Bernoulli **25**(4A), 2019, 2758–2792
https://doi.org/10.3150/18-BEJ1070

Dickman approximation in simulation, summations and perpetuities

CHINMOY BHATTACHARJEE[1] and LARRY GOLDSTEIN[2]

[1] *IMSV, Universität Bern, 3012 Bern, Switzerland.*
E-mail: chinmoy.bhattacharjee@stat.unibe.ch
[2] *Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA.*
E-mail: larry@math.usc.edu

The generalized Dickman distribution $\mathcal{D}_\theta$ with parameter $\theta > 0$ is the unique solution to the distributional equality $W =_d W^*$, where

$$W^* =_d U^{1/\theta}(W+1), \tag{1}$$

with W non-negative with probability one, $U \sim \mathcal{U}[0,1]$ independent of W, and $=_d$ denoting equality in distribution. These distributions appear in number theory, stochastic geometry, perpetuities and the study of algorithms. We obtain bounds in Wasserstein type distances between $\mathcal{D}_\theta$ and the distribution of

$$W_n = \frac{1}{n}\sum_{i=1}^{n} Y_k B_k,$$

where $B_1, \dots, B_n, Y_1, \dots, Y_n$ are independent with B_k distributed $\operatorname{Ber}(1/k)$ or $\mathcal{P}(\theta/k)$, $E[Y_k] = k$ and $\operatorname{Var}(Y_k) = \sigma_k^2$, and provide an application to the minimal directed spanning tree in $\mathbb{R}^2$. We also provide bounds with optimal rates for the Dickman convergence of weighted sums, arising in probabilistic number theory, of the form

$$S_n = \frac{1}{\log(p_n)}\sum_{k=1}^{n} X_k \log(p_k),$$

where $(p_k)_{k\ge 1}$ is an enumeration of the prime numbers in increasing order and X_k is geometric with parameter $(1 - 1/p_k)$, Bernoulli with success probability $1/(1+p_k)$ or Poisson with mean λ_k.

Lastly, we broaden the class of generalized Dickman distributions by studying the fixed points of the transformation

$$s\bigl(W^*\bigr) =_d U^{1/\theta} s(W+1)$$

generalizing (1), that allows the use of non-identity utility functions $s(\cdot)$ in Vervaat perpetuities. We obtain distributional bounds for recursive methods that can be used to simulate from this family.

Keywords: delay equation; distributional approximation; primes; utility; weighted Bernoulli sums

1. Introduction

The Dickman distribution $\mathcal{D}$ first made its appearance in [16] in the context of number theory for counting the number of integers below a fixed threshold whose prime factors lie below a given upper bound; see the more recent work [26] for a readable explanation of how the Dickman distribution arises there. Members from the broader class of generalized Dickman distributions $\mathcal{D}_\theta$ for $\theta > 0$, of which $\mathcal{D} = \mathcal{D}_1$, have since been used to approximate counts in logarithmic combinatorial structures, including permutations and partitions in [6], and more generally for the quasi-logarithmic class considered in [7], for the weighted sum of edges connecting vertices to the origin in minimal directed spanning trees in [25], and for certain weighted sums of independent random variables in [27]. Simulation of the generalized Dickman distribution has been considered in [15], and in connection with the Quickselect sorting algorithm in [22] and [19].

Following [19], for a given $\theta > 0$ and non-negative random variable W, define the θ-Dickman bias distribution of W by

$$W^* =_d U^{1/\theta}(W+1), \tag{1.1}$$

where $U \sim \mathcal{U}[0,1]$ and is independent of W, and $=_d$ denotes equality in distribution. Though the density of $\mathcal{D}_\theta$ can presently be given only by specifying it somewhat indirectly as a certain solution to a differential delay equation, it is well known [15] that the distributions $\mathcal{D}_\theta$ are characterized by satisfying $W^* =_d W$ uniquely, that is, $\mathcal{D}_\theta$ is the unique fixed point of the distributional transformation (1.1). Indeed, this property is the basis for simulating from this family using the recursion

$$W_{n+1} = U_n^{1/\theta}(W_n + 1) \qquad \text{for } n \geq 0, \text{ with } W_0 = 0, \tag{1.2}$$

where $U_m, m \geq 0$ are i.i.d. $\mathcal{U}[0,1]$ random variables and U_n is independent of W_n, see [15].

