x)}\right)g'(x). \qquad (16)$$

With all this we are ready to place bounds on the various coefficients in (13), obtained by placing bounds on the functions defined in (12). It follows immediately from (5) that

$$g(x) = \frac{\mathbb{E}\left[(h(X_1) - h(X_2))\mathbb{I}[X_1 \le x \le X_2]\right]}{\mathbb{E}\left[(X_2 - X_1)\mathbb{I}[X_1 \le x \le X_2]\right]}$$

at all $x \in \mathcal{S}(p_\infty)$, where X_1 and X_2 are independent copies of F_∞. Since, by assumption, $|h(x) - h(y)| \le |x - y|$, (15) follows. To pursue, we use [2, Lemma 2.25] to deduce

$$\begin{aligned} g'(x) &= \frac{\bar{h}(x)\,\mathcal{L}_\infty \eta(x) - \bar{\eta}(x)\,\mathcal{L}_\infty h(x)}{(\mathcal{L}_\infty \eta(x))^2} \\ &= \frac{1}{p_\infty(x)\tau_\infty(x)^2}\left(\mathbb{E}\left[h'(F_\infty)F_\infty\frac{\bar{P}_\infty(F_\infty)}{p_\infty(F_\infty)}\mathbb{I}[x \le F_\infty]\right]\mathbb{E}\left[\frac{P_\infty(F_\infty)}{p_\infty(F_\infty)}\mathbb{I}[F_\infty \le x]\right]\right. \\ &\qquad \left. -\mathbb{E}\left[h'(F_\infty)\frac{P_\infty(F_\infty)}{p_\infty(F_\infty)}\mathbb{I}[F_\infty \le x]\right]\mathbb{E}\left[\frac{\bar{P}_\infty(F_\infty)}{p_\infty(F_\infty)}\mathbb{I}[x \le X]\right]\right) \end{aligned}$$

(where $\bar{P}$ is the survival function of cdf P). The bound on the derivative then follows (see also [2, Equation (2.38)]):

$$\begin{aligned} |g'(x)| &\le \|h'\|\frac{2}{p_\infty(x)\tau_\infty(x)^2}\mathbb{E}\left[\frac{\overline{P}_\infty(F_\infty)}{p_\infty(F_\infty)}\mathbb{I}[x \le F_\infty]\right]\mathbb{E}\left[\frac{P_\infty(F_\infty)}{p_\infty(F_\infty)}\mathbb{I}[F_\infty \le x]\right] \\ &\le 2\frac{1}{\tau_\infty(x)^2}\frac{\int_{-\infty}^x P_\infty(u)\mathrm{d}u \int_x^\infty \overline{P}_\infty(u)\mathrm{d}u}{p_\infty(x)}, \end{aligned}$$

which brings (14). Furthemore, simply by taking derivatives and using the previous bounds, we get

$$\left|\left(\frac{g(x)}{r_2(x)}\right)'\right| \le \frac{1}{|r_2(x)|}\left(\frac{2R_\infty(x)}{\tau_\infty(x)^2} + \left|\frac{r_2'(x)}{r_2(x)}\right|\right)$$

as well as (using (16) to express g'' in terms of the lower order derivatives)

$$\left|\left(\frac{g'(x)}{r_1(x)}\right)'\right| \le \frac{2}{\tau_\infty(x)|r_1(x)|}\left(1 + \left|\frac{2x}{\tau_\infty(x)} - x + \frac{b'(x)}{b(x)} - \frac{r_1'(x)}{r_1(x)}\right|\frac{R_\infty(x)}{\tau_\infty(x)}\right).$$

All claims are therefore established. □

We now apply these results to the choice $B(x) \propto x|x|^k/(k+1)$ and $b(x) \propto |x|^k$ with $k \in \mathbb{N}$. Then $\mathbb{E}F_\infty = 0$ and

$$\sigma_B^2(F) = \mathbb{E}\left[F\frac{B(F)}{b(F)}\right] = \frac{\mathbb{E}F^2}{k+1},$$

so that our first assumptions become $P(F \neq 0) = 1$ and $\mathbb{E}[F^2] = k+1$. The following results then follow from direct manipulations of the definitions. The first result is the same as Lemma 2 in the main text.

