The Annals of Applied Probability
2013, Vol. 23, No. 3, 1188–1218
DOI: 10.1214/12-AAP868

DEGREE ASYMPTOTICS WITH RATES FOR PREFERENTIAL ATTACHMENT RANDOM GRAPHS[1]

BY EROL A. PEKÖZ, ADRIAN RÖLLIN AND NATHAN ROSS

Boston University, National University of Singapore and University of California, Berkeley

We provide optimal rates of convergence to the asymptotic distribution of the (properly scaled) degree of a fixed vertex in two preferential attachment random graph models. Our approach is to show that these distributions are unique fixed points of certain distributional transformations which allows us to obtain rates of convergence using a new variation of Stein's method. Despite the large literature on these models, there is surprisingly little known about the limiting distributions so we also provide some properties and new representations, including an explicit expression for the densities in terms of the confluent hypergeometric function of the second kind.

1. Introduction. Preferential attachment random graphs are random graphs that evolve by sequentially adding vertices and edges in a random way so that connections to vertices with high degree are favored. Particular versions of these models were proposed by Barabási and Albert (1999) as a mechanism to explain the appearance of the so-called power law behavior observed in some real world networks; for example, the graph derived from the world wide web by considering webpages as vertices and hyperlinks between them as edges.

Following the publication of Barabási and Albert (1999), there has been an explosion of research surrounding these (and other) random growth models. This work is largely motivated by the idea that many real world data structures can be captured in the language of networks [see Newman (2003) for a wide survey from this point of view]. However, much of this work is experimental or empirical and, by comparison, the rigorous mathematical literature on these models is less developed [see Durrett (2007) for a recent review].

For preferential attachment models, the seminal reference in the mathematics literature is Bollobás et al. (2001), in which one of the main results is a rigorous proof that the degree of a randomly chosen vertex in a particular family of preferential attachment random graph models converges to the Yule–Simon distribution. Corresponding approximation results in total variation for this and related preferential attachment models can be found in Peköz, Röllin and Ross (2012) and Ford (2009).

Received August 2011; revised April 2012.
[1]Supported in part by NUS Research Grant R-155-000-098-133.
MSC2010 subject classifications. 60C05, 60F05, 05C08.
Key words and phrases. Random graphs, preferential attachment, Stein's method, urn models.

Here we study the distribution of the degree of a fixed vertex in two preferential attachment models. In model 1 we start with a graph G_2 with two vertices labeled one and two with an edge directed from vertex two to vertex one. Given graph G_n, graph G_{n+1} is obtained by adding a vertex labeled $n+1$ and adding a single directed edge from this new vertex to a vertex labeled from the set $\{1,\dots,n\}$, where the chance that $n+1$ connects to vertex i is proportional to the degree of vertex i in G_n (here and below degree means in-degree plus out-degree). Model 2 is one studied in Bollobás et al. (2001) and allows for self-connecting edges. There, we start with a graph G_1 with a single vertex labeled one and with an edge directed from vertex one to itself. Given graph G_n, graph G_{n+1} is obtained by adding a vertex labeled $n+1$ and adding a single directed edge from this new vertex to a vertex labeled from the set $\{1,\dots,n+1\}$, where the chance that $n+1$ connects to vertex $i \in \{1,\dots,n\}$ is proportional to the degree of vertex i in G_n (a loop at a vertex contributes two to its degree) and the chance that vertex $n+1$ connects to itself is $1/(2n+1)$.

Let $W_{n,i}$ be the degree of vertex i in G_n under either of the models above. Our main result is a rate of convergence in the Kolmogorov metric (defined below) of $W_{n,i}/(\mathbb{E}W_{n,i}^2)^{1/2}$ to its distributional limit as $n \to \infty$. Although the literature on these models is large, there is surprisingly little known about these distributions. The fact that these limits exist for the first model has been shown by Móri (2005) and Backhausz (2011) and the same result for both models can be read from Janson (2006) by relation to a generalized Pólya urn, although the existing descriptions of the limits are not very explicit. A further related result from Peköz, Röllin and Ross (2012) (and the only main result there having bearing on our work here) is that for large i, the distribution of $W_{n,i}$ is approximately geometric with parameter $\sqrt{i/n}$, with the error in the approximation going to zero as $i \to \infty$. Thus it can be seen that if $i/n \to 0$ and $i \to \infty$, then the distribution of $W_{n,i}/\mathbb{E}W_{n,i}$ converges to a rate one exponential; cf. Proposition 2.5(iii) below.

The primary tool we use here to characterize the limits and obtain rates of convergence is a new distributional transformation for which the limit distributions are the unique fixed points. This transformation allows us to develop a new variation of Stein's method; we refer to Chen, Goldstein and Shao (2011), Ross (2011) and Ross and Peköz (2007) for introductions to Stein's method.

To formulate our main result we first define the family of densities

$$\kappa_s(x) = \Gamma(s)\sqrt{\frac{2}{s\pi}}\exp\left(\frac{-x^2}{2s}\right)U\left(s-1,\frac{1}{2},\frac{x^2}{2s}\right) \qquad \text{for } x>0,\ s\ge 1/2, \tag{1.1}$$

where $\Gamma(s)$ denotes the gamma function and $U(a,b,z)$ denotes the confluent hypergeometric function of the second kind (also known as the Kummer U function) [see Abramowitz and Stegun (1964), Chapter 13]. Propositions 2.3 and 2.5 below imply that κ_s is indeed a density for $s \ge 1/2$ and we denote by K_s the distribution function defined by the density κ_s. Define the Kolmogorov distance between two cumulative distribution functions P and Q as

$$d_{\mathrm{K}}(P, Q) = \sup_x \big| P(x) - Q(x) \big|.$$

THEOREM 1.1. *Let $W_{n,i}$ be the degree of vertex i in a preferential attachment graph on n vertices defined above and let $b_{n,i}^2 = \mathbb{E}W_{n,i}^2$. For model 1 with $2 \le i \le n$ and some constants $c, C > 0$ independent of n,*

$$\frac{c}{\sqrt{n}} \le d_{\mathrm{K}}\big(\mathscr{L}(W_{n,i}/b_{n,i}), K_{i-1}\big) \le \frac{C}{\sqrt{n}}.$$

For model 2 with $1 \le i \le n$ and some constants $c, C > 0$ independent of n,

$$\frac{c}{\sqrt{n}} \le d_{\mathrm{K}}\big(\mathscr{L}(W_{n,i}/b_{n,i}), K_{i-1/2}\big) \le \frac{C}{\sqrt{n}}.$$

REMARK 1.1. Using Proposition 2.5 below we see an interesting difference in the behavior of the two models. In model 1 the limit distribution for the degree of the first vertex (which by symmetry is the same as that for the second vertex) is K_1, the absolute value of a standard normal random variable, whereas in model 2 the limit distribution for the first vertex is $K_{1/2}$, the square root of an exponential random variable.

REMARK 1.2. To ease exposition we present our upper bounds as rates, but the constants are recoverable (although probably not practical especially for large i).

Theorem 1.1 will follow from a more general result derived by developing Stein's method for the distribution K_s. The key ingredient to our framework follows from observing that K_s is a fixed point of a certain distributional transformation which we will refer to as the "s-transformed double size bias" (s-TDSB) transformation, which we now describe.

Recall for a nonnegative random variable W having finite mean we say W' has the *size bias distribution* of W if

$$\mathbb{E}\{W f(W)\} = \mathbb{E}W \mathbb{E} f(W')$$

for all f such that $\mathbb{E}|W f(W)| < \infty$ [see Brown (2006) and Arratia and Goldstein (2010) for surveys and applications of size biasing]. If in addition W has finite second moment, then we will write W'' to denote a random variable having the size bias distribution of W'. Alternatively we say W'' has the *double size bias distribution* of W and it is straightforward to check that

$$\mathbb{E}\{W^2 f(W)\} = \mathbb{E}W^2 \mathbb{E} f(W''). \tag{1.2}$$

Although not used below, it is also appropriate to say that W'' has the *square bias distribution* of W since (1.2) implies that we are biasing W against its square. This terminology is used in Goldstein (2007) and Chen, Goldstein and Shao (2011) albeit under a different notation. Now, we have the following key definition.

DEFINITION 1.3. For fixed $s \geq 1/2$ let U_1 and U_2 be two independent random variables uniformly distributed on the interval $[0, 1]$, and let Y be a Bernoulli random variable with parameter $(2s)^{-1}$ independent of U_1 and U_2. Define the random variable

$$V := Y \max(U_1, U_2) + (1 - Y) \min(U_1, U_2).$$

We say that W^* has the *s-transformed double size biased* (s-TDSB) distribution of W, if

$$\mathscr{L}(W^*) = \mathscr{L}(VW''),$$

where W'', the double size bias of W, is assumed to be independent of V.

Our next result implies that the closer a distribution is to its s-TDSB transform, the closer it is to the K_s distribution. Besides the Kolmogorov metric we also consider the Wasserstein metric between two probability distribution functions P and Q, defined as

$$d_{\mathrm{W}}(P, Q) = \sup_{h:\|h'\|=1} \left| \int h(x)\, dP(x) - \int h(x)\, dQ(x) \right|.$$

THEOREM 1.2. *Let W be a nonnegative random variable with $\mathbb{E}W^2 = 1$ and let $s \geq 1$ or $s = 1/2$. Let W^* have the s-TDSB distribution of W and be defined on the same probability space as W. Then if $s \geq 1$,*

$$d_{\mathrm{W}}\big(\mathscr{L}(W), K_s\big) \leq 8s\left(s + \frac{1}{4} + \sqrt{\frac{\pi}{2}}\right)\mathbb{E}|W - W^*|, \tag{1.3}$$

and, for any $\beta \geq 0$,

$$d_{\mathrm{K}}\big(\mathscr{L}(W), K_s\big) \leq 53s\beta + 34s^{3/2}\mathbb{P}\big[|W - W^*| > \beta\big]. \tag{1.4}$$

If $s = 1/2$ then

$$d_{\mathrm{W}}\big(\mathscr{L}(W), K_{1/2}\big) \leq 2\mathbb{E}|W - W^*|,$$

and, for any $\beta \geq 0$,

$$d_{\mathrm{K}}\big(\mathscr{L}(W), K_{1/2}\big) \leq 26\beta + 8\mathbb{P}\big[|W - W^*| > \beta\big].$$

REMARK 1.4. As can easily be read from the work of Section 2 (in particular, Propositions 2.3 and 2.5), the distributions K_s can roughly be partitioned into three regions where similar behavior within the range can be expected: $s = 1/2$, $1/2 < s < 1$ and $s \geq 1$. The theorem only covers the first and last cases as this is what is needed to prove Theorem 1.1. Analogs of the results of the theorem hold in the region $1/2 < s < 1$, but we have omitted them for simplicity and brevity.