Generally, distributional characterizations and their associated transformations, such as (1.1), provide an additional avenue to study distributions and their approximation, and have been considered for the normal [21], the exponential [23], and various other distributions that may be less well known, such as one arising in the study of the degrees of vertices in certain preferential attachment graphs, see [24].

In the following, D_θ will denote a $\mathcal{D}_\theta$ distributed random variable, where the subscript may be dropped when equal to 1. In [20], the upper bound

$$d_1(W, D_\theta) \leq (1+\theta) d_1\big(W, W^*\big) \tag{1.3}$$

for the Wasserstein distance between a non-negative random variable W and D_θ was proved, where

$$d_1(X,Y) = \sup_{h \in \mathrm{Lip}_1} \big|Eh(X) - Eh(Y)\big| \tag{1.4}$$

with

$$\mathrm{Lip}_\alpha = \big\{h : \big|h(x) - h(y)\big| \leq \alpha|x-y|\big\} \qquad \text{for } \alpha \geq 0. \tag{1.5}$$

We also apply the fact that alternatively one can write

$$d_1(X,Y) = \inf E|X-Y|, \tag{1.6}$$

where the infimum is over all joint distributions having the given X, Y marginals. The infimum is achieved for variables taking values in any Polish space, see, for example, [28], and so in particular for those that are real valued. For notational simplicity, we write $d_1(X, Y)$, say, for $d_1(\mathcal{L}(X), \mathcal{L}(Y))$, where $\mathcal{L}(\cdot)$ stands for the distribution, or law, of a random variable. In [20], inequality (1.3) was used to derive a bound on the quality of the Dickman approximation for the running time of the Quickselect algorithm.

Here our aim is two fold. First, in Section 2 we study the approximation of sums that converge in distribution to Dickman, for instance, those of the form

$$W_n = \frac{1}{n}\sum_{k=1}^{n} Y_k B_k, \tag{1.7}$$

where $\{B_1, \dots, B_n, Y_1, \dots, Y_n\}$ are independent, B_k is a Bernoulli random variable with success probability $1/k$, and Y_k is non-negative with $EY_k = k$, and $\mathrm{Var}(Y_k) = \sigma_k^2$ for all $k = 1, \dots, n$. The most well known case is the one where $Y_k = k$ a.s., for which

$$W_n = \frac{1}{n}\sum_{k=1}^{n} k B_k. \tag{1.8}$$

Sums of this type arise, for instance, in the analysis of the Quickselect algorithm for finding the m^{th} smallest of a list of n distinct numbers, see [22] (also [20]), and for the sum of positions of records in a uniformly random permutation (see [30]). To state the result we will apply to such sums, we first define the Wasserstein-2 metric

$$d_{1,1}(X, Y) = \sup_{h \in \mathcal{H}_{1,1}} \bigl|Eh(Y) - Eh(X)\bigr|, \tag{1.9}$$

where, for $\alpha \ge 0, \beta \ge 0$,

$$\mathcal{H}_{\alpha,\beta} = \bigl\{h : h \in \mathrm{Lip}_\alpha, h' \in \mathrm{Lip}_\beta\bigr\}, \tag{1.10}$$

with Lip_α given in (1.5). The work [3] obtains a bound of the form $C\sqrt{\log n}/n$ between W_n in (1.8) and D in a metric weaker than $d_{1,1}$ in (1.9), requiring test functions to be three times differentiable, and with the constant C unspecified. The following theorem provides a more general result that in the specific case of (1.8) yields a bound in the stronger metric $d_{1,1}$ with a small, explicit constant.