Lemma 3. *Let all above notations prevail, and set $\tau_\infty(\cdot; k)$ to be the Stein kernel of $p_\infty(x; , k) = b(x)\gamma(x)$. If $x \neq 0$ then*

$$\tau_\infty(x;k) = 2^{k/2} e^{x^2/2} |x|^{-k} \Gamma(1 + k/2, x^2/2)$$

where $\Gamma(\alpha, x) = \int_x^\infty t^{\alpha-1} e^{-t} \mathrm{d}t$ is the (upper) incomplete gamma function.

Incomplete gamma functions are well understood. For instance, using [4], we readily obtain the next result.

Lemma 4. *The Stein kernel τ_∞ is strictly decreasing on $(0, \infty)$ and satisfies*

$$\lim_{x \to 0} |x|^k \tau_\infty(x;k) = 2^{k/2}\Gamma(1 + k/2) \text{ and } \lim_{|x| \to \infty} \tau_\infty(x;k) = 1$$

as well as the inequalities

$$\begin{aligned} &\frac{1}{|x|^k \tau_\infty(x;k)} \leq \frac{1}{2^{k/2}\Gamma(1+k/2)}, \quad \left| \frac{2}{\tau_\infty(x;k)} - 1 \right| \leq 1 \\ &\text{and } \frac{1}{|x|^{k-1}\tau_\infty(x;k)} \leq \frac{1}{2^{(k-1)/2}\Gamma(1+k/2)} \end{aligned} \tag{17}$$

for all $x \in \mathbb{R}$ and all $k \in \mathbb{N}_0$. Moreover, if we let $P_\infty(\cdot; k)$ and $\overline{P}_\infty(\cdot; k)$ be the cdf and survival function of $p_\infty(\cdot; k)$, and define

$$R_\infty(x;k) = \frac{1}{p_\infty(x;k)} \int_{-\infty}^x P_\infty(u;k) \mathrm{d}u \int_x^\infty \overline{P}_\infty(u;k) \mathrm{d}u$$

as in Lemma 2, then

$$\frac{R_\infty(x;k)}{\tau_\infty(x;k)} \leq \frac{\Gamma(k/2+1)}{\sqrt{2}\Gamma(k/2+1/2)} \tag{18}$$

for all $x \in \mathbb{R}$ and all $k \in \mathbb{N}_0$.

Remark 2. The bounds (17) and (18) are sharp because they are attained at $x \to 0$.

Using Lemmas 3 and 4 we obtain the required bounds on the constants κ_j.

Corollary 1. *Set $r_1(x) = x^k$ and $r_2(x) = 1$ in Lemma 2. Then*

$$\kappa_1(x) \leq \frac{2}{|x|^k} \frac{R_\infty(x;k)}{(\tau_\infty(x;k))^2} \leq \frac{2^{(1-k)/2}}{\Gamma((k+1)/2)}, \tag{19}$$

$$\kappa_2(x) \leq \frac{2}{|x|^k \tau_\infty(x;k)} + \frac{2}{|x|^{k-1}\tau_\infty(x;k)} \left| \frac{2}{\tau_{\infty;k}(x)} - 1 \right| \frac{R_\infty(x;k)}{\tau_\infty(x;k)} \leq 3 \frac{2^{-k/2}}{\Gamma((1+k)/2)}, \tag{20}$$

$$\kappa_3(x) \leq 1, \tag{21}$$

$$\kappa_4(x) \leq \frac{2R_\infty(x;k)}{(\tau_\infty(x;k))^2} \leq 1. \tag{22}$$

Remark 3. The bounds in (19) and (21) are sharp; the other two are not.

Upper bounding (19) by 1 and (20) by 2 (neither of these choices, nor the bound 1 in (22), are optimal in k because the true value goes to 0 as k goes to ∞), inequality (13) leads to the following result.