REMARK 1.5. From Lemma 3.8 below and the fact that for h with bounded derivative

$$\mathbb{E}\bigl|h(X) - h(Y)\bigr| \leq \|h'\| d_{\mathrm{W}}\bigl(\mathscr{L}(X), \mathscr{L}(Y)\bigr),$$

we see that for $s \geq 1$ or $s = 1/2$ and all $\varepsilon > 0$,

$$d_{\mathrm{K}}\bigl(\mathscr{L}(W), K_s\bigr) \leq \frac{d_{\mathrm{W}}(\mathscr{L}(W), K_s)}{\varepsilon} + \sqrt{2}\varepsilon.$$

Choosing $\varepsilon = 2^{-3/4}\sqrt{d_{\mathrm{W}}(\mathscr{L}(W), K_s)}$ yields

$$d_{\mathrm{K}}\bigl(\mathscr{L}(W), K_s\bigr) \leq 2^{1/4}\sqrt{d_{\mathrm{W}}\bigl(\mathscr{L}(W), K_s\bigr)}.$$

Thus we can obtain bounds in the Kolmogorov metric if $|W - W^*|$ is appropriately bounded with high probability or in expectation.

REMARK 1.6. It follows from Lemmas 3.1 and 3.11 below that $\mathscr{L}(W) = \mathscr{L}(W^*)$ if and only if $W \sim K_s$. In the case that $s = 1$, V is uniform on $(0, 1)$ and Proposition 2.5 below implies that K_1 is distributed as the absolute value of a standard normal random variable. Thus we obtain the interesting fact that $\mathscr{L}(W) = \mathscr{L}(UW'')$ for U uniform $(0, 1)$ and independent of W'' if and only if W is distributed as the absolute value of a standard normal variable. This fact can also be read from its analog for the standard normal distribution [Chen, Goldstein and Shao (2011), Proposition 2.3]: $\mathscr{L}(W) = \mathscr{L}(U_0|W|'')$ for U_0 uniform $(-1, 1)$ and independent of $|W|''$ if and only if W has the standard normal distribution [see also Pitman and Ross (2012)].

Although there are general formulations for developing Stein's method machinery for a given distribution [see Reinert (2005)], our framework below does not adhere to any of these directly since the characterizing operator we use is a second order differential operator [see (3.1) and (3.3) below]. For the distribution K_s, the usual first order Stein operator derived from the density approach of Reinert (2005) [following Stein (1986)] is a complicated expression involving special functions. However, by composing this more canonical operator with an appropriate first order operator, we are able to derive a second order Stein operator (see Lemma 3.3 below) which has a form that is amenable to our analysis. This strategy may be useful for other distributions which have first order operators that are difficult to handle.

The usual approach to developing Stein's method is to decide on the distribution of interest, find a corresponding Stein operator and then derive couplings from it.

The operator we use here was suggested by the s-TDSB transform which in turn arose from the discovery of a close coupling in the preferential attachment application. We believe this approach of using couplings to suggest a Stein operator is a potentially fruitful new strategy for extending Stein's method to new distributions and applications.

There have been several previous developments of Stein's method using fixed points of distributional transformations. Goldstein and Reinert (1997) develop Stein's method using the zero-bias transformation for which the normal distribution is a fixed point. Letting U be a uniform (0, 1) random variable independent of all else, Goldstein (2009) and Peköz and Röllin (2011) develop Stein's method for the exponential distribution using the fact that W and UW' have the same distribution if and only if W has an exponential distribution [Pakes and Khattree (1992) and Lyons, Pemantle and Peres (1995) also use this property]. We will show below that W and UW'' have the same distribution if and only if W is distributed as the absolute value of a standard normal random variable (see also Remark 1.6 above). In this light this paper can be viewed as extending the use of these types of distributional transformations in Stein's method.

The layout of the remainder of the article is as follows. In Section 2 we discuss various properties and alternative representations of K_s, in Section 3 we develop Stein's method for K_s and prove Theorem 1.2 and in Section 4 we prove Theorem 1.1 by constructing the coupling needed to apply Theorem 1.2 and bounding the appropriate terms.

2. The distribution K_s.

In this section we collect some facts about K_s. Recall the notation and definitions associated to the formula (1.1) for the density $\kappa_s(x)$. From Abramowitz and Stegun [(1964), Chapter 13], the Kummer U function, denoted $U(a,b,z)$, is the unique solution of the differential equation

$$z\frac{d^2U}{dz^2} + (b-z)\frac{dU}{dz} - aU = 0,$$

which satisfies (2.7) below. The following lemma collects some facts about $U(a,b,z)$; the right italic labeling of the formulas corresponds to the equation numbers from Abramowitz and Stegun [(1964), Chapter 13], and the notation $U'(a,b,z)$ refers to the derivative with respect to z.

LEMMA 2.1. *Let* $a, b, z \in \mathbb{R}$;

(2.1) *if* $z > 0$, $\quad U(a,b,z) = z^{1-b}U(1+a-b, 2-b, z),$ *(13.1.29)*

(2.2) *if* $a, z > 0$, *(13.2.5)*

$$U(a,b,z) = \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}\,dt,$$

(2.3) $U'(a,b,z) = -aU(a+1,b+1,z),$ *(13.4.21)*

$$(1+a-b)U(a,b-1,z) = (1-b)U(a,b,z) - zU'(a,b,z), \tag{2.4}$$ *(13.4.24)*

$$U(a,b,z) - U'(a,b,z) = U(a,b+1,z), \tag{2.5}$$ *(13.4.25)*

$$U(a-1,b-1,z) = (1-b+z)U(a,b,z) - zU'(a,b,z), \tag{2.6}$$ *(13.4.27)*

$$U(a,b,z) \sim z^{-a}, \qquad (z \to \infty), \tag{2.7}$$ *(13.5.2)*

$$\text{for } a > -\tfrac{1}{2}, \qquad U\big(a, \tfrac{1}{2}, 0\big) = \Gamma\big(\tfrac{1}{2}\big)/\Gamma\big(a+\tfrac{1}{2}\big). \tag{2.8}$$ *(13.5.10)*

As a direct consequence of (2.5) *we have*

$$\frac{\partial}{\partial z}\big(e^{-z}U(a,b,z)\big) = -e^{-z}U(a,b+1,z), \tag{2.9}$$

combining (2.3) *and* (2.7) *with* $a=0$ *we find*

$$U(0,b,z) = 1, \tag{2.10}$$

and using (2.1) *with* $a=-1/2$, $b=1/2$ *and* (2.10) *implies that for* $z>0$,

$$U\big(-\tfrac{1}{2}, \tfrac{1}{2}, z^2\big) = z. \tag{2.11}$$

By comparing integrands in (2.2), we also find the following fact.

LEMMA 2.2. *Let* $0 < a < a'$, $b < b'$ *and* $z > 0$. *Then*

$$\Gamma(a)U(a,b,z) > \Gamma(a')U(a',b,z) \quad \textit{and} \quad U(a,b,z) < U(a,b',z).$$

The next results provide simpler representations for K_s.

PROPOSITION 2.3. *If* X *and* Y *are two independent random variables having distributions*

$$X \sim \begin{cases} B(1, s-1), & \text{if } s > 1, \\ B(1/2, s-1/2), & \text{if } 1/2 < s \le 1, \end{cases}$$

where $B(a,b)$ *denotes the beta distribution, and*

$$Y \sim \begin{cases} \Gamma(1/2, 1), & \text{if } s > 1, \\ \operatorname{Exp}(1), & \text{if } 1/2 < s \le 1, \end{cases}$$

where $\Gamma(a,b)$ *denotes the gamma distribution and* $\operatorname{Exp}(\lambda)$ *the exponential distribution, then*

$$\sqrt{2sXY} \sim K_s.$$

PROOF. Let $s > 1$ and observe that by first conditioning on X, we can express the density of $\sqrt{2sXY}$ as

$$p_s(x) := \frac{\sqrt{2}(s-1)}{\sqrt{s\pi}} \int_0^1 \exp\left(\frac{-x^2}{2sy}\right) y^{-1/2}(1-y)^{s-2}\,dy. \tag{2.12}$$

After making the change of variable $y = 1/(1+t)$ in (2.12), we find

$$p_s(x) = \frac{\sqrt{2}(s-1)}{\sqrt{s\pi}} \int_0^\infty \exp\left(\frac{-x^2(t+1)}{2s}\right) t^{s-2}(1+t)^{1/2-s}\,dt,$$

and now using (2.2) with $a = s-1$ and $b = 1/2$ in the definition (1.1) of κ_s implies that $\kappa_s = p_s$.

Similarly, if $1/2 < s \le 1$, then we can express the density of $\sqrt{2sXY}$ as

$$q_s(x) := \frac{\Gamma(s)x}{s\sqrt{\pi}\Gamma(s-1/2)} \int_0^1 \exp\left(\frac{-x^2}{2sy}\right) y^{-3/2}(1-y)^{s-3/2}\,dy, \tag{2.13}$$

and after making the change of variable $y = 1/(1+t)$ in (2.13), we find

$$\begin{aligned} q_s(x) &= \frac{\Gamma(s)x}{s\sqrt{\pi}\Gamma(s-1/2)} \int_0^\infty \exp\left(\frac{-x^2(t+1)}{2s}\right) t^{s-3/2}(1+t)^{1-s}\,dt \\ &= \Gamma(s)\sqrt{\frac{2}{s\pi}} \exp\left(\frac{-x^2}{2s}\right) \frac{x}{\sqrt{2s}} U\left(s-\frac{1}{2}, \frac{3}{2}, \frac{x^2}{2s}\right), \end{aligned}$$

where we have used (2.2) with $a = s - 1/2$ and $b = 3/2$ in the second equality. Applying (2.1) with $a = s-1$ and $b = 1/2$ to this last expression implies $\kappa_s = q_s$. $\square$

The previous representations easily yield useful formulas for Mellin transforms.

PROPOSITION 2.4. *If $Z \sim K_s$ with $s \ge 1/2$, then for all $r > -1$,*

$$\mathbb{E}Z^r = \left(\frac{s}{2}\right)^{r/2} \frac{\Gamma(s)\Gamma(r+1)}{\Gamma(r/2+s)}. \tag{2.14}$$

PROOF. For $s > 1/2$, we use Proposition 2.3 and well-known formulas for the Mellin transforms of the beta and gamma distributions to find

$$\mathbb{E}Z^r = (2s)^{r/2} \frac{\Gamma(s)\Gamma(r/2+1)\Gamma(r/2+1/2)}{\Gamma(r/2+s)\Gamma(1/2)}. \tag{2.15}$$

An application of the gamma duplication formula yields

$$\Gamma\left(\frac{r}{2}+1\right)\Gamma\left(\frac{r}{2}+\frac{1}{2}\right) = \Gamma\left(\frac{1}{2}\right)2^{-r}\Gamma(r+1),$$

which combined with (2.15) implies (2.14) for the case $s > 1/2$.

The case $s = 1/2$ follows from Proposition 2.5(i) below which implies that if $\mathscr{L}(Y) = \mathrm{Exp}(1)$, then $Z \stackrel{\mathscr{D}}{=} \sqrt{Y}$. Now (2.15) easily follows from well-known Mellin transform formulas and thus (2.14) also follows. □

In a few special cases we can simplify and extend Proposition 2.3. Below $K_s(x)$ denotes the distribution function of K_s.

