

Stein’s method for the half-normal distribution with applications to limit theorems related to the simple symmetric random walk

Christian Döbler

Université du Luxembourg
Unité de Recherche en Mathématiques
Campus Kirchberg
6, rue Richard Coudenhove-Kalergi
L-1359 Luxembourg,
Luxembourg

E-mail address: christian.doebler@uni.lu

URL: http://www.uni.lu/recherche/fstc/mathematics_research_unit/people/christian_doebler

Abstract. We develop Stein’s method for the half-normal distribution and apply it to derive rates of convergence in distributional limit theorems for three statistics of the simple symmetric random walk: the maximum value, the number of returns to the origin and the number of sign changes up to a given time n . We obtain explicit error bounds with the optimal rate $n^{-1/2}$ for both the Kolmogorov and the Wasserstein metric. In order to apply Stein’s method, we compare the characterizing operator of the limiting half-normal distribution with suitable characterizations of the discrete approximating distributions, exploiting a recent technique by [Goldstein and Reinert \(2013\)](#).

1. Introduction

This article concerns the rate of convergence issue for three limit theorems in the surroundings of the one-dimensional simple symmetric random walk (SRW). By this we mean the discrete time stochastic process $(S_n)_{n \geq 0}$ defined by $S_0 := 0$ and $S_n := \sum_{j=1}^n X_j$, $n \geq 1$, where X_1, X_2, \dots , are iid random variables with $P(X_1 = 1) = P(X_1 = -1) = 1/2$. It has been known for a long time that various statistics of the process $(S_n)_{n \geq 0}$ exhibit a quite counter intuitive distributional behaviour, see e.g. Chapter 3 of [Feller \(1968\)](#) for some qualitative limit theorems. As to quantitative results, in [Döbler \(2012a\)](#) a rate of convergence for the arcsine

Received by the editors March 23, 2014; accepted March 3, 2015.

2010 Mathematics Subject Classification. 60F05, 60G50.

Key words and phrases. Stein’s method, half-normal distribution, simple random walk, density approach, Kolmogorov distance, Wasserstein distance.

law was proved using Stein's method for Beta distributions (see also [Goldstein and Reinert \(2013\)](#) for an improvement of this result with respect to the computation of an explicit constant of convergence and also [Döbler \(2014\)](#) for a further improvement of this constant).

Here, we will focus on limit theorems which state convergence towards the distribution μ of $Y := |Z|$, where $Z \sim N(0, 1)$ is standard normally distributed. This distribution is commonly known as the (standard) *half-normal distribution*. The technique of proof will be to compare the Stein characterization of Y with a suitable characterization for the approximating discrete distribution, as was proposed by [Goldstein and Reinert \(2013\)](#) and also used in [Döbler \(2012a\)](#). Generally, this technique is promising, whenever a concrete formula for the probability mass function of the discrete distribution is at hand, which yields a Stein characterization similar to the one for the limiting distribution. Note that even though the probability mass function and, hence, the discrete distribution must be known for this approach to be applicable, it might still be of interest to approximate by an easier to handle absolutely continuous distribution. In fact, already the de Moivre-Laplace theorem gives an approximation to the known binomial by the normal distribution. In the case of $Y = |Z|$ a suitable Stein characterization is easily found, using the *density approach* of Stein's method (see [Chatterjee and Shao \(2011\)](#); [Eichelsbacher and Löwe \(2010\)](#); [Chen et al. \(2011\)](#), for instance). Although we could simply quote the theory and general bounds on the solution to the Stein equation from [Chen et al. \(2011\)](#) or [Chatterjee and Shao \(2011\)](#), for example, we prefer deriving our own bounds, which usually yield better constants.

The rate of convergence results in this paper are always with respect to a certain probability metric, which is defined via test functions. Thus, if μ and ν are two probability measures on $(\mathbb{R}, \mathcal{B})$ and \mathcal{H} is some class of measurable test functions that are integrable with respect to μ and ν , then we define the distance

$$d_{\mathcal{H}}(\mu, \nu) := \sup_{h \in \mathcal{H}} \left| \int_{\mathbb{R}} h d\mu - \int_{\mathbb{R}} h d\nu \right|.$$

For example, if \mathcal{W} is the class of Lipschitz continuous functions on \mathbb{R} with Lipschitz constant not greater than 1 and if μ and ν both have first moments, then $d_{\mathcal{W}}(\mu, \nu)$ is the *Wasserstein distance* between μ and ν . On the other hand, if \mathcal{K} is the class of functions $h_z := 1_{(-\infty, z]}$, $z \in \mathbb{R}$, then we obtain the *Kolmogorov distance*

$$d_{\mathcal{K}}(\mu, \nu) = \sup_{z \in \mathbb{R}} \left| \mu((-\infty, z]) - \nu((-\infty, z]) \right|,$$

which is particularly natural from a statistician's point of view. For real-valued random variables X and Y we write $d_{\mathcal{H}}(X, Y)$ for $d_{\mathcal{H}}(\mathcal{L}(X), \mathcal{L}(Y))$.

Now, we introduce the statistics of $(S_n)_{n \geq 0}$, which converge in distribution to Y . First, consider the number K_n of times that the random walk returns to the origin up to time n , i.e.

$$K_n := |\{1 \leq k \leq n = 2m : S_k = 0\}|. \quad (1.1)$$

The variable K_n is sometimes called the *local time* or *occupation time* at 0 by time n . From Theorem 7, Section 5 of [Feller \(1949\)](#) (see also Equation (5.31) there) it is known that $K_n/\sqrt{n} \xrightarrow{\mathcal{D}} Y$ as $n \rightarrow \infty$ (see also Equation (1) of [Chung and Hunt \(1949\)](#)). Note that intuition might lead us to the (false) conclusion that the number

of returns to the origin should roughly grow linearly with the time n . Here is a theorem which gives error bounds for this distributional convergence.

Theorem 1.1. *Let $n = 2m$ be an even positive integer. Then, with $W := W_n := K_n/\sqrt{n}$*

$$d_{\mathcal{W}}(W, Y) \leq \frac{1}{\sqrt{n}} \left(\frac{2}{\pi} + 2 \right) + \frac{1}{n} \sqrt{\frac{2}{\pi}} \quad \text{and}$$

$$d_{\mathcal{K}}(W, Y) \leq \frac{1}{\sqrt{n}} \left(\frac{3 + 2\sqrt{2}}{\sqrt{2\pi}} + \frac{3}{4} \right) + \frac{3}{2n}.$$

It should be mentioned that the rate $n^{-1/2}$ for $d_{\mathcal{K}}(W, Y)$ was also given in [Peköz et al. \(2013\)](#) but they did not compute an explicit constant. Next, consider

$$M_n := \max_{0 \leq k \leq n} S_k \in \{0, 1, \dots, n\}. \quad (1.2)$$

Then, $M_n/\sqrt{n} \xrightarrow{\mathcal{D}} Y$ as $n \rightarrow \infty$. This follows by an application of the CLT to the result of Theorem 1 in Section 7 of Chapter 3 in [Feller \(1968\)](#). We will prove the following quantitative version of this result.

Theorem 1.2. *Let $n = 2m$ be an even positive integer. Then, with $W := W_n := M_n/\sqrt{n}$*

$$d_{\mathcal{W}}(W, Y) \leq \frac{1}{\sqrt{n}} \left(3 + \frac{2}{\pi} \right) \quad \text{and}$$

$$d_{\mathcal{K}}(W, Y) \leq \frac{1}{\sqrt{n}} \left(4\sqrt{\frac{2}{\pi}} + \frac{1}{2} \right) + \frac{2}{n}.$$

Finally, consider the number C_n of sign changes by the random walk up to time $n := 2m + 1$, $m \in \mathbb{N}$, i.e.

$$C_n := C_{2m+1} := |\{1 \leq k \leq 2m : S_{k-1} \cdot S_{k+1} < 0\}|. \quad (1.3)$$

Then, Theorem 2 in Section 5 of Chapter 3 in [Feller \(1968\)](#) states that $2C_{2m+1}/\sqrt{2m+1} \xrightarrow{\mathcal{D}} Y$ as $m \rightarrow \infty$. Again, we obtain a quantitative version of this result.