Theorem 1.1. *Let W_n be as in* (1.7) *and D a standard Dickman random variable. Then with the metric $d_{1,1}$ in* (1.9),

$$d_{1,1}(W_n, D) \le \frac{3}{4n} + \frac{1}{2n^2}\sum_{k=1}^{n} \frac{1}{k}\sqrt{\bigl(\sigma_k^2 + k^2\bigr)\sigma_k^2},$$

and in particular if $Y_k = k$ a.s., that is, for W_n as in (1.8),

$$d_{1,1}(W_n, D) \le \frac{3}{4n}. \tag{1.11}$$

From the first bound given by the theorem, speaking asymptotically we see that W_n in (1.7) converges to D in distribution whenever $\sum_{k=1}^{n} \frac{1}{k}\sqrt{(\sigma_k^2+k^2)\sigma_k^2} = o(n^2)$. In particular, weak convergence to the Dickman distribution occurs if $\sigma_k^2 = O(k^{2-\varepsilon})$ for some $\varepsilon > 0$. In Section 2, we provide an application of Theorem 1.1 to minimal directed spanning trees in $\mathbb{R}^2$.

We also show the following related result for a weighted sum of independent Poisson variables. For $\lambda > 0$, let $\mathcal{P}(\lambda)$ denote a Poisson random variable with mean λ.

Theorem 1.2. *For $\theta > 0$, let $\{P_1, \dots, P_n, Y_1, \dots, Y_n\}$ be independent with $P_k \sim \mathcal{P}(\theta/k)$ and Y_k non-negative with $EY_k = k$ and $\mathrm{Var}(Y_k) = \sigma_k^2$, for all $k = 1, \dots, n$. Then*

$$W_n = \frac{1}{n}\sum_{k=1}^{n} Y_k P_k \tag{1.12}$$

satisfies

$$d_{1,1}(W_n, D_\theta) \le \frac{\theta}{4n} + \frac{\theta}{n}\sum_{k=1}^{n}\frac{\sigma_k}{k} + \frac{\theta}{2n^2}\sum_{k=1}^{n}\frac{1}{k}\sqrt{\left(\sigma_k^2+k^2\right)\sigma_k^2},$$

and in particular, in the case $Y_k = k$ a.s.,

$$W_n = \frac{1}{n}\sum_{k=1}^{n} kP_k \qquad \textit{satisfies } d_{1,1}(W_n, D_\theta) \le \frac{\theta}{4n}.$$

Similar to the weighted sum of Bernoullis in (1.7), we have weak convergence to the Dickman distribution if $\sigma_k^2 = O(k^{2-\varepsilon})$ for some $\varepsilon > 0$.

Next, we study Dickman approximation of weighted geometric and Bernoulli sums that appear in probabilistic number theory. For geometric variables, we write $X \sim \mathrm{Geom}(p)$ if $P(X = m) = (1-p)^m p$ for $m \ge 0$. Let $(p_k)_{k\ge 1}$ be an enumeration of the prime numbers in increasing order and Ω_n denote the set of all positive integers having no prime factor larger than p_n. Let $X_1, \dots, X_n$ be independent with $X_k \sim \mathrm{Geom}(1 - 1/p_k)$ for $1 \le k \le n$, and let Π_n be the distribution of M_n given by

$$M_n = \prod_{k=1}^{n} p_k^{X_k} \qquad \text{and let } S_n = \frac{\log M_n}{\log(p_n)} = \frac{1}{\log(p_n)}\sum_{k=1}^{n} X_k \log(p_k). \tag{1.13}$$

One can specify (see, e.g., [26]) Π_n by

$$\Pi_n(m) = \frac{1}{\pi_n m} \qquad \text{for } m \in \Omega_n$$

with normalizing constant necessarily satisfying $\pi_n = \sum_{m\in\Omega_n} 1/m$. Distributional convergence of S_n to the standard Dickman distribution was proved in [26]. In Theorem 1.3 below, we provide a $(\log n)^{-1}$ convergence rate in the Wasserstein-2 norm.