PROPOSITION 2.5. *We have the following special cases of* K_s:

$$\text{(i)} \qquad \kappa_{1/2}(x) = 2xe^{-x^2},$$

$$\text{(ii)} \qquad \kappa_1(x) = (2/\pi)^{1/2} e^{-x^2/2},$$

$$\text{(iii)} \qquad \lim_{s\to\infty} K_s(x) = 1 - e^{-\sqrt{2}x}.$$

PROOF. The identities (i) and (ii) are immediate from (2.11) and (2.10), respectively. Using Stirling's formula for the gamma function to take the limit as $s \to \infty$ for fixed r in (2.14) yields the moments of $\mathrm{Exp}(\sqrt{2})$ which proves (iii). □

REMARK 2.1. As discussed below, the preferential attachment model we study is a special case of a generalized Pólya triangular urn scheme as studied by Janson (2006). The limiting distributions in his Theorem 1.3(v) with $\alpha = 2$ and $\delta = \gamma = 1$ include K_s. In fact, Janson (2006), Example 3.1, discusses these limits, but, with the exception of the case $s = 1$, it does not appear that the decomposition of Proposition 2.3 has previously been exposed. On the other hand, up to a scaling factor, the moment formula of Janson (2006), Theorem 1.7, simplifies to that of Proposition 2.4 for K_s. The distribution K_s also appears in this urn context in Section 9 of the survey article Janson (2010).

Additionally, if $Z \sim K_s$, then $Z^2/(2s) \sim D(1, 1/2; s)$ for $s \geq 1/2$, where $D(a, b; c)$ is a *Dufresne law* as defined in Chamayou and Letac (1999). Dufresne laws are essentially a generalization of products of independent beta and gamma random variables.

We now collect one more fact about K_s, which will also prove useful in developing the Stein's method framework below.

LEMMA 2.6 (Mills ratio for K_s). *For every* $x \geq 0$ *and* $s \geq 1$,

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y)\,dy \leq \min\left\{\sqrt{\frac{\pi}{2}}, \frac{s}{x}\right\}.$$

PROOF. Using the definition (1.1) of κ_s, making the change of variable $\frac{y^2}{2s} = z$ and then applying (2.1) with $a = s - 1$ and $b = 1/2$, (2.9) with $a = s - 1/2$ and $b = 1/2$ and then (2.7) with $a = s - 1/2$, we find

$$\begin{aligned}\int_x^\infty \kappa_s(y)\,dy &= \frac{\Gamma(s)}{\sqrt{\pi}} \int_{x^2/(2s)}^\infty z^{-1/2} \exp(-z) U\left(s-1, \frac{1}{2}, z\right) dz \\ &= \frac{\Gamma(s)}{\sqrt{\pi}} \int_{x^2/(2s)}^\infty \exp(-z) U\left(s - \frac{1}{2}, \frac{3}{2}, z\right) dz \\ &= \frac{\Gamma(s)}{\sqrt{\pi}} \exp\left(\frac{-x^2}{2s}\right) U\left(s - \frac{1}{2}, \frac{1}{2}, \frac{x^2}{2s}\right),\end{aligned}$$

so that

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y)\,dy = \sqrt{\frac{s}{2}} \frac{U(s-1/2, 1/2, x^2/(2s))}{U(s-1, 1/2, x^2/(2s))}. \tag{2.16}$$

First note that by applying (2.1) with $a = s - 1$ and $b = 1/2$ in the denominator of the final expression of (2.16) we have

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y)\,dy = \frac{s}{x} \frac{U(s-1/2, 1/2, x^2/(2s))}{U(s-1/2, 3/2, x^2/(2s))} \le \frac{s}{x}, \tag{2.17}$$

where the inequality follows by Lemma 2.2.

Now applying (2.1) to (2.16) both in the denominator as before and in the numerator with $a = s - 1/2$ and $b = 1/2$, we find

$$\frac{1}{\kappa_s(x)} \int_x^\infty \kappa_s(y)\,dy = \sqrt{\frac{s}{2}} \frac{U(s, 3/2, x^2/(2s))}{U(s-1/2, 3/2, x^2/(2s))} \le \sqrt{\frac{s}{2}} \frac{\Gamma(s-1/2)}{\Gamma(s)},$$

where again the inequality follows by Lemma 2.2. Now applying Lemma 2.7 below to this last expression and combining with (2.17) yields the lemma. □

LEMMA 2.7. *If $s \ge 1$, then*

$$1 < \frac{\sqrt{s}\,\Gamma(s-1/2)}{\Gamma(s)} \le \sqrt{\pi}.$$

PROOF. Bustoz and Ismail (1986), Theorem 1, implies that

$$\frac{\sqrt{s}\,\Gamma(s-1/2)}{\Gamma(s)} \tag{2.18}$$

is a decreasing function on $(1/2, \infty)$, so that for $s \ge 1$, (2.18) is bounded above by $\sqrt{\pi}$. Moreover, Stirling's formula implies

$$\lim_{s\to\infty} \frac{\Gamma(s)}{\sqrt{s}\,\Gamma(s-1/2)} = 1.$$

□

3. Stein's method for K_s. In this section we develop Stein's method for K_s and prove Theorem 1.2.

LEMMA 3.1 (Characterizing Stein operator). *If $Z \sim K_s$ for $s \geq 1/2$, then for every twice differentiable function f with $f(0) = f'(0) = 0$ and such that $\mathbb{E}|f''(Z)|$, $\mathbb{E}|Zf'(Z)|$ and $\mathbb{E}|f(Z)|$ are finite, we have*

$$\mathbb{E}\{sf''(Z) - Zf'(Z) - 2(s-1)f(Z)\} = 0. \tag{3.1}$$

PROOF. Let $C_s := \sqrt{2}\Gamma(s)/\sqrt{s\pi}$. First note that

$$\mathbb{E}\{sf''(Z)\} = C_s \int_0^\infty sf''(x)\exp\left(\frac{-x^2}{2s}\right)U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right)dx. \tag{3.2}$$

Using (2.7) and (2.3) with $a = s-1$ and $b = 1/2$ we find that (3.2) equals

$$\begin{aligned}
&C_s \int_0^\infty f''(x) \int_x^\infty t\exp\left(\frac{-t^2}{2s}\right)\left(U\left(s-1, \frac{1}{2}, \frac{t^2}{2s}\right) + (s-1)U\left(s, \frac{3}{2}, \frac{t^2}{2s}\right)\right)dt\,dx \\
&\quad = C_s \int_0^\infty f'(t)t\exp\left(\frac{-t^2}{2s}\right)\left(U\left(s-1, \frac{1}{2}, \frac{t^2}{2s}\right) + (s-1)U\left(s, \frac{3}{2}, \frac{t^2}{2s}\right)\right)dt \\
&\quad = \mathbb{E}\{Zf'(Z)\} + C_s \int_0^\infty f'(t)\cdot(s-1)t\exp\left(\frac{-t^2}{2s}\right)U\left(s, \frac{3}{2}, \frac{t^2}{2s}\right)dt,
\end{aligned}$$

where in the first equality we have used Fubini's theorem [justified by $\mathbb{E}|f''(Z)| < \infty$] and the fact that $f'(0) = 0$.

We also have

$$\begin{aligned}
&C_s \int_0^\infty f'(t)\cdot t\exp\left(\frac{-t^2}{2s}\right)(s-1)U\left(s, \frac{3}{2}, \frac{t^2}{2s}\right)dt \\
&\quad = C_s \int_0^\infty f'(t)\int_t^\infty 2(s-1)\exp\left(\frac{-x^2}{2s}\right) \\
&\qquad\qquad \times \left(\left(-\frac{1}{2} + \frac{x^2}{2s}\right)U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right) - \frac{x^2}{2s}U'\left(s, \frac{3}{2}, \frac{x^2}{2s}\right)\right)dx\,dt \\
&\quad = C_s \int_0^\infty f(x)\cdot 2(s-1)\exp\left(\frac{-x^2}{2s}\right) \\
&\qquad\qquad \times \left(\left(-\frac{1}{2} + \frac{x^2}{2s}\right)U\left(s, \frac{3}{2}, \frac{x^2}{2s}\right) - \frac{x^2}{2s}U'\left(s, \frac{3}{2}, \frac{x^2}{2s}\right)\right)dx \\
&\quad = C_s \int_0^\infty f(x)\cdot 2(s-1)\exp\left(\frac{-x^2}{2s}\right)U\left(s-1, \frac{1}{2}, \frac{x^2}{2s}\right)dx \\
&\quad = \mathbb{E}\{2(s-1)f(Z)\},
\end{aligned}$$

where in the second equality we have used Fubini's theorem [justified by $\mathbb{E}|Zf'(Z)| < \infty$] and the fact that $f(0) = 0$, and in the third we have used (2.6) with $a = s$ and $b = 3/2$. Hence,

$$\mathbb{E}\{sf''(Z)\} = \mathbb{E}\{Zf'(Z)\} + \mathbb{E}\{2(s-1)f(Z)\},$$

which proves the claim. $\square$

For the sake of brevity, let $V_s(x) := U(s-1, \frac{1}{2}, \frac{x^2}{2s})$.

LEMMA 3.2. *For all functions h such that $\mathbb{E}h(Z)$ exists, the second order differential equation*

$$sf''(x) - xf'(x) - 2(s-1)f(x) = h(x) - \mathbb{E}h(Z) \tag{3.3}$$

with initial conditions $f(0) = f'(0) = 0$ has solution

$$\begin{aligned} f(x) &= \frac{1}{s} V_s(x) \int_0^x \frac{1}{V_s(y)\kappa_s(y)} \int_0^y \tilde{h}(z)\kappa_s(z)\,dz\,dy \\ &= -\frac{1}{s} V_s(x) \int_0^x \frac{1}{V_s(y)\kappa_s(y)} \int_y^\infty \tilde{h}(z)\kappa_s(z)\,dz\,dy, \end{aligned} \tag{3.4}$$

where $\tilde{h} = h - \mathbb{E}h(Z)$.

In order to prove Lemma 3.2, we use the following intermediate result.

LEMMA 3.3. *If g and f are functions such that $g(0) = f(0) = 0$ and for $x > 0$,*

$$sg'(x) - s\left(\frac{x}{s} - d(x)\right)g(x) = \tilde{h}(x), \qquad f'(x) - d(x)f(x) = g(x), \tag{3.5}$$

where

$$d(x) = \frac{\partial}{\partial x} \log V_s(x) = \frac{V_s'(x)}{V_s(x)}, \tag{3.6}$$

then f solves (3.3) and $f'(0) = 0$.

Conversely, if f is a solution to (3.3) with $f(0) = f'(0) = 0$ and $g(x) = f'(x) - d(x)f(x)$, then $g(0) = 0$ and f and g satisfy (3.5).