Theorem 1.3. *Let m be a positive integer and $n := 2m + 1$. Then, with $W := W_m := 2C_{2m+1}/\sqrt{2m+1}$*

$$d_{\mathcal{W}}(W, Y) \leq \frac{1}{\sqrt{n}} \left(4 + \frac{2}{\pi} \right) + \sqrt{\frac{2}{\pi}} \frac{1}{n} + \frac{2\sqrt{2}}{\pi} \frac{1}{n^{3/2}} \quad \text{and}$$

$$d_{\mathcal{K}}(W, Y) \leq \frac{1}{\sqrt{n}} \left(\frac{2\sqrt{2} + 4}{\sqrt{\pi}} + \frac{3}{2} \right) + \frac{3}{n} + \frac{4}{\sqrt{\pi}} \frac{1}{n^{3/2}}.$$

The rest of the paper is organized as follows. In Section 2 we develop Stein's method for the half-normal distribution of Y and review the technique by [Goldstein and Reinert \(2013\)](#) of comparison with a discrete distribution. In Sections 3, 4 and 5 we present the proofs of Theorems 1.1, 1.2 and 1.3, respectively and in Section 6 we show the optimality of the obtained convergence rates. Finally, in Section 7 we give proofs of some of the results from Section 2.

2. Stein's method for the half-normal distribution and for discrete distributions

Stein's method is by now a well-established device for proving concrete error bounds in distributional convergence problems. Since its introduction by Stein in the seminal paper [Stein \(1972\)](#) on univariate normal approximation for sums of random variables, satisfying a certain mixing condition, it has undergone remarkable progress. On the one hand, the range of normal approximation problems that can be tackled by means of the method has been largely extended, particularly due to the development of certain coupling constructions (see [Chen et al. \(2011\)](#) for an introduction and overview). On the other hand, the essential idea of characterizing a given distribution by a certain differential or difference equation, was successfully carried over to other prominent distributions, like, for instance, the Poisson distribution (see e.g. [Chen \(1975\)](#) and [Barbour et al. \(1992\)](#)), the Gamma distribution (see [Luk \(1994\)](#)), the exponential distribution (see e.g. [Chatterjee et al. \(2011\)](#); [Peköz and Röllin \(2011\)](#); [Fulman and Ross \(2013\)](#)), the Beta distribution (see [Goldstein and Reinert \(2013\)](#); [Döbler \(2012b, 2014\)](#)), the Laplace distribution (see [Pike and Ren \(2014\)](#)) and, more generally, the recent article [Gaunt \(2014\)](#) on the class of Variance-Gamma distributions. Furthermore, general techniques have been proposed to develop Stein's method for a distribution with a given density, like for example the density approach (see [Stein et al. \(2004\)](#); [Chatterjee and Shao \(2011\)](#); [Eichelsbacher and Löwe \(2010\)](#); [Chen et al. \(2011\)](#)) or the general approach in [Döbler \(2014\)](#), which is adapted to a given exchangeable pair. For a nice recent generalization of the density approach also see [Ley et al. \(2014\)](#).

Before we develop Stein's method for the half-normal distribution, let us make the following remark. Note that since $Y = |Z|$ and the random variable W in each of the Theorems [1.2-1.3](#) is nonnegative, it would in principle be possible to apply Stein's method for the standard normal distribution for their proofs. Indeed, if \mathcal{H} is a given class of test functions on $[0, \infty)$ and if for $h \in \mathcal{H}$ we define the function g on \mathbb{R} by $g(x) := h(|x|)$ and denote by \mathcal{G} the class of all those functions g , when h is running through \mathcal{H} , we have that

$$\begin{aligned} d_{\mathcal{H}}(W, Y) &= \sup_{h \in \mathcal{H}} |E[h(W)] - E[h(Y)]| = \sup_{g \in \mathcal{G}} |E[g(W)] - E[g(Z)]| \\ &= \sup_{g \in \mathcal{G}} |E[\tilde{f}'_g(W) - W\tilde{f}_g(W)]|, \end{aligned}$$

where we denote by \tilde{f}_g the standard solution to the standard normal Stein equation corresponding to the test function g . Unfortunately, unless Lipschitz-continuous test functions are considered, in general the bounds on the functions \tilde{f}_g , $g \in \mathcal{G}$, and on their lower order derivatives, which could be derived from known results from Stein's method of normal approximation, are worse than the bounds we obtain for the functions f_h in Lemmas [2.2](#) and [2.3](#). As a consequence, following this route would yield worse constants in Theorems [1.1-1.3](#) than we will obtain by considering Stein's method for the half-normal distribution itself, at least as far as the Kolmogorov distance is concerned.

However, for some of our bounds, we actually will use results from Stein's method for the standard normal distribution observing that for $x \geq 0$ the solution $f_h(x)$ to the half-normal Stein equation [\(2.3\)](#) given by [\(2.5\)](#), [\(2.6\)](#) and $\tilde{f}_g(x)$ coincide, which makes it possible to shorten the exposition of some of our proofs. In our opinion,

this compromise eventually justifies developing a version of Stein's method for the half-normal distribution instead of just applying existing results on normal approximation.

We begin by developing Stein's method for the half-normal distribution. Note that μ is supported on $[0, \infty)$, since $Y = |Z|$, where $Z \sim N(0, 1)$. We denote by p and F the (continuous) density function and distribution function of Y , respectively. Thus, as a trivial computation shows, we have

$$p(x) = 2\varphi(x)1_{(0,\infty)}(x) = \sqrt{\frac{2}{\pi}}e^{-x^2/2}1_{(0,\infty)}(x) \text{ and} \quad (2.1)$$

$$F(x) = (2\Phi(x) - 1)1_{(0,\infty)}(x), \quad (2.2)$$

where φ, Φ denote the (continuous) density function and distribution function of Z , respectively. According to the density approach in Stein's method, we have the following result.

Proposition 2.1 (Stein characterization). *A random variable X with values in $[0, \infty)$ has the half-normal distribution μ if and only if*

$$E[f'(X)] = E[Xf(X)] - f(0)\sqrt{\frac{2}{\pi}}$$

for all functions $f : [0, \infty) \rightarrow \mathbb{R}$, which are absolutely continuous on every compact sub-interval of $[0, \infty)$ such that $E|f'(Y)| < \infty$.

As we do not explicitly need the result of Proposition 2.1, we omit the rather standard proof. For a given measurable function h on $[0, \infty)$ with $E|h(Y)| < \infty$ Proposition 2.1 now motivates the following *half-normal Stein equation*

$$f'(x) - xf(x) = h(x) - \mu(h), \quad (2.3)$$

where we abbreviate $\mu(h) := E[h(Y)]$. This equation is to be solved for the function f on $[0, \infty)$. If f is a solution to (2.3) and W is a given nonnegative random variable, then, taking expectations, we have the following identity:

$$E[h(W)] - E[h(Y)] = E[f'(W) - Wf(W)] \quad (2.4)$$

As a matter of fact, the right hand side of (2.4) may often be bounded more easily (even uniformly in the test functions h from some class \mathcal{H} of functions) than the left hand side, if one further tool is available. This additional tool may be a coupling, like for example the exchangeable pairs coupling (see Chen et al. (2011)), or a characterization for the distribution $\mathcal{L}(W)$, as will be exploited in this paper.

We now introduce the standard solution f_h to (2.3). Let $f_h : [0, \infty) \rightarrow \mathbb{R}$ be defined by

$$\begin{aligned} f_h(x) &:= \frac{1}{p(x)} \int_0^x (h(t) - \mu(h))p(t)dt \\ &= \frac{1}{\varphi(x)} \int_0^x (h(t) - \mu(h))\varphi(t)dt \end{aligned} \quad (2.5)$$

$$= -\frac{1}{\varphi(x)} \int_x^\infty (h(t) - \mu(h))\varphi(t)dt. \quad (2.6)$$

It is easily checked that f_h indeed solves Equation (2.3) and that the solutions of the homogeneous equation corresponding to (2.3) have the form $ce^{x^2/2}$ for some

constant $c \in \mathbb{R}$. This particularly shows that if f_h is bounded, then it is the only bounded solution to (2.3) and also the only solution f with $\lim_{x \rightarrow \infty} e^{-x^2/2} f(x) = 0$. Note that the half-normal Stein equation (2.3) is the same as the standard normal one (see e.g. [Chen et al. \(2011\)](#)) except that we only consider functions on $[0, \infty)$, here. Thus, there should also be some correspondence between the solutions. Indeed, if as above, for given h on $[0, \infty)$, we consider the function g on \mathbb{R} given by $g(x) := h(|x|)$ and denote by \tilde{f}_g the standard solution to the Stein equation for the standard normal distribution $\mathcal{L}(Z)$ and the test function g , we obtain for each $x \geq 0$ that

$$\begin{aligned} \tilde{f}_g(x) &= -\frac{1}{\varphi(x)} \int_x^\infty (g(t) - E[g(Z)]) \varphi(t) dt \\ &= -\frac{1}{\varphi(x)} \int_x^\infty (h(t) - E[h(Y)]) \varphi(t) dt \\ &= f_h(x) \end{aligned} \tag{2.7}$$

by (2.6). Thus, \tilde{f}_g coincides with f_h on $[0, \infty)$. This allows us to derive properties of the solutions f_h to (2.3) from those of the functions \tilde{f}_g , which are well-studied. For example note that, if h is Lipschitz on $[0, \infty)$ with Lipschitz constant $L > 0$, then for $x, y \in \mathbb{R}$

$$|g(x) - g(y)| = |h(|x|) - h(|y|)| \leq L||x| - |y|| \leq L|x - y|. \tag{2.8}$$

So, g is also Lipschitz with the same constant L .