Theorem 1.3. *For D a standard Dickman random variable and S_n as in (1.13) with $X_1, \dots, X_n$ independent variables with $X_k \sim \mathrm{Geom}(1 - 1/p_k)$, we have*

$$d_{1,1}(S_n, D) \le \frac{C}{\log n}$$

for some universal constant C. Moreover, the order is not improvable.

One may instead consider the distribution Π_n' over Ω_n', the set of square-free integers with largest prime factor less than or equal to p_n, with $\Pi_n'(m)$ proportional to $1/m$ for all $m \in \Omega_n'$. Then $M_n = \prod_{k=1}^n p_k^{X_k}$ has distribution Π_n' when $X_k \sim \mathrm{Ber}(1/(1 + p_k))$ and are independent (see, e.g., [13]). That $S_n = \log M_n / \log(p_n)$ converges in distribution to the standard Dickman was proved in [13] and very recently a $(\log \log n)^{3/2} (\log n)^{-1}$ rate was provided in [3] in a metric defined as a supremum over a class of three times differentiable functions. We provide the improved $(\log n)^{-1}$ convergence rate in the stronger Wasserstein-2 norm.

Theorem 1.4. *For D a standard Dickman random variable and S_n as in (1.13) with $X_1, \dots, X_n$ independent variables with $X_k \sim \mathrm{Ber}(1/(1 + p_k))$, we have*

$$d_{1,1}(S_n, D) \le \frac{C}{\log n}$$

for some universal constant C. Moreover, the order is not improvable.

In Examples 2.1 and 2.2, we also provide such bounds when the X_k's in (1.13) are distributed as Poisson random variables with parameters $\lambda_k > 0$ given by certain functions of p_k. For our results in probabilistic number theory, we closely follow the arguments in [3].

In Section 3, we consider the connection between the class of Dickman distributions and perpetuities. By approaching from the view of utility, we extend the scope of the Dickman distributions past the currently known class. The recursion (1.2) was interpreted by Vervaat, see [37], as the relation between the values of a perpetuity at two successive times. In particular, during the n^{th} time period a deposit of some fixed value, scaled to be unity, is added to the value of an asset. During that time period, a multiplicative factor in [0, 1], accounting for depreciation is applied; in (1.2) that factor is taken to be $U^{1/\theta}$. The generalized Dickman distributions arise as fixed points of this recursion, that is, solutions to $W^* =_d W$ where W^* is given in (1.1).

Measuring the value of an asset directly by its monetary value corresponds to the case where the utility function $s(\cdot)$ of an asset is taken to be the identity. We consider the generalization of (1.2) to

$$s(W_{n+1}) = U_n^{1/\theta} s(W_n + 1). \tag{1.14}$$

In [9], see also the translation [10], Daniel Bernoulli argued that utility should be given as a concave function of the value of an asset, typically justified by observing that receiving one unit of currency would be of more value to an individual who has very few resources than one who has resources in abundance, see [17]. We may then interpret (1.14) in a manner similar to (1.2), but now in terms of utility. Again, during the n^{th} time period, a constant value, scaled to be one, is added to an asset. Then, at time $n+1$, the utility of the asset is given by some discount factor applied to the incremented utility of the asset. When $s(\cdot)$ is invertible, as for the most common Vervaat perpetuities, one can now gain insight into their long term behavior by studying fixed points of the transformation

$$W^* =_d s^{-1}\big(U^{1/\theta}s(W+1)\big). \tag{1.15}$$

Theorem 3.3 in Section 3 shows that under mild and natural conditions on the utility function $s(\cdot)$ the transformation (1.15) has a unique fixed point, say $D_{\theta,s}$, which we say has the (θ,s)-Dickman distribution, denoted here as $\mathcal{D}_{\theta,s}$. As the identity function $s(x)=x$ recovers the class of generalized Dickman distributions, this extended class strictly contains them. The parameter $\theta>0$ here plays the same role for $\mathcal{D}_{\theta,s}$ as it does for $\mathcal{D}_\theta$, in particular in its appearance in the distributional bounds for simulation using recursive schemes. Theorem 3.4 generalizes the bound (1.3) of [20] to the $\mathcal{D}_{\theta,s}$ family, providing the inequality

$$d_1(W,D_{\theta,s}) \le (1-\rho)^{-1}d_1\big(W^*,W\big) \tag{1.16}$$

with a parameter ρ given by a bound on an integral involving θ and $s(\cdot)$, see (3.10) and (3.11).