PROOF. Assume f and g satisfy (3.5) and $f(0) = g(0) = 0$. The fact that $f'(0) = 0$ follows easily from the second equation of (3.5). To show that (3.5) yields a solution to (3.3), differentiate the second equality in (3.5) and combine the resulting equations to obtain

$$sf''(x) - xf'(x) - \big(sd'(x) + sd(x)^2 - xd(x)\big)f(x) = \tilde{h}(x).$$

Hence, we only need to show that

$$sd'(x) + sd(x)^2 - xd(x) = 2(s-1). \tag{3.7}$$

In order to simplify the calculations, let us introduce

$$D(z) = \frac{\partial}{\partial z} \log U\left(s-1, \frac{1}{2}, z\right) = \frac{U'(s-1, 1/2, z)}{U(s-1, 1/2, z)};$$

note that $d(x) = \frac{x}{s} D(\frac{x^2}{2s})$. With this and $z = \frac{x^2}{2s}$, (3.7) becomes

$$\left(\tfrac{1}{2} - z\right) D(z) + zD'(z) + zD(z)^2 = s - 1. \tag{3.8}$$

The left-hand side of (3.8) is equal to

$$\begin{aligned} &\frac{(1/2 - z)U'(s-1, 1/2, z) + zU''(s-1, 1/2, z)}{U(s-1, 1/2, z)} \\ &\quad = (s-1)\frac{(-1/2 + z)U(s, 3/2, z) - zU'(s, 3/2, z)}{U(s-1, 1/2, z)} = s - 1, \end{aligned}$$

where we have used (2.3) with $a = s - 1$ and $b = 1/2$ to handle the derivatives in the first equality and then (2.6) with $a = s$ and $b = 3/2$ in the second. Hence, (3.7) holds, as desired.

If f is a solution to (3.3) with $f(0) = f'(0) = 0$ and $g(x) = f'(x) - d(x)f(x)$, then obviously $g(0) = 0$ and the second assertion of the lemma follows from the previous calculations. □

PROOF OF LEMMA 3.2. Lemma 3.3 implies that we only need to solve (3.5). Note first that the general differential equation

$$F'(x) - A'(x)F(x) = H(x), \qquad x > 0, F(0) = 0,$$

has solution

$$F(x) = e^{A(x)} \int_0^x H(z) e^{-A(z)}\, dz.$$

Hence, noticing that

$$\frac{x}{s} - d(x) = -\frac{\partial}{\partial x} \log \kappa_s(x),$$

the solution to the first equation in (3.5) is

$$g(y) = \frac{1}{\kappa_s(y)} \int_0^y \frac{\tilde{h}(z)}{s} \kappa_s(z)\, dz, \tag{3.9}$$

whereas the solution to the second equation in (3.5) is

$$f(x) = V_s(x) \int_0^x \frac{g(y)}{V_s(y)}\, dy,$$

which is the first identity of (3.4); the second follows by observing that $\int_0^\infty \tilde{h}(x)\kappa_s(x)\,dx = 0$. □

Before developing the Stein's method machinery further we need two more lemmas, the first of which is well known and easily read from Gordon (1941).

LEMMA 3.4 (Gaussian Mills ratio). *For* $x, s > 0$,

$$\exp\left(\frac{x^2}{2s}\right)\int_x^\infty \exp\left(\frac{-t^2}{2s}\right)dt \le \min\left\{\sqrt{\frac{s\pi}{2}}, \frac{s}{x}\right\}.$$

LEMMA 3.5. *If* $d(x)$ *is defined by* (3.6), *then for* $s \ge 1$ *and* $x > 0$

$$0 \le -d(x) \le \frac{\sqrt{2}\Gamma(s)}{\sqrt{s}\Gamma(s-1/2)} < \sqrt{2},$$
$$0 \le -xd(x) \le 2(s-1).$$

PROOF. To prove the first assertion note that (2.3) with $a = s-1$ and $b = 1/2$ followed by (2.1) with $a = s - 1/2$ and $b = 1/2$ and Lemma 2.2 imply

$$\begin{aligned}
-d(x) &= -\frac{x}{s}\frac{U'(s-1, 1/2, x^2/(2s))}{U(s-1, 1/2, x^2/(2s))} \\
&= \frac{\sqrt{2}(s-1)}{\sqrt{s}}\frac{U(s-1/2, 1/2, x^2/(2s))}{U(s-1, 1/2, x^2/(2s))} \qquad (3.10)\\
&\le \frac{\sqrt{2}(s-1)\Gamma(s-1)}{\sqrt{s}\Gamma(s-1/2)}.
\end{aligned}$$

The claimed upper bound now follows from Lemma 2.7. The lower bound follows from the final expression of (3.10), since for $s > 1$, the integral representation (2.2) implies all terms in the quotient are nonnegative, and for $s = 1$, (2.10) implies $d(x) = 0$.

For the second assertion, we use (2.5) with $a = s - 1$ and $b = 1/2$ in the second equality below to find

$$\begin{aligned}
-xd(x) &= -\frac{x^2}{s}\frac{U'(s-1, 1/2, x^2/(2s))}{U(s-1, 1/2, x^2/(2s))} \qquad (3.11)\\
&= 2\left(s - \frac{1}{2}\right)\frac{U(s-1, -1/2, x^2/(2s))}{U(s-1, 1/2, x^2/(2s))} - 1.
\end{aligned}$$

Applying Lemma 2.2 to (3.11) proves the remaining upper bound. The second lower bound follows from the first. □

LEMMA 3.6. *If g satisfies the first equation of* (3.5) *with* $g(0)=0$, *then*

$$g(x)=\frac{1}{s\kappa_s(x)}\int_0^x \kappa_s(y)\tilde{h}(y)\,dy=-\frac{1}{s\kappa_s(x)}\int_x^\infty \kappa_s(y)\tilde{h}(y)\,dy.$$

- *If h is nonnegative and bounded*, *then for all* $x>0$ *and* $s\geq 1$,

$$\left|g(x)\right|\leq \|h\|\min\left\{\frac{1}{s}\sqrt{\frac{\pi}{2}},\frac{1}{x}\right\}. \tag{3.12}$$

- *If h is absolutely continuous with bounded derivative*, *then for all* $s\geq 1$

$$\|g\|\leq \|h'\|\left(1+\frac{1}{s}\sqrt{\frac{\pi}{2}}\right). \tag{3.13}$$

PROOF. The first assertion is a restatement of (3.9), recorded in this lemma for convenient future reference.

If $h(x)\geq 0$ for $x\geq 0$ with $\|h\|<\infty$, then for all $s\geq 1$ and $x>0$,

$$\left|g(x)\right|\leq \frac{\|\tilde{h}\|}{s\kappa_s(x)}\int_x^\infty \kappa_s(y)\,dy\leq \min\left\{\sqrt{\frac{\pi}{2}},\frac{s}{x}\right\}\frac{\|h\|}{s},$$

where we have used Lemma 2.6; this shows (3.12).

Let h be absolutely continuous with $\|h'\|<\infty$, and without loss of generality assume that $h(0)=0$ so that for $x\geq 0$, $|h(x)|\leq \|h'\|x$. In particular, if $Z_s\sim K_s$, then $\tilde{h}(x)\leq (x+\mathbb{E}Z_s)\|h'\|$ and noting that $\mathbb{E}Z_s\leq \sqrt{\mathbb{E}Z_s^2}=1$ (using Proposition 2.4), we can apply Lemma 2.6 to find that for $x>0$,

$$\left|g(x)\right|\leq \frac{\|h'\|}{s\kappa_s(x)}\int_x^\infty (y+1)\kappa_s(y)\,dy\leq \frac{\|h'\|}{s}\left(\frac{\int_x^\infty y\kappa_s(y)\,dy}{\kappa_s(x)}+\sqrt{\frac{\pi}{2}}\right).$$

To bound the integral in this last expression, we make the change of variable $\frac{y^2}{2s}=z$ and apply (2.9) with $a=s-1$, $b=-1/2$ and (2.7) with $a=s-1$ to find

$$\begin{aligned}\frac{\int_x^\infty y\kappa_s(y)\,dy}{\kappa_s(x)}&=\frac{s}{\kappa_s(x)}\int_{x^2/(2s)}^\infty e^{-z}U\left(s-1,\frac{1}{2},z\right)dz\\&=s\frac{U(s-1,-1/2,x^2/(2s))}{U(s-1,1/2,x^2/(2s))}\leq s,\end{aligned}$$

where the last inequality follows from Lemma 2.2. $\square$

LEMMA 3.7. *Let f be defined as in* (3.4) *with* $f(0)=f'(0)=0$.

- *If h is nonnegative and bounded and* $s\geq 1$, *then*

$$\|f'\|\leq \sqrt{2\pi}\|h\|. \tag{3.14}$$

- *If h is nonnegative, bounded and absolutely continuous with bounded derivative and $s \geq 1$, then*

$$\|f''\| \leq 2\left(\pi\sqrt{s} + \frac{1}{s}\right)\|h\|. \tag{3.15}$$

If $s = 1/2$, then

$$\|f''\| \leq 4\|h\|. \tag{3.16}$$

- *If h is absolutely continuous with bounded derivative and $s \geq 1$, then*

$$\|f'''\| \leq 8\left(s + \frac{1}{4} + \sqrt{\frac{\pi}{2}}\right)\|h'\|. \tag{3.17}$$

If $s = 1/2$, then

$$\|f'''\| \leq 4\|h'\|. \tag{3.18}$$

PROOF. From (3.4) of Lemma 3.2 we have that

$$f(x) = V_s(x)\int_0^x \frac{g(y)}{V_s(y)}\,dy,$$

where g is as in Lemma 3.6. If either h is bounded or absolutely continuous with bounded derivative, then recall that Lemma 3.6 implies g is bounded. If $s \geq 1$, then (2.3) and (2.7) with $a = s-1$ and $b = 1/2$ imply that $V_s(x) = U(s-1, \frac{1}{2}, \frac{x^2}{2s})$ is nonincreasing and positive for positive x, so that

$$|f(x)| \leq x\|g\|. \tag{3.19}$$

Now, again by (3.5), we have

$$|f'(x)| \leq |d(x)f(x)| + \|g\| \leq \|g\|\left(|xd(x)| + 1\right) \leq \|g\|(2s-1), \tag{3.20}$$

where we have used (3.19) in the first inequality and Lemma 3.5 in the second. Applying the bound (3.12) proves (3.14).