In order to make good use of identity (2.4) one needs bounds on the solutions f_h and their lower order derivatives. The following lemma gives bounds for bounded measurable or Lipschitz test functions h .

Lemma 2.2. *Let $h : [0, \infty) \rightarrow \mathbb{R}$ be Borel-measurable.*

- (i) *If h is bounded, then f_h is Lipschitz and with $z_{0.75} := \Phi^{-1}(3/4)$ we have*
 - (a) $\|f_h\|_\infty \leq \frac{\|h - \mu(h)\|_\infty}{4\varphi(z_{0.75})}$,
 - (b) $\|f'_h\|_\infty \leq 2\|h - \mu(h)\|_\infty$.
- (ii) *If h is Lipschitz, then f_h is continuously differentiable with a Lipschitz continuous derivative and we have the bounds*
 - (a) $\|f_h\|_\infty \leq \|h'\|_\infty$,
 - (b) $\|f'_h\|_\infty \leq \sqrt{\frac{2}{\pi}} \|h'\|_\infty$,
 - (c) $\|f''_h\|_\infty \leq 2\|h'\|_\infty$.

Proof: Assertion (a) of (i) is proved in Section 7. By (2.7) $f_h(x)$ coincides with the solution $\tilde{f}_g(x)$ to the Stein equation for the standard normal distribution corresponding to the test function $g(x) = h(|x|)$. Since

$$\sup_{x \in \mathbb{R}} |g(x) - E[g(Z)]| = \sup_{x \geq 0} |h(x) - E[h(Y)]|, \quad E[g(Z)] = E[h(Y)] \quad \text{and,}$$

by (2.8), $\|g'\|_\infty = \|h'\|_\infty$ bound (b) of (i) and the bounds in (ii) follow from well-known bounds in the standard normal case (see [Chen et al. \(2011\)](#))). □

Lipschitz test functions h with Lipschitz constant $\|h'\|_\infty \leq 1$ yield the Wasserstein distance. The Kolmogorov distance, in contrast, is induced by the class of

functions $h_z = 1_{(-\infty, z]}$, $z \in \mathbb{R}$. Since we only compare distributions with support on $[0, \infty)$ we may restrict ourselves to the case $z \geq 0$. We write $f_z := f_{h_z}$. Since the functions h_z are uniformly bounded by 1, we immediately get bounds on $\|f_z\|_\infty$ and $\|f'_z\|_\infty$ from Lemma 2.2 (i). But for this particular class of functions, a special analysis indeed yields smaller constants. If we denote by \tilde{f}_z the standard solution to Stein's equation for the standard normal distribution with respect to the test function h_z , where $z \in \mathbb{R}$, we obtain from (2.7) for $z \geq 0$ that $f_z(x) = \tilde{f}_z(x) - \tilde{f}_{-z}(x)$ for each $x \geq 0$. Indeed, letting $g_z(x) := h_z(|x|)$, $x \in \mathbb{R}$, we have $g_z(x) = h_z(x) - 1_{(-\infty, -z)}(x)$ which equals $h_z(x) - h_{-z}(x)$ for every $x \neq -z$, leading to this representation of f_z . However, we were not able to use this representation and known properties of the functions \tilde{f}_z to derive the properties of f_z stated in the following lemma. This is why a complete proof is given in Section 7.

Lemma 2.3. *For each $x, y, z \geq 0$ we have:*

- (a) *The function $[0, \infty) \ni x \mapsto xf_z(x) \in \mathbb{R}$ is increasing and $0 \leq xf_z(x) \leq 2\Phi(z) - 1 < 1$.*
- (b) *$0 < f_z(x) \leq \frac{1}{2}$.*
- (c) *$\|f'_z\|_\infty \leq 1$ and $|f'_z(x) - f'_z(y)| \leq 1$.*

Remark 2.4. The bound in Lemma 2.3 (b) is not optimal. More precisely, computer algebra systems suggest that

$$\sup_{z \geq 0} \|f_z\|_\infty = 0.456296\dots$$

However, the bounds in Lemma 2.3 (c) are optimal, which is proved in Section 7.

In the following, we review the technique of finding a suitable Stein type characterization for a discrete distribution on the integers by Goldstein and Reinert [Goldstein and Reinert \(2013\)](#). A finite integer interval is a set I of the form $I = [a, b] \cap \mathbb{Z}$ for some integers $a \leq b$. Given a probability mass function $p : \mathbb{Z} \rightarrow \mathbb{R}$ with $p(k) > 0$ for $k \in I$ and $p(k) = 0$ for $k \in \mathbb{Z} \setminus I$, we consider the function $\psi : I \rightarrow \mathbb{R}$ given by the formula

$$\psi(k) := \frac{\Delta p(k)}{p(k)} = \frac{p(k+1) - p(k)}{p(k)}, \quad (2.9)$$

where for a function f on the integers $\Delta f(k) := f(k+1) - f(k)$ denotes the *forward difference operator*. The next result, a version of Corollary 2.1 from [Goldstein and Reinert \(2013\)](#), yields various Stein characterizations for the distribution corresponding to p . For such a probability mass function p with support a finite integer interval $I = [a, b] \cap \mathbb{Z}$, let $\mathcal{F}(p)$ denote the class of all real-valued functions f on \mathbb{Z} such that $f(a-1) = 0$.

Proposition 2.5. *Let p be a probability mass function which is supported on the finite integer interval $I = [a, b] \cap \mathbb{Z}$ and is positive there. Let $c : [a-1, b] \cap \mathbb{Z} \rightarrow \mathbb{R}$ be a function with $c(k) \neq 0$ for all $k \in I$. Then, in order that a given random variable X with support I is distributed according to p it is necessary and sufficient that for all functions $g \in \mathcal{F}(p)$ we have*

$$E \left[c(X-1)\Delta g(X-1) + [c(X)\psi(X) + \Delta c(X-1)]g(X) \right] = 0. \quad (2.10)$$

3. The number of returns to the origin

Recall the definition of K_n from (1.1). In this section we give the proof of Theorem 1.1. It is known (see e.g. Feller (1968), Problem 9 in Chapter 3) that for each $r \in \{0, 1, \dots, m\}$

$$p(r) := P(K_n = r) = \frac{1}{2^{n-r}} \binom{n-r}{n/2} = \frac{1}{2^{2m-r}} \binom{2m-r}{m}. \quad (3.1)$$

Using (3.1) as well as the relation

$$\binom{n+1}{k} = \frac{n+1}{n-k+1} \binom{n}{k}$$

we obtain that

$$\begin{aligned} \psi(r) &:= \frac{p(r+1) - p(r)}{p(r)} = \frac{2^{-(2m-r-1)} \binom{2m-r-1}{m} - 2^{-(2m-r)} \binom{2m-r}{m}}{2^{-(2m-r)} \binom{2m-r}{m}} \\ &= 2 \frac{m-r}{2m-r} - 1 = \frac{-r}{2m-r} \end{aligned} \quad (3.2)$$

for all $r \in \{0, 1, \dots, m\}$. We thus define $c(r) := 2m - r$ for $r = -1, 0, \dots, m$ and obtain

- $\Delta c(r-1) = c(r) - c(r-1) = 2m - r - (2m - r + 1) = -1$
- $c(r)\psi(r) = -r$ and
- $\gamma(r) := c(r)\psi(r) + \Delta c(r-1) = -r - 1 = -(r+1)$.

Proposition 2.5 thus yields the following characterization of the distribution of K_n .

Lemma 3.1. *A random variable X with support $[0, m] \cap \mathbb{Z}$ has probability mass function p if and only if for every function $g \in \mathcal{F}(p)$*

$$E[(2m - X + 1)\Delta g(X-1) - (X+1)g(X)] = 0.$$

Letting $g(k) := 1$ for $k \geq 0$ and $g(k) := 0$ for $k < 0$ we obtain that

$$\begin{aligned} E[K_n] &= (2m+1)P(K_n = 0) - 1 \\ &= (2m+1)2^{-2m} \binom{2m}{m} - 1 \leq 2m2^{-2m} \binom{2m}{m} \\ &\leq \frac{2m}{\sqrt{\pi m}} = \sqrt{\frac{2}{\pi}} \sqrt{n}, \end{aligned} \quad (3.3)$$

where the last inequality follows from Stirling's formula.