Theorem 1. *If $F_\infty \sim \mathbb{F}_\infty$ has density $p_\infty(x) \propto |x|^k \varphi(x)$, for a given $k \in \mathbb{N}$ and $F \sim \mathbb{F}$ is some random variable with mean 0 such that $P(F \neq 0) = 1$ and $\mathbb{E}[F^2] = k+1$, then there exists a random variable $F^\star$ which uniquely satisfies*

$$\mathbb{E}\left[\frac{f'(F^\star)}{|F^\star|^k}\right] = \mathbb{E}\left[\frac{f(F)}{|F|^{k-1}}\right]$$

for all f such that both integrals exist, and

$$\begin{aligned} d_{\mathrm{W}}(\mathbb{F}, \mathbb{F}_\infty) \leq\ & \mathbb{E}\left|F^k \tau_\infty(F;k) - (F^\star)^k \tau_\infty(F^\star;k)\right| + 2\mathbb{E}\left[|F|^k \tau_\infty(F;k)|F - F^\star|\right] \\ & + \mathbb{E}\left|F(\tau_\infty(F;k) - 1) - F^\star(\tau_\infty(F^\star;k) - 1)\right| \\ & + \mathbb{E}\left[\left|F(\tau_\infty(F;k) - 1)\right||F - F^\star|\right] \end{aligned} \tag{23}$$

where $\tau_\infty(x;k)$ is the Stein kernel given in equation (**??**) *from the main text.*

Remark 4. The random variable $F^\star$ in the above statement is the (b,γ)-bias transform from Definition 1, here with $b(x) = |x|^k$ and γ the standard Gaussian density. In the sequel we will refer to such $F^\star$ as having the k-radial-bias distribution of F.

2.3. Cases $k = 0$, $k = 1$ and a proof of Proposition ??

The upper bounds in Propositions **??** and **??** from the main text are direct corollaries of the above result. We have already proved the upper bound from Proposition **??** in Example **??** by other means. We therefore concentrate on Proposition **??**.

The following can be shown directly from the definitions:

$$(x\tau_\infty(x,1))' \leq 1, \quad |x(\tau_\infty(x,1)-1)| \leq 1 \text{ and } |\,(x(\tau_\infty(x,1)-1))'\,| \leq 1.$$

Plugging these into (23) gives

$$\begin{aligned} d_{\mathrm{W}}(\mathbb{F},\mathbb{F}_\infty) &\leq 3\mathbb{E}\left[|F-F^\star|\right] + 2\mathbb{E}\left[(|F|(\tau_\infty(F;1)-1)|F-F^\star|\right] + 2\mathbb{E}\left[|F||F-F^\star|\right] \\ &= \mathbb{E}\left[(5+2|F|)\,|F-F^\star|\right] \\ &\leq (5+2x_1)(N-1)^{-1}\sum_{i=1}^{N-1}|x_{i+1}-x_i| \\ &= (5+2x_1)2x_1/(N-1) \\ &= O(\log N/N), \end{aligned}$$

where the second inequality follows from a $k = 1$ version of the coupling argument given in [7, Proof of Corollary 3.7] along with a version of [7, Lemma 4.8] showing that $x_1 = O(\sqrt{\log N})$, as required.

2.4. Cases $k \geq 2$

For $k \geq 2$, we once again call upon [4] to obtain the following lemma.

Lemma 5. (Stein kernel.) *Let all previous notations prevail. Given j, k two integers define*

$$a_j(k) = 2^j\frac{\Gamma(1+k/2)}{\Gamma(1+k/2-j)}$$

with the convention that $a_j(k) = 0$ for all $j \geq k$. Then, for all $x \in \mathbb{R}$ and all $k \in \mathbb{N}$, we have

$$\tau_\infty(x;k) = \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor}\frac{a_j(k)}{x^{2j}} + \frac{a_{\lceil k/2 \rceil}(k)}{\sqrt{2}}\epsilon_k(x)$$

where

$$\epsilon_k(x) = \begin{cases} 0 & \text{if } k = 2\ell \\ e^{x^2/2}|x|^{-(2\ell-1)}\Gamma(1/2, x^2/2) & \text{if } k = 2\ell - 1 \end{cases}$$

Moreover the remainder satisfies

$$0 \le \epsilon_k(x) \le 2^{(k+1)/2} x^{-(k+1)} \frac{\Gamma(1+k/2)}{\Gamma(1/2)}$$

for all $k \in \mathbb{N}$ *and all* x.