To bound f'' for h having $\|h'\| < \infty$, let $s \geq 1/2$ and differentiate (3.3) to find

$$f'''(x) - \frac{x}{s}f''(x) = \frac{2s-1}{s}f'(x) + \frac{h'(x)}{s}, \tag{3.21}$$

which implies

$$\frac{d}{dx}\left(\exp\left(\frac{-x^2}{2s}\right)f''(x)\right) = \exp\left(\frac{-x^2}{2s}\right)\left(\frac{2s-1}{s}f'(x) + \frac{h'(x)}{s}\right).$$

Integrating, we obtain

$$\exp\left(\frac{-x^2}{2s}\right)f''(x) = -\int_x^\infty \exp\left(\frac{-y^2}{2s}\right)\left(\frac{2s-1}{s}f'(y) + \frac{h'(y)}{s}\right)dy,$$

so that Lemma 3.4 yields

$$|f''(x)| \le (2s-1)\|f'\| \min\left\{\sqrt{\frac{\pi}{2s}}, \frac{1}{x}\right\} + \frac{1}{s}\exp\left(\frac{x^2}{2s}\right)\int_x^\infty \exp\left(\frac{-y^2}{2s}\right)h'(y)\,dy. \tag{3.22}$$

If $\|h\| < \infty$, then an integration by parts yields a bound on the second term of (3.22) which yields

$$|f''(x)| \le (2s-1)\|f'\| \min\left\{\sqrt{\frac{\pi}{2s}}, \frac{1}{x}\right\} + \frac{2\|h\|}{s}.$$

If $s \ge 1$, then apply the bound (3.14) above on $\|f'\|$ to find (3.15); for $s = 1/2$, (3.16) follows immediately. Now, we can apply Lemma 3.4 directly to (3.22) to find

$$|f''(x)| \le \big((2s-1)\|f'\| + \|h'\|\big) \min\left\{\sqrt{\frac{\pi}{2s}}, \frac{1}{x}\right\}. \tag{3.23}$$

Finally, (3.21) implies

$$s|f'''(x)| \le |xf''(x)| + (2s-1)\|f'\| + \|h'\|; \tag{3.24}$$

the first term can be bounded by (3.23), and if $s \ge 1$, a subsequent application of (3.20) on $\|f'\|$ and then (3.13) on $\|g\|$ yields (3.17). If $s = 1/2$, then (3.18) follows from (3.24) and (3.23). □

In order to obtain the bounds for the Kolmogorov metric, we need to introduce the smoothed half-line indicator function

$$h_{a,\varepsilon}(x) = \frac{1}{\varepsilon}\int_0^\varepsilon \mathrm{I}[x \le a + t]\,dt. \tag{3.25}$$

LEMMA 3.8. *If $Z \sim K_s$ and W is a nonnegative random variable and $s \ge 1$, then, for all $\varepsilon > 0$,*

$$d_{\mathrm{K}}\big(\mathscr{L}(W), K_s\big) \le \sup_{a \ge 0}\big|\mathbb{E}h_{a,\varepsilon}(W) - \mathbb{E}h_{a,\varepsilon}(Z)\big| + \varepsilon\sqrt{2}.$$

If $s = 1/2$, then, for all $\varepsilon > 0$,

$$d_{\mathrm{K}}\big(\mathscr{L}(W), K_{1/2}\big) \le \sup_{a \ge 0}\big|\mathbb{E}h_{a,\varepsilon}(W) - \mathbb{E}h_{a,\varepsilon}(Z)\big| + \varepsilon\sqrt{2/e}.$$

PROOF. The lemma follows from a well-known argument and the following bounds on the density $\kappa_s(x)$ defined by (1.1). If $s \ge 1$, then by (2.3) with $a = s-1$ and (2.2) with $a = s$, $\kappa_s(x)$ is nonincreasing in x and from (2.8) with $a = s-1$,

$$\kappa_s(0) = \frac{\Gamma(s)\sqrt{2}}{\Gamma(s-1/2)\sqrt{s}} \le \sqrt{2},$$

where the inequality is by Lemma 2.7. If $s = 1/2$, then $\kappa_s(x) = 2xe^{-x^2}$ which has maximum $\sqrt{2/e}$. $\square$

We will also need the following "indirect" concentration inequality; it follows from the arguments of the proof of Lemma 3.8 immediately above.

LEMMA 3.9. *If $Z \sim K_s$ and W is a nonnegative random variable and $s \geq 1$, then, for all $0 \leq a < b$,*

$$\mathbb{P}(a < W \leq b) \leq \sqrt{2}(b-a) + 2d_{\mathrm{K}}\bigl(\mathscr{L}(W), K_s\bigr).$$

If $s = 1/2$, then, for all $0 \leq a < b$,

$$\mathbb{P}(a < W \leq b) \leq \sqrt{2/e}(b-a) + 2d_{\mathrm{K}}\bigl(\mathscr{L}(W), K_s\bigr).$$

LEMMA 3.10. *If f satisfies (3.3) for $h_{a,\varepsilon}$ and $s \geq 1$, then for $x \geq 0$,*

$$s\bigl|f''(x+t) - f''(x)\bigr| \leq |t|\bigl(2x(\pi\sqrt{s}+1) + (2s-1)\sqrt{2\pi}\bigr) + \frac{1}{\varepsilon}\int_{t\wedge 0}^{t\vee 0} \mathrm{I}[a < x+u \leq a+\varepsilon]\,du.$$

If $s = 1/2$, then for $x \geq 0$,

$$\frac{1}{2}\bigl|f''(x+t) - f''(x)\bigr| \leq 4|t|x + \frac{1}{\varepsilon}\int_{t\wedge 0}^{t\vee 0} \mathrm{I}[a < x+u \leq a+\varepsilon]\,du.$$

PROOF. Using (3.3), we obtain

$$s\bigl(f''(x+t) - f''(x)\bigr) = x\bigl(f'(x+t) - f'(x)\bigr) + tf'(x+t) + 2(s-1)\bigl(f(x+t) - f(x)\bigr) + h_{a,\varepsilon}(x+t) - h_{a,\varepsilon}(x),$$

hence,

$$s\bigl|f''(x+t) - f''(x)\bigr| \leq |t|\bigl(x\|f''\| + \|f'\| + 2(s-1)\|f'\|\bigr) + \frac{1}{\varepsilon}\int_{t\wedge 0}^{t\vee 0} \mathrm{I}[a < x+u \leq a+\varepsilon]\,du.$$

Applying the bounds of Lemma 3.7 yields the claim. $\square$

LEMMA 3.11. *Let W be a nonnegative random variable with $\mathbb{E}W^2 = 1$ and let W^* be the s-TDSB of W as in Definition 1.3 for some $s \geq 1/2$. For every twice differentiable function f with $f(0) = f'(0) = 0$ and such that the expectations below are well defined, we have*

$$s\mathbb{E}f''(W^*) = \mathbb{E}\bigl\{Wf'(W) + 2(s-1)f(W)\bigr\}.$$

PROOF. The lemma will follow from two facts:

- If W'' has the double size bias distribution of W, then for all g with $\mathbb{E}|W^2 g(W)| < \infty$,

$$\mathbb{E}g(W'') = \mathbb{E}\{W^2 g(W)\}.$$

- If g is a function such that $g'(0) = g(0) = 0$ and for V as defined in Definition 1.3, $\mathbb{E}|g''(V)| < \infty$, then

$$s\mathbb{E}g''(V) = g'(1) + 2(s-1)g(1).$$

The first item above is easy to verify from the definition of the size bias distribution and the fact that $\mathbb{E}W^2 = 1$, and the second follows from a simple calculation after noting that V has density $(2 - \frac{1}{s}) - 2x(1 - \frac{1}{s})$ for $0 < x < 1$.

By conditioning on W'' and using the second fact above for $g(t) = f(tW'')/(W'')^2$, we find

$$s\mathbb{E}f''(W^*) = \mathbb{E}\left\{\frac{f'(W'')}{W''} + 2(s-1)\frac{f(W'')}{(W'')^2}\right\},$$

and applying the first fact above proves the lemma. □

PROOF OF WASSERSTEIN BOUND OF THEOREM 1.2. Making use of Lemma 3.2 and Lemma 3.11, we have

$$\begin{aligned}\mathbb{E}h(W) - \mathbb{E}h(Z) &= \mathbb{E}\{sf''(W) - Wf'(W) - 2(s-1)f(W)\}\\ &= s\mathbb{E}\{f''(W) - f''(W^*)\},\end{aligned}$$

where f is given by (3.4). If h is Lipschitz continuous, then f is three times differentiable almost everywhere and we have

$$|\mathbb{E}h(W) - \mathbb{E}h(Z)| \le s\|f'''\|\mathbb{E}|W - W^*|.$$

We now obtain (1.3) by invoking (3.17) and (3.18) of Lemma 3.7. □

PROOF OF KOLMOGOROV BOUND OF THEOREM 1.2. Fix $a > 0$ and let $\varepsilon > 0$, to be chosen later. Let f be as in (3.4) with $\tilde{h}$ replaced by $h_{a,\varepsilon} - \mathbb{E}h_{a,\varepsilon}(Z)$, where $h_{a,\varepsilon}$ is defined by (3.25). Define the indicator random variable $J = \mathrm{I}[|W - W^*| \le \beta]$. Now,

$$\begin{aligned}&\mathbb{E}h_{a,\varepsilon}(W) - \mathbb{E}h_{a,\varepsilon}(Z)\\ &\quad = s\mathbb{E}\{f''(W) - f''(W^*)\}\\ &\quad = s\mathbb{E}\{J(f''(W) - f''(W^*))\} + s\mathbb{E}\{(1-J)(f''(W) - f''(W^*))\}\\ &\quad =: R_1 + R_2.\end{aligned}$$

If $s \ge 1$, using (3.15) from Lemma 3.7 implies

$$|R_2| \le 4(\pi s^{3/2} + 1)\mathbb{P}(|W - W^*| > \beta) \le 17 s^{3/2}\mathbb{P}(|W - W^*| > \beta).$$

Applying Lemma 3.10,

$$|R_1| \le \beta\big(2\mathbb{E}W(\pi\sqrt{s}+1)+(2s-1)\sqrt{2\pi}\big)+\frac{1}{\varepsilon}\int_{-\beta}^{\beta}\mathbb{P}(a < W+u \le a+\varepsilon)\,du.$$

Noticing that $\mathbb{E}W \le 1$ and applying Lemma 3.9 to the integrand,

$$|R_1| \le 12s\beta + 2\beta\varepsilon^{-1}(\sqrt{2}\varepsilon + 2\delta) \le 15s\beta + 4\beta\varepsilon^{-1}\delta,$$

where $\delta = d_{\mathrm{K}}(\mathscr{L}(W), K_s)$.

From Lemma 3.8, we have

$$\delta \le \sqrt{2}\varepsilon + 15s\beta + 4\beta\varepsilon^{-1}\delta + 17s^{3/2}\mathbb{P}\big(|W - W^*| > \beta\big).$$

Choosing $\varepsilon = 8\beta$ and solving for δ,

$$\delta \le 16\sqrt{2}s\beta + 30s\beta + 34s^{3/2}\mathbb{P}\big(|W - W^*| > \beta\big),$$

which yields (1.4).

A nearly identical argument yields the statement for $s = 1/2$. □

4. Proof of Theorem 1.1. We first reformulate Theorem 1.1 in terms of a generalized Pólya urn model. An urn initially contains i black balls and j white balls and at each step a ball is drawn. If the ball drawn is black, it is returned to the urn along with an additional α black balls and β white balls; if the ball drawn is white, the ball is returned to the urn along with an additional γ black balls and δ white balls. We use the notation $(\alpha, \beta; \gamma, \delta)^n_{i,j}$ to denote the distribution of the number of white balls in the urn after n draws and replacements. For example, $(\alpha, \beta; \gamma, \delta)^0_{i,j}$ has a single point mass at j and also note that $(1, 0; 0, 1)^n_{i,j}$ corresponds to the classical Pólya urn model.