We apply (1.16) to assess the quality of the recursive scheme

$$W_{n+1} = s^{-1}\big(U_n^{1/\theta}s(W_n+1)\big) \qquad \text{for } n\ge 0 \text{ and } W_0=0, \tag{1.17}$$

for the simulation of variables having the $\mathcal{D}_{\theta,s}$ distribution. Simulation by these means for the $\mathcal{D}_\theta$ family was considered in [15], though no bounds on its accuracy were provided. An algorithmic method for the exact simulation from the $\mathcal{D}_\theta$ family was given in [18] with bounds on the expected running time. In brief, the method in [18] depends on the use of a multigamma coupler as an update function for the kernel $K(x,\cdot) := \mathcal{L}(U^{1/\theta}(x+1))$, and on finding a dominating chain so that one can simulate from its stationary distribution, a shifted geometric distribution in this case. To extend this approach to the more general family $\mathcal{D}_{\theta,s}$, one would consider the kernel $K(x,\cdot) := \mathcal{L}(U^{1/\theta}s(x+1))$, and though one can generalize the multigamma coupler for use as an update function for this kernel, finding a suitable dominating chain in this generality may not be straightforward.

The efficacy of a simpler recursive scheme for simulation from this family is addressed in (3.17) of Corollary 3.2 where we show that the iterates generated by (1.17) obey the inequality

$$d_1(W_n,D_{\theta,s}) \le (1-\rho)^{-1}\left(\frac{\theta}{\theta+1}\right)^n E\big[s^{-1}\big(U^{1/\theta}\big)\big],$$

and which thus exhibit exponentially fast convergence. In Section 3.3, we present some instances from the $\mathcal{D}_{\theta,s}$ family that arise as limiting distributions for perpetuities when taking our utilities $s(\cdot)$ from those studied in economics.

We obtain our results by extensions of [19] for the Stein's method framework for the Dickman distribution. The application of Stein's method, as unveiled in [34] and further developed in [35], begins with a characterizing equation for a given target distribution. Such a characterization is then used as the basis to form a Stein equation, which is usually a difference or differential equation involving test functions in a class corresponding to a desired probability metric, such as the class of Lip_1 functions for the Wasserstein distance in (1.4). One key step of the method requires bounds on the smoothness of solutions over the given class of test functions. For a modern treatment of Stein's method, see [14] and [31].

Theorems 1.4 improves on results of [3]. That work applies a different version of Stein's method, and in particular does not consider any form of the Stein equation, such as (1.18) or (1.20). Consequently [3] does not obtain bounds on a Stein solution for any Dickman case, as is achieved here in Theorems 4.7 and 4.9. Indeed, there it is noted in [2] that this last step can be an "extremely difficult problem".

In [19] the Stein equation used for the $\mathcal{D}_\theta$ family was of the integral type

$$g(x) - A_{x+1}g = h(x) - E\big[h(D_\theta)\big], \tag{1.18}$$

where the averaging operator $A_x g$ was given by

$$A_x g = \begin{cases} g(0) & \text{for } x = 0, \\ \dfrac{\theta}{x^\theta}\displaystyle\int_0^x g(u)u^{\theta-1}\,du & \text{for } x > 0. \end{cases}$$

To handle the $\mathcal{D}_{\theta,s}$ family, over the range $x > 0$ we generalize the form of the averaging operator to

$$A_x g = \frac{1}{t(x)}\int_0^x g(u)t'(u)\,du, \tag{1.19}$$

where $t(x) = s^\theta(x)$. Smoothness bounds for solutions of (1.18), with A_x as in (1.19) and D_θ replaced by $D_{\theta,s}$, are given in Theorem 4.7 in Section 4 for a wide range of functions $s(\cdot)$. This generalization requires significant extensions of existing methods.