Proof. The claim follows from the following representation of the incomplete gamma function (available e.g. from [4, Theorem 3 and Proposition 13]): for all $a > 0$ and all $x > 0$,

$$\Gamma(a, x) = e^{-x} x^{a-1} \sum_{j=0}^{\lfloor a \rfloor - 1} P_j(a) x^{-j} + r(a, x)$$

where $P_j(a) = \Gamma(a)/\Gamma(a-j)$ (and $P_j(a) = 0$ for all $j \ge a$), and, setting $[a] = a - \lfloor a \rfloor$, $r(a, x) = P_{\lfloor a \rfloor}(a)\Gamma([a], x)$ which satisfies

$$0 \le r(a, x) \le e^{-x} P_{\lfloor a \rfloor}(a) x^{[a]-1}.$$

The claim follows.

□

In [7] we considered $b(x) = x^2$. The argument from that paper is now extended to arbitrary non-negative integers k in the next result.

Theorem 2. *Instate all previous notations and let*

$$a_r(k) = \frac{\Gamma(k/2+1)}{\Gamma(k/2-r+1)} 2^r.$$

Then, with the convention that sums over empty sets are defined as 0, the following holds: for even non-negative integers $k = 2\ell$

$$d_W(F, F_\infty) \le \sum_{j=0}^{\ell} a_{\ell-j}(2\ell) \left(\mathbb{E} \left| F^{2j} - (F^\star)^{2j} \right| + 2\mathbb{E} \ \left[|F|^{2j} \left| F - F^\star \right| \right] \right)$$
$$+ \sum_{j=1}^{\ell} a_j(2\ell) \left(\mathbb{E} \left| \frac{1}{F^{2j-1}} - \frac{1}{(F^\star)^{2j-1}} \right| + \mathbb{E} \left[\frac{1}{|F|^{2j-1}} \left| F - F^\star \right| \right] \right)$$

and for odd positive integers $k = 2\ell - 1$

$$\begin{aligned}
d_W(F, F_\infty) \leq &\sum_{j=1}^{\ell} a_{\ell-j}(2\ell-1)\Bigg(\mathbb{E}\left|F^{2j-1} - (F^\star)^{2j-1}\right| + 2\mathbb{E}\ \left[|F|^{2j-1}\,|F - F^\star|\right]\Bigg)\\
&+ \sum_{j=1}^{\ell-1} a_j(2\ell-1)\Bigg(\mathbb{E}\left|\frac{1}{F^{2j-1}} - \frac{1}{(F^\star)^{2j-1}}\right| + \mathbb{E}\left[\frac{1}{|F|^{2j-1}}\Big|F - F^\star\Big|\right]\Bigg)\\
&+ 3a_\ell(2\ell-1)\mathbb{E}\left[\left(2 + \frac{2}{|F|^{2(\ell-1)}} + \frac{1}{|F^\star|^{2(\ell-1)}}\right)|F - F^\star|\right]\\
&+ 3a_\ell(2\ell-1)\mathbb{E}\left[\left(|F| + |F|^{2\ell-1}\right)\left|\frac{1}{F^{2\ell-1}} - \frac{1}{(F^\star)^{2\ell-1}}\right|\right].
\end{aligned}$$