Remark 3.2. It is not easy to obtain the above formula for $E[K_n]$ directly from the definition of the expected value, but some combinatorial tricks are needed. Also note that bound (3.3) on $E[K_n]$ is quite accurate. Furthermore, since $K_n = \sum_{j=1}^m 1_{A_j}$, where $A_j := \{S_{2j} = 0\}$, we have $E[K_n] = \sum_{j=1}^m 2^{-2j} \binom{2j}{j}$ and this is the partial sum of the diverging series which is often used to prove recurrence of the simple symmetric random walk. Thus, we see how the Stein characterization in Lemma 3.1 gives us information about how fast this series diverges. This indicates that a lot of information might be encoded in such a Stein identity.

Proof of Theorem 1.1: Let h be a Borel-measurable test function and consider the corresponding standard solution $f := f_h$ to Stein's equation (2.3) given by (2.5)

and (2.6). We also let $f_h(x) := 0$ for each $x < 0$ and define $g(k) := f_h(k/\sqrt{n})$ for $k \in \mathbb{Z}$. Writing $\Delta_y f(x) := f(x+y) - f(x)$ we obtain from Lemma 3.1

$$\begin{aligned} E[Wf(W)] &= \frac{1}{\sqrt{n}}E[K_n g(K_n)] = \frac{1}{\sqrt{n}}E[(n - K_n + 1)\Delta g(K_n - 1) - g(K_n)] \\ &= E\left[(\sqrt{n} - W + n^{-1/2})\Delta_{n^{-1/2}}f(W - n^{-1/2})\right] - \frac{1}{\sqrt{n}}E[f(W)] \\ &= \sqrt{n}E[\Delta_{n^{-1/2}}f(W - n^{-1/2})] + E_1, \end{aligned} \quad (3.4)$$

where

$$\begin{aligned} |E_1| &= \left|E\left[(n^{-1/2} - W)\Delta_{n^{-1/2}}f(W - n^{-1/2})\right] - \frac{1}{\sqrt{n}}E[f(W)]\right| \\ &\leq \|f'\|_\infty \left(\frac{1}{n} + \frac{1}{\sqrt{n}}E[W]\right) + \frac{\|f\|_\infty}{\sqrt{n}} \\ &\leq \frac{\|f'\|_\infty}{\sqrt{n}} \left(\frac{1}{\sqrt{n}} + \sqrt{\frac{2}{\pi}}\right) + \frac{\|f\|_\infty}{\sqrt{n}} \end{aligned} \quad (3.5)$$

by inequality (3.3) and because $W = n^{-1/2}K_n$. Now, using (3.4) and the fact that $f = f_h$ is a solution to the half-normal Stein equation (2.3) we obtain

$$\begin{aligned} |E[h(W)] - E[h(Y)]| &= |E[f'(W) - Wf(W)]| \\ &\leq |E[f'(W) - \sqrt{n}\Delta_{n^{-1/2}}f(W - n^{-1/2})]| + |E_1| \\ &= \sqrt{n}\left|E\left[\int_{W-n^{-1/2}}^W (f'(W) - f'(t))dt\right]\right| + |E_1| \\ &= |E_2| + |E_1|, \end{aligned} \quad (3.6)$$

where,

$$E_2 := \sqrt{n}E\left[\int_{W-n^{-1/2}}^W (f'(W) - f'(t))dt\right]. \quad (3.7)$$

If h is Lipschitz, then we know from Lemma 2.2 (ii) that f' is also Lipschitz with Lipschitz constant $2\|h'\|_\infty$ and we can bound

$$\begin{aligned} |E_2| &\leq \sqrt{n}\|f''\|_\infty E\left[\int_{W-n^{-1/2}}^W (W-t)dt\right] = \sqrt{n}\|f''\|_\infty \int_0^{n^{-1/2}} udu \\ &= \sqrt{n}\|f''\|_\infty \frac{1}{2n} = \frac{\|f''\|_\infty}{2\sqrt{n}} \leq \frac{\|h'\|_\infty}{\sqrt{n}}. \end{aligned} \quad (3.8)$$

More generally, if h is measurable and $E|h(Y)| < \infty$, then, again using that $f = f_h$ is a solution to (2.3) we conclude

$$\begin{aligned} E_2 &= \sqrt{n}E\left[\int_{W-n^{-1/2}}^W (Wf(W) - tf(t) + h(W) - h(t))dt\right] \\ &= \sqrt{n}E\left[\int_{W-n^{-1/2}}^W (Wf(W) - tf(t))dt\right] + \sqrt{n}E\left[\int_{W-n^{-1/2}}^W (h(W) - h(t))dt\right] \\ &=: E_3 + E_4. \end{aligned} \quad (3.9)$$

By the fundamental theorem of calculus we have

$$\begin{aligned}
|E_3| &= \sqrt{n} \left| E \left[\int_{W-n^{-1/2}}^W \int_t^W (f(s) + sf'(s)) ds dt \right] \right| \\
&\leq \sqrt{n} \|f\|_\infty E \left[\int_{W-n^{-1/2}}^W (W-t) dt \right] + \sqrt{n} \|f'\|_\infty E \left[\int_{W-n^{-1/2}}^W \int_t^W |s| ds dt \right] \\
&\leq \sqrt{n} \|f\|_\infty \frac{1}{2n} + \sqrt{n} \|f'\|_\infty E \left[\max(n^{-1/2}, W) \int_{W-n^{-1/2}}^W (W-t) dt \right] \\
&= \frac{\|f\|_\infty}{2\sqrt{n}} + \frac{\|f'\|_\infty}{2\sqrt{n}} E \left[\max(n^{-1/2}, W) \right] \\
&\leq \frac{\|f\|_\infty}{2\sqrt{n}} + \frac{\|f'\|_\infty}{2n} + \frac{\|f'\|_\infty}{2\sqrt{n}} \sqrt{\frac{2}{\pi}}, \tag{3.10}
\end{aligned}$$

where we have used the inequality $\max(x, y) \leq x + y$ valid for $x, y \geq 0$ and (3.3) for the last step. Note that for Lipschitz h we have $|h(W) - h(t)| \leq \|h'\|_\infty |W - t|$ and hence

$$|E_4| \leq \sqrt{n} \|h'\|_\infty E \left[\int_{W-n^{-1/2}}^W (W-t) dt \right] = \frac{\|h'\|_\infty}{2\sqrt{n}},$$

which together with (3.10) yields a worse bound than the one obtained in (3.8), if we plug in the bounds on f and f' from Lemma 2.2 (ii). If h is not Lipschitz, then $|E_4|$ cannot be bounded that easily. Having in mind the Kolmogorov distance, we restrict ourselves to the test functions $h_z = 1_{(-\infty, z]}$, $z > 0$. Note that for $h = h_z$ we can write

$$|E_4| = -E_4 = \sqrt{n} E \left[\int_{\mathbb{R}} 1_{\{W-n^{-1/2} \leq t \leq W\}} (h(t) - h(W)) dt \right]$$

and that

$$\begin{aligned}
1_{\{W-n^{-1/2} \leq t \leq W\}} (h(t) - h(W)) &= 1_{\{W-n^{-1/2} \leq t \leq W\}} 1_{\{W-n^{-1/2} \leq t \leq z < W\}} \\
&= 1_{\{W-n^{-1/2} \leq t \leq z\}} 1_{\{W-n^{-1/2} \leq z < W\}} \\
&\leq 1_{\{W-n^{-1/2} \leq t \leq W\}} 1_{\{W-n^{-1/2} \leq z < W\}}. \tag{3.11}
\end{aligned}$$

Thus, using (3.11) we obtain

$$\begin{aligned}
|E_4| &= \sqrt{n} E \left[\int_{\mathbb{R}} 1_{\{W-n^{-1/2} \leq t \leq z\}} 1_{\{W-n^{-1/2} \leq z < W\}} dt \right] \\
&\leq \sqrt{n} E \left[\int_{W-n^{-1/2}}^W 1 dt 1_{\{W-n^{-1/2} \leq z < W\}} \right] \\
&= P(W - n^{-1/2} \leq z < W) = P(K_n - 1 \leq \sqrt{n}z < K_n) \\
&= P(K_n = 1 + \lfloor \sqrt{n}z \rfloor). \tag{3.12}
\end{aligned}$$