THEOREM 1.1(a). *Let $n \ge 1$ and $i \ge 0$ be integers and $\mathscr{L}(W_{n,i}) = (2, 0; 1, 1)^n_{i,1}$. If $b^2_{n,i} = \mathbb{E}W^2_{n,i}$, then, for some constants $c, C > 0$ independent of n,*

$$\frac{c}{\sqrt{n}} \le d_{\mathrm{K}}\big(\mathscr{L}(W_{n,i}/b_{n,i}), K_{(i+1)/2}\big) \le \frac{C}{\sqrt{n}}.$$

Theorem 1.1 follows immediately from Theorem 1.1(a) after noting that for model 1 with $n \ge i \ge 2$, the degree of vertex i in G_n, the graph with n vertices and $n-1$ edges, has distribution $(2, 0; 1, 1)^{n-i}_{2i-3,1}$; this is because the degree of vertex i in G_i is 1 and the sum of the degrees of the remaining vertices is $2i-3$ (since G_i has $i-1$ edges). For model 2 with $n \ge i \ge 1$, the degree of vertex i in G_n, the graph with n vertices and n edges, has distribution $(2, 0; 1, 1)^{n-i+1}_{2i-2,1}$; this is because the sum of the degrees of G_{i-1} is $2i-2$ and vertex i has probability $1/(2i-1)$ of self-attachment when forming G_i from G_{i-1}.

The lower bound of the theorem follows from the following general result combined with the fact from Lemma 4.7 below that $\mathbb{E}W^2_{n,i} \le 2(1+2n)$.

LEMMA 4.1. *Let μ be a probability distribution with a density f such that for all x in some interval (a, b), $f(x) > \varepsilon > 0$. If $(X_n)_{n\ge 1}$ is a sequence of integer-valued random variables and $(a_n)_{n\ge 1}$ is a sequence of nonnegative numbers tending to zero, then*

$$d_{\mathrm{K}}\big(\mathscr{L}(a_n X_n), \mu\big) \ge c a_n$$

for some positive constant c independent of n.

PROOF. Let F be the distribution function of μ and note that the hypothesis on the density f implies that if $a \le x < y \le b$, then

$$F(y) - F(x) \ge \varepsilon(y - x). \tag{4.1}$$

Since $\lim_n a_n = 0$, there exists N such that for all $n \ge N$, there is an integer k_n such that $[a_n k_n, a_n(k_n + 1)] \subset (a, b)$. From (4.1), for $n \ge N$ we have

$$F\big(a_n(k_n + 1)\big) - F(a_n k_n) \ge a_n \varepsilon,$$

and now using the continuity of F on (a, b) and the fact that the distribution function G_n of $a_n X_n$ is constant on the interval $I_n := [a_n k_n, a_n(k_n + 1))$, it follows that for $n \ge N$,

$$d_{\mathrm{K}}\big(\mathscr{L}(a_n X_n), \mu\big) \ge \sup_{x\in I_n}\big|G_n(x) - F(x)\big| \ge \frac{\varepsilon}{2} a_n.$$

Since G_n is the distribution function of a discrete random variable and F is continuous, it follows that $d_{\mathrm{K}}(\mathscr{L}(a_n X_n), \mu) > 0$ for all $n \in \mathbb{N}$ (and in particular $n < N$), so that we may choose $c > 0$. $\square$

REMARK 4.1. As mentioned in the Introduction we write our results as rates, but the constants are recoverable. For the sake of clarity, we have not been careful to optimize the bounds in our arguments, but it is clear that sharper statements can be read from the proofs below. For example, the constant in both the lower bound and upper bounds of Theorem 1.1(a) depend crucially on the scaling factor $\mathbb{E}W_{n,i}^2$. For our purposes Lemma 4.7 below is acceptable, but note that exact results are available [see (4.23) and (4.24) in the proof of Lemma 4.7].

Now let $W := W_{n,i}$ have distribution $(2, 0; 1, 1)_{i,1}^n$. We will use (1.4) to prove the upper bounds of Theorem 1.1(a) and so we will show that there is a close coupling of W and VW'', where V is as in Definition 1.3 with $s = (i + 1)/2$. This result will follow from the following lemmas proved at the end of this section.

LEMMA 4.2. *There is a coupling (R, W'') of $(2, 0; 1, 1)_{i,3}^{n-1}$ and the double size bias distribution of $(2, 0; 1, 1)_{i,1}^n$ satisfying $\mathbb{P}(R \ne W'') \le C/\sqrt{n}$.*

LEMMA 4.3. *The distribution* $(2, 0; 1, 1)_{i,1}^{n}$ *can be expressed as a mixture of the distributions* $(2, 0; 1, 1)_{i+1,2}^{n-1}$ *and* $(2, 0; 1, 1)_{i+2,1}^{n-1}$ *with respective probabilities* $1/(1+i)$ *and* $1 - 1/(1+i)$.

In the next lemma we use the notation $(\alpha, \beta; \gamma, \delta)_{i,j}^{N}$ for a nonnegative integer-valued random variable N to denote a mixture of the distributions $(\alpha, \beta; \gamma, \delta)_{i,j}^{n}$ for $n = 0, 1, 2, \dots$ that are mixed with respective probabilities $\mathbb{P}(N = n)$ for $n = 0, 1, 2, \dots$.

LEMMA 4.4. *Let* $\mathscr{L}(R) = (2, 0; 1, 1)_{i,3}^{n-1}$, *let* $\mathscr{L}(X_1) = (1, 0; 0, 1)_{1,2}^{R-3}$ *and let* $\mathscr{L}(X_2) = (1, 0; 0, 1)_{2,1}^{R-3}$. *Then*

$$\mathscr{L}(X_1) = (2, 0; 1, 1)_{i+1,2}^{n-1} \quad \textit{and} \quad \mathscr{L}(X_2) = (2, 0; 1, 1)_{i+2,1}^{n-1}.$$

LEMMA 4.5. *Let* U_1 *and* U_2 *be uniform* $(0, 1)$ *random variables, independent of each other and of* R, *defined as in Lemma* 4.2. *Then there exist random variables* X_1 *with distribution* $(1, 0; 0, 1)_{1,2}^{R-3}$ *and* X_2 *with distribution* $(1, 0; 0, 1)_{2,1}^{R-3}$ *such that*

$$|X_1 - R\max(U_1, U_2)| < 3 \quad \textit{and} \quad |X_2 - R\min(U_1, U_2)| < 3 \qquad \textit{a.s.}$$

From these lemmas we can now prove Theorem 1.1(a); here and below we use C to denote a generic constant that may differ from line to line.

PROOF OF THEOREM 1.1(a). Let $W = W_{n,i}$ and $b = b_{n,i}$, let (R, W'') be defined as in Lemma 4.2 above and, as per Definition 1.3, let Y be a Bernoulli$(1/(1+i))$ random variable and $V = Y\max(U_1, U_2) + (1-Y)\min(U_1, U_2)$, where U_1 and U_2 are independent uniform $(0, 1)$ variables independent of Y. Lemmas 4.3, 4.4 and 4.5 imply that we can couple W and VR together so that $|W - VR| < 3$ almost surely. Thus, using Lemma 4.2,

$$\mathbb{P}(|W - VW''| > 3) \le \mathbb{P}(W'' \ne R) \le C/\sqrt{n},$$

and recalling that VW'' has the s-TDSB distribution, the theorem follows from (1.4) taking $\beta = 3/b$, noting that for $c > 0$, $(cW)' \stackrel{\mathscr{D}}{=} cW'$ and using $b^2 \ge Cn$ from Lemma 4.7 below. □

We have left to prove Lemmas 4.2–4.5 and 4.7; Lemma 4.3 is immediate after considering the urn process corresponding to $(2, 0; 1, 1)_{i,1}^{n}$ and conditioning on the color of the first ball drawn, which is white with probability $1/(1+i)$.

PROOF OF LEMMA 4.4. Consider an urn with i green balls, 1 black ball and 2 white balls. A ball is drawn at random and replaced in the urn along with another ball of the same color plus an additional green ball.

If X is the number of times a nongreen ball is drawn in $n-1$ draws, the number of white balls in the urn after $n-1$ draws is distributed as $(1,0;0,1)_{1,2}^{X}$. Since $X+3$ is distributed as $(2,0;1,1)_{i,3}^{n-1}$ (which is also that of R) and the number of white balls in the urn after $n-1$ draws has distribution $(2,0;1,1)_{i+1,2}^{n-1}$, the first equation follows. The second equation follows from similar considerations. □

PROOF OF LEMMA 4.5. We will show that for U_1 and U_2 independent uniform $(0,1)$ random variables, there exist random variables N and M such that $\mathscr{L}(N) = (1,0;0,1)_{1,2}^{n-3}$, $\mathscr{L}(M) = \mathscr{L}(n-N)$ and

$$\left|N - n\max(U_1,U_2)\right| < 3 \quad \text{and} \quad \left|M - n\min(U_1,U_2)\right| < 3 \qquad \text{a.s.}$$

The lemma follows from these "conditional" almost sure statements after noting that $\mathscr{L}(M) = (1,0;0,1)_{2,1}^{n-3}$ since we can think of $n-N$ as the number of black balls in the $(1,0;0,1)_{1,2}^{n-3}$ urn.

The formulas of Durrett [(2010), page 206] imply that $(1,0;0,1)_{1,2}^{n-3}$ has distribution function

$$F(k) = \left(\frac{k}{n-1}\right)\left(\frac{k-1}{n-2}\right), \qquad k = 1,\dots,n-1, \tag{4.2}$$

and it is straightforward to verify that

$$N := \max\bigl(\lceil (n-1)U_1\rceil, 1 + \lceil (n-2)U_2\rceil\bigr)$$

has the same distribution. We find $|N - n\max(U_1,U_2)| < 3$ and thus a coupling satisfying the first claim above. Defining

$$M := \min\bigl(\lceil (n-1)U_1\rceil, 1 + \lceil (n-2)U_2\rceil\bigr),$$

(4.2) implies $\mathscr{L}(M) = \mathscr{L}(n-N)$ and $|M - n\min(U_1,U_2)| < 3$. □

Before proving Lemma 4.2, we provide a useful construction for the double size bias distribution of a sum of indicators.

LEMMA 4.6. *Let $W = \sum_{i=1}^{n} X_i$, where the X_i are Bernoulli random variables and $b^2 := \mathbb{E}W^2$. For each $j,k \in \{1,\dots,n\}$, let $(X_i^{(j,k)})_{i\notin\{j,k\}}$ have the distribution of $(X_i)_{i\notin\{j,k\}}$ conditional on $X_j = X_k = 1$ and let J and K be random variables independent of the variables above satisfying*

$$\mathbb{P}(J=j, K=k) = \frac{\mathbb{E}(X_j X_k)}{b^2}, \qquad j,k \ge 1.$$

Then,

$$W'' = \sum_{i\notin\{J,K\}} X_i^{(J,K)} + 2 - \mathrm{I}[J=K]$$

has the double size bias distribution of W.