Proof. The claim for even integers k is immediate. If $k = 2\ell - 1$ is an odd integer then

$$\begin{aligned}
d_W(F, F_\infty) \leq &\sum_{j=1}^{\ell} a_{\ell-j}(2\ell-1)\Bigg(\mathbb{E}\left|F^{2j-1} - (F^\star)^{2j-1}\right| + 2\mathbb{E}\ \left[|F|^{2j-1}\,|F - F^\star|\right]\Bigg)\\
&+ \sum_{j=1}^{\ell-1} a_j(2\ell-1)\Bigg(\mathbb{E}\left|\frac{1}{F^{2j-1}} - \frac{1}{(F^\star)^{2j-1}}\right| + \mathbb{E}\left[\frac{1}{|F|^{2j-1}}\Big|F - F^\star\Big|\right]\Bigg)\\
&+ \frac{a_\ell(2\ell-1)}{\sqrt{2}}\Psi_\ell(F, F^\star)
\end{aligned}$$

where

$$\begin{aligned}
\Psi_\ell(F, F^\star) = &\ \mathbb{E}\left|F^{2\ell-1}\epsilon_{2\ell-1}(F) - (F^\star)^{2\ell-1}\epsilon_{2\ell-1}(F^\star)\right| + 2\mathbb{E}\left[|F|^{2\ell-1}\epsilon_{2\ell-1}(F)|F - F^\star|\right]\\
&+ \mathbb{E}\left|F\epsilon_{2\ell-1}(F) - (F^\star)\epsilon_{2\ell-1}(F^\star)\right| + \mathbb{E}\left[|F|\epsilon_{2\ell-1}(F)|F - F^\star|\right]
\end{aligned}$$

Here we aim to bound

$$\begin{aligned}
\Psi_\ell(F, F^\star) = &\ \mathbb{E}\left|F^{2\ell-1}\epsilon_{2\ell-1}(F) - (F^\star)^{2\ell-1}\epsilon_{2\ell-1}(F^\star)\right| + 2\mathbb{E}\left[|F|^{2\ell-1}\epsilon_{2\ell-1}(F)|F - F^\star|\right]\\
&+ \mathbb{E}\left|F\epsilon_{2\ell-1}(F) - (F^\star)\epsilon_{2\ell-1}(F^\star)\right| + \mathbb{E}\left[|F|\epsilon_{2\ell-1}(F)|F - F^\star|\right]\\
=:&\ I + II + III + IV
\end{aligned}$$

where $\epsilon_{2\ell-1}(x) = e^{x^2/2}|x|^{-(2\ell-1)}\Gamma(1/2, x^2/2)$. We first note that, for $x > 0$, the function $x \mapsto \nu(x) := e^{x^2/2}\Gamma(1/2, x^2/2)$ is strictly decreasing as $|x| \to \infty$, with maximum value $\sqrt{\pi}$ at $x = 0$. Hence

$$II + IV \leq 2\sqrt{\pi}\mathbb{E}\left[|F - F^\star|\right] + \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F|^{2(\ell-1)}}|F - F^\star|\right].$$

Also, $|\nu'(x)| \le \sqrt{2}$ so that $|\nu(x) - \nu(y)| \le \sqrt{2}|y - x|$, hence

$$\begin{aligned}|\epsilon_{2\ell-1}(x) - \epsilon_{2\ell-1}(y)| &\le |x|^{-(2\ell-1)}|\nu(x) - \nu(y)| + |x^{-(2\ell-1)} - y^{-(2\ell-1)}|\nu(y) \\ &\le \sqrt{2}|x|^{-(2\ell-1)}|x - y| + \sqrt{\pi}|x^{-(2\ell-1)} - y^{-(2\ell-1)}|.\end{aligned}$$

This gives

$$\begin{aligned}I &\le \mathbb{E}\left[|F|^{2\ell-1}\left|\epsilon_{2\ell-1}(F) - \epsilon_{2\ell-1}(F^\star)\right|\right] + \mathbb{E}\left[\left|F^{2\ell-1} - (F^\star)^{2\ell-1}\right|\epsilon_{2\ell-1}(F^\star)\right] \\ &\le \mathbb{E}\left[|F|^{2\ell-1}\left|\epsilon_{2\ell-1}(F) - \epsilon_{2\ell-1}(F^\star)\right|\right] + \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F^\star|^{2\ell-1}}\left|F^{2\ell-1} - (F^\star)^{2\ell-1}\right|\right] \\ &\le \sqrt{2}\mathbb{E}\left[|F - F^\star|\right] + \sqrt{\pi}\mathbb{E}\left[|F|^{2\ell-1}\left|F^{-(2\ell-1)} - (F^\star)^{-(2\ell-1)}\right|\right] \\ &\quad + \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F^\star|^{2\ell-1}}\left|F^{2\ell-1} - (F^\star)^{2\ell-1}\right|\right].\end{aligned}$$