Now note that by unimodality of binomial coefficients for each $r = 0, 1, \dots, m$ we can bound

$$\begin{aligned} P(K_n = r) &= \frac{1}{2^{2m-r}} \binom{2m-r}{m} \leq \frac{1}{2^{2m-r}} \begin{cases} \binom{2m-r}{m-r/2}, & r \text{ even} \\ \binom{2m-r}{m-(r+1)/2}, & r \text{ odd} \end{cases} \\ &\leq \begin{cases} (\pi(m-r/2))^{-1/2}, & r \text{ even} \\ (\pi(m-(r+1)/2))^{-1/2}, & r \text{ odd} \end{cases} \\ &\leq \frac{\sqrt{2}}{\sqrt{\pi m}}, \end{aligned} \quad (3.13)$$

where the next to last inequality comes from

$$2^{-2k-1} \binom{2k+1}{k} \leq 2^{-2k} \binom{2k}{k} \leq \frac{1}{\sqrt{\pi k}}$$

by Stirling's formula. Note that $P(K_n = r) = 0$ for $r > m = n/2$ and, hence, $P(K_n = 1 + \lfloor \sqrt{n}z \rfloor) = 0$ unless $z \leq 2^{-1}n^{1/2} - n^{-1/2}$. Thus, from (3.13) we have that for each $z > 0$

$$|E_4| \leq \frac{\sqrt{2}}{\sqrt{\pi m}} = \frac{2}{\sqrt{\pi n}}. \quad (3.14)$$

Collecting terms we see from (3.6), (3.5), (3.8) and Lemma 2.2 (ii) that for h Lipschitz on $[0, \sqrt{n}]$ we have

$$|E[h(W)] - E[h(Y)]| \leq \frac{\|h'\|_\infty}{\sqrt{n}} \left(\frac{1}{\sqrt{n}} \sqrt{\frac{2}{\pi}} + \frac{2}{\pi} + 2 \right) \quad (3.15)$$

and letting $h = h_z$ for $z \geq 0$ we see from (3.6), (3.5), (3.9), (3.10), (3.14) and Lemma 2.3 that

$$|P(W \leq z) - P(Y \leq z)| \leq \frac{1}{\sqrt{n}} \left(\frac{3 + 2\sqrt{2}}{\sqrt{2\pi}} + \frac{3}{4} \right) + \frac{3}{2n}. \quad (3.16)$$

Theorem 1.1 now follows from (3.15) and (3.16). \square

4. The maximum of the simple random walk

Recall the definition of M_n from (1.2), where $n = 2m$ is an even positive integer. In this section we give a proof of Theorem 1.2, which is a little bit more complicated than the one of Theorem 1.1 given in the last section. This is so because in this case the function ψ from (2.9) vanishes on the odd integers which is inconvenient for our method of proof. So we will introduce an auxiliary variable V and use the triangle inequality

$$d_{\mathcal{H}}(W, Y) \leq d_{\mathcal{H}}(W, V) + d_{\mathcal{H}}(V, Y) \quad (4.1)$$

to prove the theorem.

For $r = 0, 1, \dots, n = 2m$ let $p(r) := P(M_n = r)$. Then, it is known (see Feller (1968), Theorem 1 in Section 8 of Chapter 3) that

$$p(r) = P(M_n = r) = p_{n,r} + p_{n,r+1}, \quad (4.2)$$

where

$$p_{n,k} := P(S_n = k) = \binom{n}{\frac{n+k}{2}} 2^{-n}$$

is the probability that the random walk is in position k at time n . We use the convention that $\binom{y}{x} = 0$ unless x is a nonnegative integer. Hence, if $p_{n,k} \neq 0$, then n and k must have the same parity and since we always assume that n is even, we have $p_{n,k} = 0$ whenever k is odd.

Remark 4.1. From (4.2) one can easily get that for $z \geq 0$

$$\left| P(M_n \leq \sqrt{n}z) - (2\Phi(z) - 1) \right| = \left| 2(P(S_n \leq z) - \Phi(z)) + P(S_n = 1 + \lfloor \sqrt{n}z \rfloor) \right|, \quad (4.3)$$

which, together with the Berry-Esseen Theorem for Bernoulli random variables, yields the second part of Theorem 1.2, maybe with different constants. Using the fact that

$$d_{\mathcal{W}}(\mu, \nu) = \int_{\mathbb{R}} |F(x) - G(x)| dx,$$

where F and G are the distribution functions corresponding to the distributions μ and ν , respectively, one might also be able to derive a bound comparable to the first bound of Theorem 1.2 from (4.3) and a quantitative version of the mean CLT, see Goldstein (2010), for instance. However, we prefer giving a full proof of Theorem 1.2 which does not rely on any external quantitative CLT results but is just based on the Stein characterizations of the two distributions involved.

From (4.2) it easily follows that with ψ given by (2.9) for all $r \in \{0, 1, \dots, n\}$

$$\psi(r) := \frac{p_{n,r+2} - p_{n,r}}{p_{n,r} + p_{n,r+1}} = \begin{cases} \frac{p_{n,r+2} - p_{n,r}}{p_{n,r}}, & r \text{ even} \\ 0, & r \text{ odd}. \end{cases} \quad (4.4)$$

This is why we introduce the auxiliary random variables

$$N_n := \left\lfloor \frac{M_n + 1}{2} \right\rfloor \quad \text{and} \quad V := V_n := \frac{2N_n}{\sqrt{n}}. \quad (4.5)$$

Then, N_n only takes the values $0, 1, \dots, m$ and for $s \in \{0, 1, \dots, m\}$ we obtain from (4.2) that

$$\begin{aligned} q(s) &:= P(N_n = s) = P(M_n + 1 = 2s) + P(M_n + 1 = 2s + 1) \\ &= P(M_n = 2s - 1) + P(M_n = 2s) = 2P(M_n = 2s) = 2p(2s) \\ &= 2 \binom{2m}{m+s} 2^{-2m}. \end{aligned} \quad (4.6)$$

Denoting by ϱ the corresponding difference quotient of the probability mass function q we obtain for $s = 0, 1, \dots, m$ that

$$\begin{aligned} \varrho(s) &:= \frac{q(s+1) - q(s)}{q(s)} = \frac{2p(2s+2) - 2p(2s)}{2p(2s)} \\ &= \frac{p(2s+2) - p(2s)}{p(2s)} = \frac{\binom{2m}{m+s+1} - \binom{2m}{m+s}}{\binom{2m}{m+s}} \\ &= \frac{\binom{2m}{m+s} \left(\frac{m-s}{m+s+1} - 1 \right)}{\binom{2m}{m+s}} = \frac{m-s}{m+s+1} - \frac{m+s+1}{m+s+1} \\ &= -\frac{2s+1}{m+s+1}. \end{aligned} \quad (4.7)$$

So, we define the function $c : [-1, m] \rightarrow \mathbb{R}$ by $c(s) := m + s + 1$. Then, for $s = 0, 1, \dots, m$

$$\begin{aligned} c(s)g(s) &= -2s - 1 = -(2s + 1) \\ \Delta c(s - 1) &= c(s) - c(s - 1) = m + s + 1 - (m + s) = 1 \\ \gamma(s) &:= c(s)g(s) + \Delta c(s - 1) = -2s - 1 + 1 = -2s. \end{aligned}$$

Hence, Proposition 2.5 implies the following lemma.

Lemma 4.2. *A random variable X with values in $I := [0, m] \cap \mathbb{Z}$ has the probability mass function q , if and only if for all functions $g \in \mathcal{F}(q)$*

$$E[(m + X)\Delta g(X - 1) - 2Xg(X)] = 0.$$

Letting $g(k) := 1$ for $k \geq 0$ and $g(k) := 0$ for $k < 0$, we see from Lemma 4.2 that

$$\begin{aligned} E[N_n] &= \frac{mP(N_n = 0)}{2} = \frac{2m2^{-2m} \binom{2m}{m}}{2} = m2^{-2m} \binom{2m}{m} \\ &\leq \sqrt{\frac{m}{\pi}} = \frac{1}{\sqrt{2\pi}} \sqrt{n} \\ \Rightarrow E[V] &= \frac{2}{\sqrt{n}} E[N_n] \leq \sqrt{\frac{2}{\pi}}. \end{aligned} \tag{4.8}$$

The next lemma gives bounds on the distance from V to W .