PROOF. We have

$$\begin{aligned}\mathbb{E}f(W'') &= b^{-2}\sum_{j,k}\mathbb{E}(X_jX_k)\mathbb{E}f\Big(\sum_{i\notin\{j,k\}}X_i^{(j,k)}+2-\mathrm{I}[j=k]\Big)\\ &= b^{-2}\sum_{j,k}\mathbb{E}(X_jX_k)\mathbb{E}\{f(W)|X_j=X_k=1\}\\ &= b^{-2}\sum_{j,k}\mathbb{E}\{X_jX_kf(W)\}=b^{-2}\mathbb{E}\{W^2f(W)\};\end{aligned}$$

this is exactly (1.2). □

To simplify the notation we consider i fixed in what follows. We write

$$W_n=\sum_{j=0}^{n}X_j,$$

where for $j\geq 1$, X_j is the indicator that a white ball is drawn on draw j from the $(2,0;1,1)_{i,1}$ urn and $X_0=1$ to represent the initial white ball in the urn. We will then define random variables $M_n^{j,k}$ such that

$$\mathscr{L}(M_n^{j,k})=\mathscr{L}(W_n|X_j=X_k=1), \tag{4.3}$$

so that by Lemma 4.6, if J and K are random variables independent of $M_n^{j,k}$ satisfying

$$\mathbb{P}(K=k,J=j)=\frac{\mathbb{E}(X_kX_j)}{b^2},\qquad j,k\geq 0 \tag{4.4}$$

for $b^2:=\mathbb{E}W^2$, then $M_n^{J,K}$ has the double size bias distribution of W.

In order to generate a variable satisfying (4.3) for $j<k$, we use the following lemma that yields a method to construct an urn process having the law of the $(2,0;1,1)_{i,1}$ urn process up to time n conditional on $X_k=X_j=1$. This conditioned process follows the law of the $(2,0;1,1,)_{i,3}$ urn process up to (and including) draw $j-1$. At draw j, exactly one *black* ball is added and then draws $j+1$ through $k-1$ follow the $(2,0;1,1)$ urn law. Again at draw k exactly one *black* ball is added and then the process continues to draw n following the $(2,0;1,1)$ urn rule. We write $M_n^{j,k}$ to denote the number of white balls in the urn after n draws in this process, and we refer to this process as the $M^{j,k}$ process. Our next main result shows that this construction of $M_n^{j,k}$ has the distribution specified in (4.3). First we state a technical lemma; the proof can be found at the end of this section.

LEMMA 4.7. *Fix $i\geq 1$ and let $W_n=\sum_{j=0}^{n}X_j$ where for $j\geq 1$, X_j is the indicator that a white ball is drawn on draw j from the $(2,0;1,1)_{i,1}$ urn and $X_0=1$. If $1\leq j<k\leq n$, then*

$$\mathbb{P}(X_j=1|X_k=1,W_{j-1})=\frac{1+W_{j-1}}{2j+i}, \tag{4.5}$$

and if $1 \le l < j < k \le n$,

$$\mathbb{P}(X_l = 1 | X_k = 1, X_j = 1, W_{l-1}) = \frac{2 + W_{l-1}}{2l + i + 1}. \tag{4.6}$$

For all $1 \le j < k \le n$,

$$\mathbb{E}W_n \le \sqrt{2\pi}\sqrt{\frac{n}{i+2} + \frac{1}{2}}, \qquad \mathbb{E}X_j \le \frac{\sqrt{\pi}}{\sqrt{(i+1)(i+2j-1)}},$$

$$(2 - \sqrt{\pi})\frac{i + 2n + 1}{i + 1} \le \mathbb{E}W_n^2 \le 2\frac{i + 2n + 1}{i + 1},$$

$$\mathbb{E}(X_j X_k) \le \frac{\sqrt{2\pi}(1 + \sqrt{\pi})}{(i+2)\sqrt{(i+2j)(i+2k-1)}}.$$

If $\mathscr{L}(R_t) = (2, 0; 1, 1)_{i,3}^t$, *then for some constant* C *independent of* t *and* i,

$$\mathbb{E}R_t \le C\sqrt{t/i}. \tag{4.7}$$

LEMMA 4.8. *Let* $1 \le j < k \le n$ *and* $M_n^{j,k}$, W_n *and* $(X_l)_{l \ge 1}$ *be defined as in Lemma* 4.7 *and the preceding two paragraphs. Then*

$$\mathscr{L}(M_n^{j,k}) = \mathscr{L}(W_n | X_j = X_k = 1).$$

PROOF. Let $M_l^{j,k} = 3 + \sum_{t=1}^{l} m_t^{j,k}$, where for $t \ne j, k$, $m_t^{j,k}$ is the indicator that draw t in the $M^{j,k}$ urn process is white and $m_j^{j,k} = m_k^{j,k} = 0$. From the definition of the process, for $l < j$,

$$\mathbb{P}(m_l^{j,k} = 1 | M_{l-1}^{j,k}) = \frac{M_l^{j,k}}{2l + i + 1}. \tag{4.8}$$

And for $j < l < k$, since $m_j^{j,k} = 0$,

$$\mathbb{P}(m_l^{j,k} = 1 | M_{l-1}^{j,k}) = \frac{M_l^{j,k}}{2l + i}. \tag{4.9}$$

Note also that $M_0^{j,k} = W_0 + 2 = 3$, $m_j^{j,k} = m_k^{j,k} = 0$ and draw $l > k$ in the $M^{j,k}$ urn process follows the $(2, 0; 1, 1)$ urn law. Now comparing (4.8) to (4.6) and (4.9) to (4.5), we find the sequential conditional probabilities agree and so the lemma follows. □

We are now ready to prove Lemma 4.2, and we first give the following remark about the argument. The $(2, 0; 1, 1)_{i,3}$ process and the $M^{j,k}$ process defined above differ only in that, in the latter process, after each of draws j and k a single black ball is added into the urn regardless of what is drawn; in the former process, the two balls added to the urn in these draws depend on the color drawn. This difference turns out to be small enough to allow a close coupling as stated in the lemma.

PROOF OF LEMMA 4.2. For each t we construct $(r_t, m_t^{j,k})$ to, respectively, denote the indicator for the event that a white ball is added to the urn after draw number t for the $(2, 0; 1, 1)_{i,3}$ process and for the $M^{j,k}$ process, and we write

$$R_{n-1} = 3 + \sum_{t=1}^{n-1} r_t, \qquad M_n^{j,k} = 3 + \sum_{t=1}^{n} m_t^{j,k}$$

to denote the number of white balls in the urn after draw $n - 1$ and n, respectively, for each process. Let U_t be independent uniform $(0, 1)$ random variables. We define

$$r_t = \mathrm{I}\left[U_t < \frac{R_{t-1}}{i + 2t + 1}\right] \tag{4.10}$$

and for $t \neq k, t \neq j$ we define

$$m_t^{j,k} = \mathrm{I}\left[U_t < \frac{M_{t-1}^{j,k}}{i + 2t + 1 - \mathrm{I}[t > j] - \mathrm{I}[t > k]}\right].$$

We also set $m_k^{j,k} = m_j^{j,k} = 0$ since a single black ball is added after draws j and k. Writing the event $M_n^{j,k} \neq R_{n-1}$ as a union of the events that index t is the least index such that $r_t \neq m_t^{j,k}$, and also using that $m_k^{j,k} = m_j^{j,k} = 0$, we find that for $0 < j < k$,

$$\begin{aligned}
&\mathbb{P}(M_n^{j,k} \neq R_{n-1}) \\
&\quad \leq \mathbb{E}(r_k + r_j) + \mathbb{P}\left(U_n < \frac{R_{n-1}}{i + 2n - 1}\right) \\
&\qquad + \sum_{t=j}^{n-1} \mathbb{P}\left(\frac{R_{t-1}}{i + 2t + 1} < U_t < \frac{R_{t-1}}{i + 2t - 1}\right).
\end{aligned}$$

From (4.10), $\mathbb{E}r_t = \mathbb{E}R_{t-1}/(i + 2t + 1)$, so that we find

$$\begin{aligned}
&\mathbb{P}(M_n^{j,k} \neq R_{n-1}) \\
&\quad \leq \frac{\mathbb{E}R_{k-1}}{i + 2k + 1} + \frac{\mathbb{E}R_{j-1}}{i + 2j + 1} + \frac{\mathbb{E}R_{n-1}}{i + 2n - 1} \\
&\qquad + \sum_{t=j}^{n-1} \mathbb{E}R_{t-1}\left(\frac{1}{i + 2t - 1} - \frac{1}{i + 2t + 1}\right) \\
&\quad \leq C\sqrt{\frac{k}{i}}\left(\frac{1}{i + 2k + 1}\right) + C\sqrt{\frac{j}{i}}\left(\frac{1}{i + 2j + 1}\right) + C\sqrt{\frac{n}{i}}\left(\frac{1}{i + 2n - 1}\right)
\end{aligned}$$

$$+ C \sum_{t=j}^{n-1} \sqrt{\frac{t}{i}} \left(\frac{1}{i+2t-1} - \frac{1}{i+2t+1} \right)$$
$$\leq C j^{-1/2},$$

where we have used (4.7). Defining J and K as in (4.4) we now have

$$\begin{aligned} &\mathbb{P}(M_n^{J,K} \neq R_{n-1}) \\ &\quad \leq \mathbb{P}(J=0) + \mathbb{P}(K=0) + \mathbb{P}(J=K) + \frac{2C}{b^2} \sum_{j<k} j^{-1/2} \mathbb{E} X_j X_k. \end{aligned} \tag{4.11}$$

Since $X_i^2 = X_i$, $X_0 = 1$ and using (4.4), we have

$$\mathbb{P}(J=K) = \mathbb{P}(K=0) = \mathbb{P}(J=0) = \mathbb{E} W_n / b^2,$$

and now using the bounds on $\mathbb{E} X_i X_j$, $\mathbb{E} W_n$ and $\mathbb{E} W_n^2$ from Lemma 4.7, we find that (4.11) is bounded above by

$$\frac{3\mathbb{E}W_n}{b^2} + \frac{C}{b^2} \sum_{j<k} \frac{1}{\sqrt{jk}} \frac{1}{\sqrt{j}} \leq \frac{C}{\sqrt{n}} + \frac{C}{n} \sum_{j<k} j^{-3/2} \leq \frac{C}{\sqrt{n}} + C \sum_j j^{-3/2} \leq \frac{C}{\sqrt{n}}. \quad \square$$

PROOF OF LEMMA 4.7. Let $A_m = \{X_m = 1\}$. By the definition of conditional probability,

$$\mathbb{P}(A_l | A_k, A_j, W_{l-1}) = \frac{\mathbb{P}(A_l | W_{l-1}) \mathbb{P}(A_k A_j | A_l, W_{l-1})}{\mathbb{P}(A_k A_j | W_{l-1})} \tag{4.12}$$

and

$$\mathbb{P}(A_j | A_k, W_{j-1}) = \frac{\mathbb{P}(A_j | W_{j-1}) \mathbb{P}(A_k | A_j, W_{j-1})}{\mathbb{P}(A_k | W_{j-1})}, \tag{4.13}$$

and we next will calculate the probabilities above. For $j \geq 1$, we have

$$\mathbb{P}(A_j | W_{j-1}) = \frac{W_{j-1}}{i + 2j - 1}, \tag{4.14}$$

which implies that for $k \geq j$,

$$\mathbb{P}(A_k | W_{j-1}) = \frac{\mathbb{E}(W_{k-1} | W_{j-1})}{i + 2k - 1} \tag{4.15}$$

and

$$\mathbb{P}(A_k | A_j, W_{j-1}) = \frac{\mathbb{E}(W_{k-1} | A_j, W_{j-1})}{i + 2k - 1}. \tag{4.16}$$