Similarly,

$$\begin{aligned}III &\le \mathbb{E}\left[|F|\left|\epsilon_{2\ell-1}(F) - \epsilon_{2\ell-1}(F^\star)\right|\right] + \mathbb{E}\left[|F - F^\star|\,\epsilon_{2\ell-1}(F^\star)\right] \\ &\le \mathbb{E}\left[|F|\left|\epsilon_{2\ell-1}(F) - \epsilon_{2\ell-1}(F^\star)\right|\right] + \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F^\star|^{2\ell-1}}|F - F^\star|\right] \\ &\le \sqrt{2}\mathbb{E}\left[|F|^{-2(\ell-1)}|F - F^\star|\right] + \sqrt{\pi}\mathbb{E}\left[|F|\left|F^{-(2\ell-1)} - (F^\star)^{-(2\ell-1)}\right|\right] \\ &\quad + \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F^\star|^{2\ell-1}}|F - F^\star|\right].\end{aligned}$$

Combining these bounds leads to

$$\begin{aligned}\Psi_\ell(F, F^\star) \le 2\sqrt{\pi}\mathbb{E}\left[|F - F^\star|\right] &+ \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F|^{2(\ell-1)}}|F - F^\star|\right] \\ &+ \sqrt{2}\mathbb{E}\left[|F - F^\star|\right] + \sqrt{\pi}\mathbb{E}\left[|F|^{2\ell-1}\left|F^{-(2\ell-1)} - (F^\star)^{-(2\ell-1)}\right|\right] \\ &+ \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F^\star|^{2\ell-1}}\left|F^{2\ell-1} - (F^\star)^{2\ell-1}\right|\right] \\ &+ \sqrt{2}\mathbb{E}\left[|F|^{-2(\ell-1)}|F - F^\star|\right] + \sqrt{\pi}\mathbb{E}\left[|F|\left|F^{-(2\ell-1)} - (F^\star)^{-(2\ell-1)}\right|\right] \\ &+ \sqrt{\pi}\mathbb{E}\left[\frac{1}{|F^\star|^{2\ell-1}}|F - F^\star|\right]\end{aligned}$$

which, after bounding all constants by $3\sqrt{2}$ for simplicity, gives

$$\begin{aligned}\Psi_\ell(F, F^\star) \le 3\sqrt{2}\mathbb{E}&\left[\left(2 + \frac{2}{|F|^{2(\ell-1)}} + \frac{1}{|F^\star|^{2(\ell-1)}}\right)|F - F^\star|\right] \\ &+ 3\sqrt{2}\mathbb{E}\left[\left(|F| + |F|^{2\ell-1}\right)\left|F^{-(2\ell-1)} - (F^\star)^{-(2\ell-1)}\right|\right],\end{aligned}$$

which is the claim. □

Corollary 2. *In terms of the notation in Theorem 2, the following bounds hold:*

- $k = 2$ *(Maxwell case):* $p_\infty(x) = x^2\varphi(x)$, $\mathbb{E}[F^2] = 3$, $\mathbb{E}[f'(F^\star)/(F^\star)^2] = \mathbb{E}[f(F)/|F|]$ *for all* f, *and*

$$d_W(F, F_\infty) \leq \mathbb{E}\left[\left(3 + 2|F| + \frac{2}{|F|} + \frac{2}{|F||F^\star|}\right)|F - F^\star|\right].$$