Lemma 4.3. *For each $n = 2m$ we have*

$$d_{\mathcal{V}}(V, W) \leq \frac{1}{\sqrt{n}} \quad \text{and} \quad d_{\mathcal{K}}(V, W) \leq \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}}.$$

Proof: We have

$$\begin{aligned} \frac{M_n - 1}{2} = \frac{M_n + 1}{2} - 1 &< N_n \leq \frac{M_n + 1}{2} \\ \Rightarrow -1 &< 2N_n - M_n \leq 1 \\ \Rightarrow -\frac{1}{\sqrt{n}} &< V - W \leq \frac{1}{\sqrt{n}}. \end{aligned} \tag{4.9}$$

Hence, if h is Lipschitz on $[0, \sqrt{n}]$ with constant 1, then from (4.9) follows that

$$\left| E[h(V) - h(W)] \right| \leq \|h'\|_{\infty} E|V - W| \leq \frac{1}{\sqrt{n}},$$

which proves the first claim. As to the second claim, note that the Kolmogorov distance is scale-invariant which implies that $d_{\mathcal{K}}(V, W) = d_{\mathcal{K}}(2N_n, M_n)$ and that for $z \in [0, \infty)$

$$\left| P(2N_n \leq z) - P(M_n \leq z) \right| = \left| P(2N_n \leq \lfloor z \rfloor) - P(M_n \leq \lfloor z \rfloor) \right|.$$

Hence, we need only consider $z \in \mathbb{N}_0$. If $z = 2k$ is even, then

$$\{2N_n \leq 2k\} = \left\{ \left\lfloor \frac{M_n + 1}{2} \right\rfloor \leq k \right\} = \{M_n \leq 2k\}$$

and, thus,

$$P(2N_n \leq 2k) = P(M_n \leq 2k). \tag{4.10}$$

On the other hand, if $z = 2k + 1$, then

$$\{2N_n \leq 2k + 1\} = \{2N_n \leq 2k\} = \{M_n \leq 2k\} \subseteq \{M_n \leq 2k + 1\}$$

and

$$\{M_n \leq 2k+1\} \setminus \{2N_n \leq 2k+1\} = \{M_n = 2k+1\}.$$

Hence,

$$\begin{aligned} |P(2N_n \leq 2k+1) - P(M_n \leq 2k+1)| &= P(M_n = 2k+1) = p_{n,2k+2} \\ &= \binom{2m}{m+k+1} 2^{-2m} \leq \binom{2m}{m} 2^{-2m} \\ &\leq \frac{1}{\sqrt{\pi m}} = \frac{\sqrt{2}}{\sqrt{\pi n}}, \end{aligned} \quad (4.11)$$

by Stirling's formula. Thus, for every $z \in \mathbb{R}$ we have the bound

$$|P(2N_n \leq z) - P(M_n \leq z)| \leq \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}},$$

proving the second claim of the lemma. \square

Having bounded the distance from V to W , we may now derive bounds on the distance from V to Y in a similar way as in Section 3.

Lemma 4.4. *For the distance from V to Y we have*

$$\begin{aligned} d_{\mathcal{W}}(V, Y) &\leq \frac{1}{\sqrt{n}} \left(2 + \frac{4}{\pi}\right) + 2\sqrt{\frac{2}{\pi}} \frac{1}{n} \quad \text{and} \\ d_{\mathcal{K}}(V, Y) &\leq \frac{1}{\sqrt{n}} \left(3\sqrt{\frac{2}{\pi}} + \frac{1}{2}\right) + \frac{2}{n}. \end{aligned}$$

Proof: Let h be a Borel-measurable test function and consider the corresponding standard solution $f := f_h$ to Stein's equation (2.3) given by (2.5) and (2.6). We also let $f_h(x) := 0$ for each $x < 0$ and define $g(k) := f_h(2k/\sqrt{n})$ for $k \in \mathbb{Z}$. Writing $\Delta_y f(x) := f(x+y) - f(x)$ we obtain from Lemma 4.2

$$\begin{aligned} E[Vf(V)] &= \frac{2}{\sqrt{n}} E[N_n g(N_n)] = \frac{1}{\sqrt{n}} E[(m + N_n) \Delta g(N_n - 1)] \\ &= E\left[\left(\frac{\sqrt{n}}{2} + \frac{V}{2}\right) \Delta_{2n^{-1/2}} f(V - 2n^{-1/2})\right]. \end{aligned} \quad (4.12)$$

Thus, since f solves Stein's equation (2.3) we obtain from (4.12)

$$\begin{aligned} |E[h(V)] - E[h(Y)]| &= |E[f'(V) - Vf(V)]| \\ &= \left|E\left[f'(V) - \left(\frac{\sqrt{n}}{2} + \frac{V}{2}\right) \Delta_{2n^{-1/2}} f(V - 2n^{-1/2})\right]\right| \\ &= \left|E\left[f'(V) - \frac{\sqrt{n}}{2} \Delta_{2n^{-1/2}} f(V - 2n^{-1/2})\right] - E_1\right| \end{aligned} \quad (4.13)$$

where

$$\begin{aligned} |E_1| &= \left|E\left[\frac{V}{2} \Delta_{2n^{-1/2}} f(V - 2n^{-1/2})\right]\right| \leq \frac{2n^{-1/2}}{2} \|f'\|_{\infty} E[V] \\ &\leq \|f'\|_{\infty} \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}} \end{aligned} \quad (4.14)$$

by inequality (4.8). Similarly to the proof of Theorem 1.1 in Section 3 we have

$$\begin{aligned}
|E_2| &:= \left| E \left[f'(V) - \frac{\sqrt{n}}{2} \Delta_{2n^{-1/2}} f(V - 2n^{-1/2}) \right] \right| \\
&= \frac{\sqrt{n}}{2} \left| E \left[\int_{V - \frac{2}{\sqrt{n}}}^V (f'(V) - f'(t)) dt \right] \right| \\
&= \frac{\sqrt{n}}{2} \left| E \left[\int_{V - \frac{2}{\sqrt{n}}}^V (Vf(V) - tf(t) + h(V) - h(t)) dt \right] \right| \\
&\leq \frac{\sqrt{n}}{2} \left| E \left[\int_{V - \frac{2}{\sqrt{n}}}^V (Vf(V) - tf(t)) dt \right] \right| + \frac{\sqrt{n}}{2} \left| E \left[\int_{V - \frac{2}{\sqrt{n}}}^V (h(V) - h(t)) dt \right] \right| \\
&=: |E_3| + |E_4|. \tag{4.15}
\end{aligned}$$

As for (3.8), if h is Lipschitz than by bound (c) from Lemma 2.2 (ii) we easily get

$$|E_2| \leq \|f''\|_\infty \frac{\sqrt{n}}{2} \int_{V - \frac{2}{\sqrt{n}}}^V (V - t) dt = \frac{\sqrt{n}}{2} \|f''\|_\infty \frac{2}{n} \leq \frac{2\|h'\|_\infty}{\sqrt{n}} \tag{4.16}$$

Similar computations as those in Section 3 yield

$$\begin{aligned}
|E_3| &\leq \frac{1}{\sqrt{n}} \left(\|f\|_\infty + \|f'\|_\infty \left(\frac{2}{\sqrt{n}} + E[V] \right) \right) \\
&\leq \frac{1}{\sqrt{n}} \left(\|f\|_\infty + \|f'\|_\infty \left(\frac{2}{\sqrt{n}} + \sqrt{\frac{2}{\pi}} \right) \right) \tag{4.17}
\end{aligned}$$

and

$$|E_4| \leq \frac{\|h'\|_\infty}{\sqrt{n}} \tag{4.18}$$

for Lipschitz continuous functions h . With a similar computation as the one leading to (3.12) one can show that for $h = h_z$, where $z \geq 0$,

$$|E_4| \leq P \left(N_n = \left\lfloor \frac{\sqrt{nz}}{2} \right\rfloor + 1 \right) \leq \max_{s=0, \dots, m} q(s). \tag{4.19}$$

But from (4.6) we see that

$$\max_{s=0, \dots, m} q(s) = q(0) \leq \frac{2}{\sqrt{\pi m}} = \sqrt{\frac{2}{\pi}} \frac{2}{\sqrt{n}}, \tag{4.20}$$

again by Stirling's formula. From (4.18) and (4.20) we get that for each $z \in \mathbb{R}$

$$|E_4| \leq \sqrt{\frac{2}{\pi}} \frac{2}{\sqrt{n}}. \tag{4.21}$$

Collecting terms, we see from (4.13), (4.14), (4.16) and Lemma 2.2 (ii) that for h Lipschitz on $[0, \sqrt{n}]$ we have

$$\left| E[h(V)] - E[h(Y)] \right| \leq \|h'\|_\infty \left(\frac{1}{\sqrt{n}} \left(2 + \frac{2}{\pi} \right) \right) \tag{4.22}$$

and from (4.13), (4.14), (4.15), (4.17), (4.21) and Lemma 2.3 we see that for each $z \in \mathbb{R}$ we have the bound

$$\left| P(V \leq z) - P(Y \leq z) \right| \leq \frac{1}{\sqrt{n}} \left(3\sqrt{\frac{2}{\pi}} + \frac{1}{2} \right) + \frac{2}{n}. \tag{4.23}$$

The claims of the lemma now follow from (4.22) and (4.23), respectively. \square

Proof of Theorem 1.2: The claims of Theorem 1.2 now follow easily from Lemmas 4.3, 4.4 and from the triangle inequality (4.1). \square

5. The number of sign changes

Recall the definition of C_{2m+1} from (1.3). It is known (see Theorem 1 in Section 5 of Chapter 3 of Feller (1968), for example) that for $s = 0, 1, \dots, m$

$$p(s) := P(C_{2m+1} = s) = 2p_{2m+1, 2s+1} = 2 \binom{2m+1}{m+s+1} 2^{-2m-1}. \quad (5.1)$$

Thus, p is very similar to the probability mass function q of N_n from Section 4 (see (4.6)). This allows for a proof of Theorem 1.3, which is completely analogous to the proof of Lemma 4.4 and an analogue of Remark 4.1 is also valid, here. This is why we omit the details of the proof but just give the suitable Stein characterization of C_{2m+1} , which easily follows from (5.1) and Proposition 2.5.