Now to compute the conditional expectations appearing above note first that

$$\mathbb{E}(W_k | W_{k-1}) = W_{k-1} + \frac{W_{k-1}}{i + 2k - 1} = \left(\frac{i + 2k}{i + 2k - 1} \right) W_{k-1}. \tag{4.17}$$

Conditioning on W_{j-1} and taking expectations yields

$$\mathbb{E}(W_k|W_{j-1}) = \left(\frac{i+2k}{i+2k-1}\right)\mathbb{E}(W_{k-1}|W_{j-1}),$$

and then iterating and substituting $k-1$ for k yields

$$\mathbb{E}(W_{k-1}|W_{j-1}) = \prod_{t=1}^{k-j}\left(\frac{i+2(k-t)}{i+2(k-t)-1}\right)W_{j-1}. \tag{4.18}$$

Using (4.18) with j substituted for $j-1$, we also find

$$\mathbb{E}(W_{k-1}|A_j, W_{j-1}) = \prod_{t=1}^{k-j-1}\left(\frac{i+2(k-t)}{i+2(k-t)-1}\right)(1+W_{j-1}); \tag{4.19}$$

note here that conditioning on A_j and W_{j-1} is equivalent to conditioning on W_j and the event $\{W_j = W_{j-1}+1\}$. We use a similar approach to obtain

$$\begin{aligned}\mathbb{E}(W_k^2|W_{k-1}) &= W_{k-1}^2\left(1-\frac{W_{k-1}}{i+2k-1}\right) + (W_{k-1}+1)^2\frac{W_{k-1}}{i+2k-1}\\ &= \left(\frac{i+2k+1}{i+2k-1}\right)W_{k-1}^2 + \frac{W_{k-1}}{i+2k-1},\end{aligned}$$

which can then be added to (4.17) while letting $D_k = W_k(1+W_k)$ to obtain

$$\mathbb{E}(D_k|W_{k-1}) = \frac{i+2k+1}{i+2k-1}D_{k-1},$$

and thus

$$\mathbb{E}(D_k|W_{j-1}) = \frac{i+2k+1}{i+2k-1}\mathbb{E}(D_{k-1}|W_{j-1}).$$

Iterating and substituting $k-1$ for k gives

$$\mathbb{E}(D_{k-1}|W_{j-1}) = \frac{i+2k-1}{i+2j-1}D_{j-1} = \frac{i+2k-1}{i+2j-1}W_{j-1}(1+W_{j-1}), \tag{4.20}$$

and using (4.20) with j substituted for $j-1$, we also find

$$\mathbb{E}(D_{k-1}|A_j, W_{j-1}) = \frac{i+2k-1}{i+2j+1}(W_{j-1}+1)(W_{j-1}+2).$$

Letting

$$c = \frac{1}{(i+2j-1)(i+2k-1)}\prod_{t=1}^{k-j-1}\left(\frac{i+2(k-t)}{i+2(k-t)-1}\right)$$

and applying (4.14), (4.16), (4.19) and (4.20) we have

$$\begin{aligned} \mathbb{P}(A_jA_k|W_{l-1}) &= \mathbb{E}\big(\mathbb{P}(A_j|W_{j-1})\mathbb{P}(A_k|A_j, W_{j-1})|W_{l-1}\big) \\ &= c\mathbb{E}(D_{j-1}|W_{l-1}) \\ &= c\frac{i+2j-1}{i+2l-1}W_{l-1}(1+W_{l-1}), \end{aligned} \tag{4.21}$$

and by substituting l for $l-1$ in (4.21), we also find

$$\mathbb{P}(A_jA_k|A_l, W_{l-1}) = c\frac{i+2j-1}{i+2l+1}(1+W_{l-1})(2+W_{l-1}). \tag{4.22}$$

Substituting (4.14)–(4.16), (4.18), (4.19), (4.21) and (4.22) appropriately into (4.12) and (4.13) proves (4.5) and (4.6).

From (4.18) we have

$$\mathbb{E}W_n = \prod_{t=1}^{n}\frac{i+2t}{i+2t-1} = \frac{\Gamma((i+1)/2)\Gamma(n+(i+1)/2+1/2)}{\Gamma((i+1)/2+1/2)\Gamma(n+(i+1)/2)}, \tag{4.23}$$

and from (4.14) we find

$$\mathbb{E}X_j = \frac{\mathbb{E}W_{j-1}}{i+2j-1}.$$

Now using (4.16) and (4.19) yields

$$\mathbb{E}(X_jX_k) = \frac{\mathbb{E}(1+W_{j-1})\mathbb{E}X_j}{i+2k-1}\prod_{t=j+1}^{k-1}\frac{i+2t}{i+2t-1},$$

and using (4.20) we find

$$\mathbb{E}W_n^2 = 2\frac{i+2n+1}{i+1} - \mathbb{E}W_n. \tag{4.24}$$

Lemma 2.7 applied to (4.23) implies

$$\frac{1}{\sqrt{\pi}}\sqrt{\frac{2n}{i+2}+1} \le \prod_{t=1}^{n}\frac{i+2t}{i+2t-1} \le \sqrt{\pi}\sqrt{\frac{2n}{i+2}+1}, \tag{4.25}$$

and collecting the appropriate facts above yields the bounds on $\mathbb{E}W_n$, $\mathbb{E}W_n^2$, $\mathbb{E}X_iX_j$ and $\mathbb{E}X_i$.

For the bound on $\mathbb{E}R_t$, an argument similar to (4.18) leading to (4.23) yields that

$$\mathbb{E}R_t = 3\prod_{m=1}^{t}\frac{i+2m+2}{i+2m+1} = \frac{\Gamma((i+3)/2)\Gamma(t+(i+3)/2+1/2)}{\Gamma((i+3)/2+1/2)\Gamma(t+(i+3)/2)},$$

which can be bounded by Lemma 2.7 resulting in inequalities which are similar to (4.25). □

Acknowledgments. We are grateful to Carl Morris for valuable advice on this project and an anonymous referee for very helpful and detailed comments. The authors have also learned that Robert Gaunt (University of Oxford) has independently been working on a second order Stein operator for the variance-gamma distribution, but there appears to be little overlap between his work and ours.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A., eds. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

ARRATIA, R. and GOLDSTEIN, L. (2010). Size bias, sampling, the waiting time paradox, and infinite divisibility: When is the increment independent? Unpublished manuscript. Available at arXiv:1007.3910 [math.PR].

BACKHAUSZ, Á. (2011). Limit distribution of degrees in random family trees. *Electron. Commun. Probab.* **16** 29–37. MR2763526

BARABÁSI, A.-L. and ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286** 509–512. MR2091634

BOLLOBÁS, B., RIORDAN, O., SPENCER, J. and TUSNÁDY, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18** 279–290. MR1824277

BROWN, M. (2006). Exploiting the waiting time paradox: Applications of the size-biasing transformation. *Probab. Engrg. Inform. Sci.* **20** 195–230. MR2261286

BUSTOZ, J. and ISMAIL, M. E. H. (1986). On gamma function inequalities. *Math. Comp.* **47** 659–667. MR0856710

CHAMAYOU, J.-F. and LETAC, G. (1999). Additive properties of the Dufresne laws and their multivariate extension. *J. Theoret. Probab.* **12** 1045–1066. MR1729469

CHEN, L. H. Y., GOLDSTEIN, L. and SHAO, Q.-M. (2011). *Normal Approximation by Stein's Method*. Springer, Heidelberg. MR2732624

DURRETT, R. (2007). *Random Graph Dynamics*. Cambridge Univ. Press, Cambridge. MR2271734

DURRETT, R. (2010). *Probability: Theory and Examples*, 4th ed. Cambridge Univ. Press, Cambridge. MR2722836

FORD, E. (2009). Barabási–Albert random graphs, scale-free distributions and bounds for approximation through Stein's method. Ph.D. thesis, Univ. Oxford.

GOLDSTEIN, L. (2007). L^1 bounds in normal approximation. *Ann. Probab.* **35** 1888–1930. MR2349578

GOLDSTEIN, L. (2009). Personal communication and unpublished notes. Stein workshop, January 2009, Singapore.

GOLDSTEIN, L. and REINERT, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7** 935–952. MR1484792

GORDON, R. D. (1941). Values of Mills' ratio of area to bounding ordinate and of the normal probability integral for large values of the argument. *Ann. Math. Statistics* **12** 364–366. MR0005558

JANSON, S. (2006). Limit theorems for triangular urn schemes. *Probab. Theory Related Fields* **134** 417–452. MR2226887

JANSON, S. (2010). Moments of gamma type and the Brownian supremum process area. *Probab. Surv.* **7** 1–52. MR2645216

LYONS, R., PEMANTLE, R. and PERES, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** 1125–1138. MR1349164

MÓRI, T. F. (2005). The maximum degree of the Barabási–Albert random tree. *Combin. Probab. Comput.* **14** 339–348. MR2138118

NEWMAN, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev.* **45** 167–256 (electronic). MR2010377

PAKES, A. G. and KHATTREE, R. (1992). Length-biasing, characterizations of laws and the moment problem. *Austral. J. Statist.* **34** 307–322. MR1193781

PEKÖZ, E. A. and RÖLLIN, A. (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **39** 587–608. MR2789507

PEKÖZ, E., RÖLLIN, A. and ROSS, N. (2012). Total variation error bounds for geometric approximation. *Bernoulli*. To appear. Available at arXiv:1005.2774 [math.PR].

PITMAN, J. and ROSS, N. (2012). Archimedes, Gauss, and Stein. *Notices Amer. Math. Soc.* **59** 1416–1421.

REINERT, G. (2005). Three general approaches to Stein's method. In *An Introduction to Stein's Method. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **4** 183–221. Singapore Univ. Press, Singapore. MR2235451

ROSS, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293. MR2861132

ROSS, S. and PEKÖZ, E. (2007). *A Second Course in Probability*. www.ProbabilityBookstore.com, Boston, MA.

STEIN, C. (1986). *Approximate Computation of Expectations*. IMS, Hayward, CA. MR0882007

E. A. PEKÖZ
SCHOOL OF MANAGEMENT
BOSTON UNIVERSITY
595 COMMONWEALTH AVENUE
BOSTON, MASSACHUSETTS 02215
USA

A. RÖLLIN
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
6 SCIENCE DRIVE 2
SINGAPORE 117546

N. ROSS
DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
367 EVANS HALL #3860
BERKELEY, CALIFORNIA 94720-3860
USA