- $k = 3$*:* $p_\infty(x) \propto |x|^3\varphi(x)$, $\mathbb{E}[F^2] = 4$, $\mathbb{E}[f'(F^\star)/|F^\star|^3] = \mathbb{E}[f(F)/F^2]$ *for all* f, *and*

$$d_W(F, F_\infty) \leq \mathbb{E}\left[\left(21 + 6|F| + 2F^2 + \frac{3}{|F|} + \frac{18}{F^2} + \frac{9}{(F^\star)^2} + \frac{3}{|F||F^\star|}\right)|F - F^\star|\right] + \mathbb{E}|F^2 - (F^\star)^2| + 9\mathbb{E}\left[(|F| + |F|^3)\left|\frac{1}{F^3} - \frac{1}{(F^\star)^3}\right|\right].$$

Remark 5. We note that the Maxwell bound ($k = 2$) is the same as in [7, Equation (24)] but with improved constants.

2.5. Upper bounds on the rate of convergence via coupling

In this section we apply Theorem 2 to give explicit upper bounds on the accuracy of $F = F_N \sim \mathbb{F}_N$ (defined after Lemma 3 of the main text) in approximating the radial distributions with density $p_\infty(x) \propto |x|^k\varphi(x)$ for $k = 2, \ldots, 14$.

Two crucial steps in controlling the various error terms in Theorem 2 when $k \geq 1$ (in which case there is a singularity in p_∞ at the origin) are to obtain good approximations to x_1 and x_m (where $m = N/2$), which (when necessary) we write as $x_{1,k}$ and $x_{m,k}$, to reflect their dependence on k. It is easily shown by extending [7, Lemma 4.8] that

$$x_{1,k} = O(\sqrt{\log N})$$

for all $k \geq 0$. Concerning $x_{m,k}$, we have $x_{m,k} \geq 1/\sqrt{N}$ for all $k \geq 1$, cf. [7, proof of Corollary 3.7]. However, for $k \geq 3$ a lower bound of order $1/\sqrt{N}$ is too small to control all error terms in Theorem 2, and it seems very challenging to improve this lower bound analytically.

A way around this difficulty is to examine the numerical behavior of $x_{m,k}$. We find that the following scaling relation provides a remarkably accurate approximation:

$$x_{m,k} \asymp N^{-1/r_k}, \quad \text{where } r_k = 3/2 + 4k/5 \tag{24}$$

for $k = 2, 3, \ldots, 14$, see Figure 1. We expect this scaling relation holds for all $k \geq 2$.

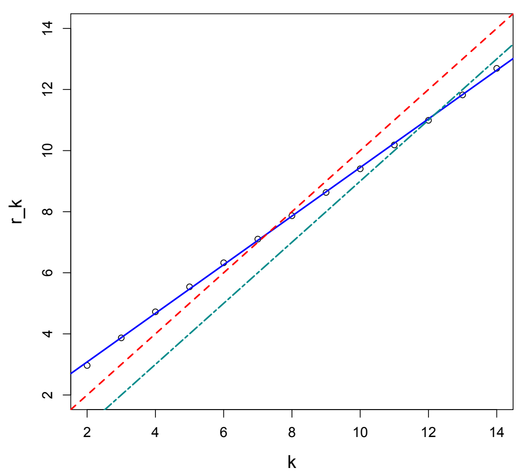

FIGURE 1: The graph of $r_k = 3/2 + 4k/5$ (blue solid line) obtained by fitting the scaling relation (24) for $k = 2, \ldots, 14$. The fitted line approximates data points (circles) derived from estimates of the slope when regressing $\log x_{m,k}$ against $\log N$, $N = 14, 24, \ldots, 114$, separately for each value of k. It is remarkable that there is virtually no scatter around any of the linear fits. For k odd, the fitted line needs to fall above the diagonal (red dashed) line to ensure adequate control of (25) when $l = k$. For k even, the fitted line needs to fall above the dot-dashed (cyan) line to obtain control when $l = k - 1$.

The next step is to use the scaling relation to obtain upper bounds on the error terms in Theorem 2 of the form

$$\mathbb{E}\left|\frac{1}{F^l}-\frac{1}{(F^\star)^l}\right|, \quad 1\le l\le k \tag{25}$$

where $F\sim\mathbb{F}_N$, the empirical distribution of $x_1>\ldots>x_N$, and $F^\star\sim\mathbb{F}_N^\star$ is the corresponding k-radial-bias distribution, as given in the following lemma, which is a consequence of [7, Proposition 3.5].