Lemma 5.1. *A random variable X with values in $I := [0, m] \cap \mathbb{Z}$ has the probability mass function p , if and only if for all functions $g \in \mathcal{F}(p)$*

$$E[(m+1+X)\Delta g(X-1) - 2(X+1)g(X)] = 0.$$

6. Optimality of the rates

In this section we give an argument why the rate $n^{-1/2}$ in Theorems 1.2-1.3 is optimal for both the Kolmogorov and the Wasserstein distance. Let us explain this by means of the example of the number of returns K_n . To see that the rate is optimal for the Kolmogorov distance, it suffices to take $z = 0$ and observe that

$$|P(K_n \leq 0) - P(Y \leq 0)| = P(K_n = 0) = 2^{-2m} \binom{2m}{m} \sim \frac{1}{\sqrt{\pi m}} = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}}$$

by Stirling's formula. Hence, $n^{-1/2}$ is optimal and the best constant is no less than $\sqrt{\frac{2}{\pi}}$. To show that the rate is also optimal with respect to the Wasserstein distance, we consider the function $h(x) := x$ which is 1-Lipschitz on \mathbb{R} . By (3.3) we have

$$\begin{aligned} E[W] &= \frac{1}{\sqrt{n}} E[K_n] = \frac{1}{\sqrt{2m}} (2m+1) 2^{-2m} \binom{2m}{m} - 1 \\ &\sim \frac{\sqrt{2m}}{\sqrt{\pi m}} - \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{\pi m} \sqrt{n}} \\ &= \sqrt{\frac{2}{\pi}} - \frac{1}{\sqrt{n}} + \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{n}}. \end{aligned}$$

Since $E[Y] = \sqrt{\frac{2}{\pi}}$ this shows that

$$|E[W] - E[Y]| \sim \frac{1}{\sqrt{n}},$$

yielding the optimality of the rate $n^{-1/2}$. In a similar fashion one may prove that the convergence rates in Theorems 1.2 and 1.3 are best possible.

7. Proofs from Section 2

In this section we give proofs of some of the results from Section 2. Most of them are quite standard in Stein's method, but in many cases we obtain better constants than we would by quoting the general bounds from the literature, for example from [Chatterjee and Shao \(2011\)](#). Recall the density function p and the distribution function F of the half-normal distribution from (2.1) and (2.2), respectively. Also, in this section, we let

$$\psi(x) := \frac{d}{dx} \log p(x) = \frac{p'(x)}{p(x)} = -x, \quad x \geq 0 \quad (7.1)$$

denote the logarithmic derivative of the density function p . Note that ψ is a decreasing function. This property suffices to prove more explicit bounds on the solution to Stein's equation in the general density approach than those currently given in the literature (see [Chen et al. \(2011\)](#); [Chatterjee and Shao \(2011\)](#)).

We will several times make use of the following well-known *Mill's ratio inequality*, which is valid for $x > 0$:

$$\frac{x}{1+x^2} \varphi(x) \leq 1 - \Phi(x) \leq \frac{\varphi(x)}{x}. \quad (7.2)$$

Also, for a Borel-measurable test function $h : [0, \infty) \rightarrow \mathbb{R}$ such that $E|h(Y)| < \infty$ we let

$$\tilde{h}(x) := h(x) - E[h(Y)].$$

Proof of Lemma 2.2: From (2.5) and (2.6) one immediately gets

$$|f_h(x)| \leq \|\tilde{h}\|_\infty \min(M(x), N(x)), \quad (7.3)$$

where

$$M(x) := \frac{F(x)}{p(x)} \quad \text{and} \quad N(x) := \frac{1-F(x)}{p(x)}. \quad (7.4)$$

We have

$$M'(x) = \frac{p(x)^2 - p'(x)F(x)}{p(x)^2} = \frac{1}{p(x)} (p(x) - \psi(x)F(x)) \quad \text{and} \quad (7.5)$$

$$N'(x) = \frac{-1}{p(x)} (p(x) + \psi(x)(1-F(x))). \quad (7.6)$$

Note that

$$\begin{aligned} p(x) &= p(0) + \int_0^x p'(t) dt = p(0) + \int_0^x \psi(t)p(t) dt \\ &\geq p(0) + \psi(x) \int_0^x p(t) dt = p(0) + \psi(x)F(x) \\ &\geq \psi(x)F(x). \end{aligned}$$

Hence, from (7.5) we conclude that M is increasing on $[0, \infty)$. Similarly, one shows that N is decreasing. Since $N(0) > 0 = M(0)$ and $\lim_{x \rightarrow \infty} M(x) > \lim_{x \rightarrow \infty} N(x) = 0$ there is a unique point $x \geq 0$ such that $M(x) = N(x)$ and, clearly, at this point the function $\min(M, N)$ attains its maximum value. But

$$M(x) = N(x) \Leftrightarrow F(x) = 1 - F(x) \Leftrightarrow F(x) = \frac{1}{2}$$

and, thus, $x = m$ is the median of F . In our case of the half-normal distribution we have $F(x) = 2\Phi(x) - 1$ and so

$$m = \Phi^{-1}(3/4) =: z_{0.75}.$$

Since $p(x) = 2\varphi(x)$ and $F(m) = 1/2$ we obtain from (7.3) that

$$\|f_h\|_\infty \leq \frac{\|\tilde{h}\|_\infty}{4\varphi(z_{0.75})}, \quad (7.7)$$

proving the first claim of (i). For the second claim note that since f_h is a solution to the half-normal Stein equation (2.3)

$$|f'_h(x)| = |\tilde{h}(x) - \psi(x)f_h(x)| \leq \|\tilde{h}\|_\infty + \|\psi f_h\|_\infty. \quad (7.8)$$

Using (7.3) and (7.2) we have that

$$\begin{aligned} |\psi(x)f_h(x)| &\leq \|\tilde{h}\|_\infty |\psi(x)| \min(M(x), N(x)) \leq \|\tilde{h}\|_\infty xN(x) \\ &= \|\tilde{h}\|_\infty \frac{x(1 - \Phi(x))}{\varphi(x)} \leq \|\tilde{h}\|_\infty. \end{aligned} \quad (7.9)$$

Thus,

$$\|\psi f_h\|_\infty \leq \|\tilde{h}\|_\infty$$

and (7.8) yields the second claim of (i). \square

Proof of Lemma 2.3: For $z, x \geq 0$ we have the representation

$$\begin{aligned} f_z(x) &= \frac{F(x \wedge z) - F(x)F(z)}{p(x)} \\ &= \begin{cases} \frac{(1-F(z))F(x)}{p(x)}, & x \leq z \\ \frac{(1-F(x))F(z)}{p(x)}, & x > z \end{cases} = \begin{cases} (1-F(z))M(x), & x \leq z \\ F(z)N(x), & x > z. \end{cases} \end{aligned} \quad (7.10)$$

In particular, f_z is positive everywhere. To prove (a) note that

$$0 \leq x f_z(x) = \begin{cases} (1-F(z))xM(x), & x \leq z \\ F(z)xN(x), & x > z. \end{cases} \quad (7.11)$$

Now, using (7.6) we obtain

$$\begin{aligned} \frac{d}{dx} xN(x) &= N(x) + xN'(x) = \frac{(1-x^2)(1-\Phi(x)) - x\varphi(x)}{\varphi(x)} \\ &\geq 0 \end{aligned}$$

by (7.2). Thus, from (7.11) we see that $x f_z(x)$ is increasing on $[z, \infty)$. Similarly, using (7.5) we have

$$\begin{aligned} \frac{d}{dx} xM(x) &= M(x) + xM'(x) = \frac{(1+x^2)(\Phi(x) - \frac{1}{2}) + x\varphi(x)}{\varphi(x)} \\ &\geq 0 \end{aligned}$$

and $x f_z(x)$ is also increasing on $[0, z]$. Thus, using (7.2) again

$$0 \leq x f_z(x) \leq \lim_{y \rightarrow \infty} y f_z(y) = 2\Phi(z) - 1,$$

proving (a).