Lemma 6. *The k-radial-bias distribution $\mathbb{F}_N^\star$ of $\mathbb{F}_N$ is defined, and has density*

$$p^\star(x)\propto|x|^k\left[\sum_{i=1}^n\frac{x_i}{|x_i|^k}\right]$$

for $x_{n+1}<x\le x_n$ $(n=1,\ldots,N-1)$, and $p^\star(x)=0$ if $x>x_1$ or $x\le x_N$.

From Lemma 6 and the recursion satisfied by $x_1>\ldots>x_N$, it follows that $p^\star(x)$ puts mass $1/(N-1)$ on each interval between successive x_n, so there exists a coupling of $F\sim\mathbb{P}_N$ with $F^\star\sim p^\star(x)$ such that

$$|F-F^\star|\le|x_n-x_{n+1}|$$

when $F\in[x_{n+1},x_n]$. For a detailed proof of such a coupling, see the construction given in [6]. Now decompose (25) as

$$\mathbb{E}\left|\frac{1}{F^l}-\frac{1}{(F^\star)^l}\right|=\mathbb{E}\left|\frac{1}{F^l}-\frac{1}{(F^\star)^l}\right|1_{F^\star\in(x_{m+1},x_m]}+2\sum_{n=1}^{m-1}\mathbb{E}\left|\frac{1}{F^l}-\frac{1}{(F^\star)^l}\right|1_{F^\star\in(x_{n+1},x_n]}.$$

From Proposition 6 note that $p^\star(x)\propto|x|^k$ for $x\in(x_{m+1},x_m]$. Using the fact that $p^\star(x)$ puts mass $1/(N-1)$ on this interval, the first term above can be written

$$\frac{2(k+1)}{x_m^{k+1}(N-1)}\int_0^{x_m}\left(\frac{1}{x^l}-\frac{1}{x_m^l}\right)x^k\,dx\asymp\frac{1}{x_m^lN}\asymp N^{l/r_k-1}\to0,$$

provided $l<3/2+4k/5$ by (24). The second term is bounded above by the telescoping sum

$$\frac{2}{N-1}\sum_{n=1}^{m-1}\left(\frac{1}{x_{n+1}^l}-\frac{1}{x_n^l}\right)=\frac{2}{N-1}\left(\frac{1}{x_m^l}-\frac{1}{x_1^l}\right)=O\left(N^{l/r_k-1}\right),$$

so we conclude

$$\mathbb{E}\left|\frac{1}{F^l} - \frac{1}{(F^\star)^l}\right| = O\left(N^{l/r_k - 1}\right).$$

This bound gives the desired control of (25) for any $l \leq k \leq 7$, see Figure 1. For even k, we only need to consider $l \leq k-1$, so we have control for $k = 8, 10$ and 12, as well.

References

[1] L. H. Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg, 2011.

[2] M. Ernst, G. Reinert, and Y. Swan. First-order covariance inequalities via Stein's method. *Bernoulli*, 26(3):2051–2081, 2020.

[3] M. Ernst and Y. Swan. Distances between distributions via Stein's method. *Journal of Theoretical Probability*, 2020.

[4] G. Jameson. The incomplete gamma functions. *The Mathematical Gazette*, 100(548):298–306, 2016.

[5] C. Ley, G. Reinert, and Y. Swan. Distances between nested densities and a measure of the impact of the prior in Bayesian statistics. *The Annals of Applied Probability*, 27(1):216–241, 2017.

[6] I. W. McKeague and B. Levin. Convergence of empirical distributions in an interpretation of quantum mechanics. *The Annals of Applied Probability*, 26(4):2540–2555, 2016.

[7] I. W. McKeague, E. A. Peköz, and Y. Swan. Stein's method and approximating the quantum harmonic oscillator. *Bernoulli*, 25(1):89–111, 02 2019.