Since M is increasing and N is decreasing, we conclude that for each $z \geq 0$

$$\sup_{x \geq 0} |f_z(x)| = f_z(z) = \frac{(1 - F(z))F(z)}{p(z)} = \frac{(1 - \Phi(z))(2\Phi(z) - 1)}{\varphi(z)} =: g(z). \quad (7.12)$$

To show that $g(z) \leq \frac{1}{2}$ for each $z \geq 0$, we define

$$D_1(x) := \frac{1}{2}\varphi(x) - (1 - \Phi(x))(2\Phi(x) - 1), \quad x \geq 0$$

and need to show that $D_1 \geq 0$. We have $D_1(0) = \frac{1}{2\sqrt{2\pi}} > 0$ and $\lim_{x \rightarrow \infty} D_1(x) = 0$. Thus, it suffices to show that D_1 is a decreasing function on $[0, \infty)$. We have

$$\begin{aligned} D_1'(x) &= -\frac{1}{2}x\varphi(x) + \varphi(x)(2\Phi(x) - 1) - 2\varphi(x)(1 - \Phi(x)) \\ &= \varphi(x)\left(-\frac{x}{2} + 4\Phi(x) - 3\right) =: \varphi(x)D_2(x). \end{aligned} \quad (7.13)$$

Thus, we want to show that $D_2(x) \leq 0$ for $x \geq 0$. Note that $D_2(0) = -1$ and $\lim_{x \rightarrow \infty} D_2(x) = -\infty$. Furthermore,

$$D_2'(x) = -\frac{1}{2} + 4\varphi(x) \quad (7.14)$$

$$D_2''(x) = -4x\varphi(x) < 0, \quad x > 0. \quad (7.15)$$

From (7.15) we conclude that D_2 is strictly concave. Hence, there is a unique $x_0 \in (0, \infty)$ such that $D_2(x_0) = \sup_{x \geq 0} D_2(x)$ and $D_2'(x_0) = 0$. By (7.14) we have

$$\begin{aligned} D_2'(x_0) = 0 &\Leftrightarrow \varphi(x_0) = \frac{1}{8} \\ &\Leftrightarrow -\frac{1}{2}(\log 2 + \log \pi) - \frac{x_0^2}{2} = -3 \log 2 \\ &\Leftrightarrow x_0^2 = 5 \log 2 - \log \pi = \log\left(\frac{32}{\pi}\right) \\ &\Leftrightarrow x_0 = \sqrt{\log\left(\frac{32}{\pi}\right)} = 1.52348\dots \end{aligned} \quad (7.16)$$

Since $D_2(x_0) = -0.01701\dots < 0$, the claim of (a) follows.

To prove (c) note that since f_z solves Equation (2.3) for $h = h_z$

$$f_z'(x) = xf_z(x) + 1_{(-\infty, z]}(x) - F(z). \quad (7.17)$$

Hence, for $x \leq z$ by (7.5), (7.17), (7.2) and (a) it follows that

$$0 < f_z'(x) = 2(1 - \Phi(z)) + xf_z(x) \quad (7.18)$$

$$\leq 2(1 - \Phi(z)) + zf_z(z) \quad (7.19)$$

$$= 2(1 - \Phi(z)) + (2\Phi(z) - 1)\frac{z(1 - \Phi(z))}{\varphi(z)}$$

$$\leq 2(1 - \Phi(z)) + (2\Phi(z) - 1) = 1.$$

Similarly, for $x > z$ we have by (7.6), (7.17), (7.2) and (a) that

$$\begin{aligned}
 0 &> f'_z(x) = xf_z(x) - 2\Phi(z) + 1 \\
 &\geq zf_z(z) - 2\Phi(z) + 1 \\
 &= 1 - 2\Phi(z) + (2\Phi(z) - 1) \frac{z(1 - \Phi(z))}{\varphi(z)} \\
 &\geq 1 - 2\Phi(z) + (2\Phi(z) - 1) \frac{z^2}{1 + z^2} = -(2\Phi(z) - 1) \frac{1}{1 + z^2} \\
 &\geq 1 - 2\Phi(z) > -1,
 \end{aligned} \tag{7.20}$$

which already prove the first statement of (c). Furthermore, taking $x = 0$ and letting $z \downarrow 0$ in (7.18) shows the optimality of the first bound in (c). For the second part of (c) note that by (7.19) and (7.20) for all $x, y \geq 0$ we have

$$|f'_z(x) - f'_z(y)| \leq 2(1 - \Phi(z)) + zf_z(z) - (zf_z(z) - 2\Phi(z) + 1) = 1.$$

Optimality of this bound follows by first taking $x = 0$ and then letting $z \downarrow 0$ and $y \rightarrow \infty$.

□

Acknowledgements

The author would like to thank the anonymous referees for useful suggestions concerning the presentation of this work as well as for pointing out several typos.

References

- A.D. Barbour, L. Holst and S. Janson. *Poisson approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York (1992). ISBN 0-19-852235-5. Oxford Science Publications. [MR1163825](#).
- S. Chatterjee, J. Fulman and A. Röllin. Exponential approximation by Stein's method and spectral graph theory. *ALEA Lat. Am. J. Probab. Math. Stat.* **8**, 197–223 (2011). [MR2802856](#).
- S. Chatterjee and Q.-M. Shao. Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie-Weiss model. *Ann. Appl. Probab.* **21** (2), 464–483 (2011). [MR2807964](#).
- L.H.Y. Chen. Poisson approximation for dependent trials. *Ann. Probability* **3** (3), 534–545 (1975). [MR0428387](#).
- 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). ISBN 978-3-642-15006-7. [MR2732624](#).
- K.L. Chung and G.A. Hunt. On the zeros of $\sum_1^n \pm 1$. *Ann. of Math. (2)* **50**, 385–400 (1949). [MR0029488](#).
- C. Döbler. A rate of convergence for the arcsine law by Stein's method. *ArXiv Mathematics e-prints* (2012a). [arXiv: 1207.0533](#).
- C. Döbler. New developments in Stein's method with applications (2012b). Ph.D.-Thesis - Ruhr-Universität Bochum.
- C. Döbler. Stein's method of exchangeable pairs for the Beta distribution and generalizations. *ArXiv Mathematics e-prints* (2014). [arXiv: 1411.4477](#).

- P. Eichelsbacher and M. Löwe. Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **15**, no. 30, 962–988 (2010). [MR2659754](#).
- W. Feller. Fluctuation theory of recurrent events. *Trans. Amer. Math. Soc.* **67**, 98–119 (1949). [MR0032114](#).
- W. Feller. *An introduction to probability theory and its applications. Vol. I*. Third edition. John Wiley & Sons, Inc., New York-London-Sydney (1968). [MR0228020](#).
- J. Fulman and N. Ross. Exponential approximation and Stein's method of exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** (1), 1–13 (2013). [MR3083916](#).
- R.E. Gaunt. Variance-gamma approximation via Stein's method. *Electron. J. Probab.* **19**, no. 38, 33 (2014). [MR3194737](#).
- L. Goldstein. Bounds on the constant in the mean central limit theorem. *Ann. Probab.* **38** (4), 1672–1689 (2010). [MR2663641](#).
- Larry Goldstein and Gesine Reinert. Stein's method for the beta distribution and the Pólya-Eggenberger urn. *J. Appl. Probab.* **50** (4), 1187–1205 (2013). [MR3161381](#).
- C. Ley, G. Reinert and Y. Swan. Approximate computation of expectations: a canonical Stein operator. *ArXiv Mathematics e-prints* (2014). [arXiv: 1408.2998](#).
- H.M. Luk. *Stein's method for the Gamma distribution and related statistical applications*. ProQuest LLC, Ann Arbor, MI (1994). Thesis (Ph.D.)—University of Southern California. [MR2693204](#).
- E. Peköz and A. Röllin. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** (2), 587–608 (2011). [MR2789507](#).
- E. Peköz, A. Röllin and N. Ross. Generalized gamma approximation with rates for urns, walks and trees. *ArXiv Mathematics e-prints* (2013). [arXiv: 1309.4183](#).
- J. Pike and H. Ren. Stein's method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** (1), 571–587 (2014). [MR3283586](#).
- C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602. Univ. California Press, Berkeley, Calif. (1972). [MR0402873](#).
- C. Stein, P. Diaconis, S. Holmes and G. Reinert. 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.*, pages 1–26. Inst. Math. Statist., Beachwood, OH (2004). [MR2118600](#).