Use of the Stein equation (1.18) is appropriate when the variable W of interest can be coupled to some W^* with its θ-Dickman bias distribution. However, such direct couplings appear elusive for all our examples in Section 2, including in particular those in probabilistic number theory, and a different approach is needed. To handle these examples we consider instead a new Stein equation, of differential-delay type, given by

$$(x/\theta)f'(x) + f(x) - f(x+1) = h(x) - E\big[h(D_\theta)\big]. \tag{1.20}$$

To apply the method, uniform bounds on the smoothness of the solution $f(\cdot)$ over test functions $h(\cdot)$ in some class $\mathcal{H}$ is required; we achieve such bounds for the class $\mathcal{H}_{1,1}$ in Theorem 4.9 in Section 4.

Throughout the paper, for a real-valued measurable function $f(\cdot)$ on a domain $S \subset \mathbb{R}$, $\|f\|_\infty$ denotes its essential supremum norm defined by

$$\|f\|_\infty = \operatorname*{ess\,sup}_{x\in S}\big|f(x)\big| = \inf\big\{b \in \mathbb{R} : m\big(\{x : f(x) > b\}\big) = 0\big\}, \tag{1.21}$$

where m denotes the Lebesgue measure on $\mathbb{R}$. For any real valued function defined on $A \subset S$, we define its supremum norm on A by

$$\|f\|_A = \sup_{x\in A}\big|f(x)\big|. \tag{1.22}$$

Unless otherwise specifically noted, integration will be with respect to m, which for simplicity will be denoted by, say, dv when the variable of integration is v.

This work is organized as follows. We focus on sums, such as the Bernoulli and Poisson weighted sums in (1.7) and (1.12), and sums arising in probabilistic number theory as (1.13), in Section 2. We focus on perpetuities, with examples, in Section 3, and in Section 4 we prove smoothness bounds on the two types of Stein solutions considered here.

2. Dickman approximation of sums

We will prove Theorems 1.1 and 1.2, starting with a simple application of the former, in Section 2.1, and then provide the proofs of Theorems 1.3 and 1.4, in probabilistic number theory, in Section 2.2. In this section, we deal with the form (1.20) of the Stein equation. That is, in the proofs of Theorems 1.1, 1.2 and 2.1, we take a fixed $\theta > 0$ and $h \in \mathcal{H}_{1,1}$, the function class defined in (1.10), and let $f \in \mathcal{H}_{\theta,\theta/2}$ be the solution of the Stein equation (1.20) that is guaranteed by Theorem 4.9. Substituting our W_n of interest for x in (1.20) and taking expectation yields

$$E\big[h(W_n)\big] - E\big[h(D_\theta)\big] = E\big[(W_n/\theta)f'(W_n) - \big(f(W_n+1) - f(W_n)\big)\big]. \tag{2.1}$$

2.1. Weighted Bernoulli and Poisson sums

We begin with a simple application of Theorem 1.1 to the minimal directed spanning tree, or MDST, following [11], first pausing to describe the construction of the MDST.

For two points (u_1, v_1) and (u_2, v_2) in $\mathbb{R}^2$, we write $(u_1, v_1) \preceq (u_2, v_2)$ if $u_1 \le u_2$ and $v_1 \le v_2$, and write $(u_1, v_1) \npreceq (u_2, v_2)$ otherwise. For any set of points $\mathcal{V}$ in $\mathbb{R}^2$, we say $(u, v) \in \mathcal{V}$ is a minimal point, or sink, of $\mathcal{V}$ if $(a, b) \npreceq (u, v)$ for all $(a, b) \in \mathcal{V}$, $(a, b) \neq (u, v)$.

For $n \in \mathbb{N}$, consider a set of $n + 1$ distinct points $\mathcal{V} = \{(a_i, b_i), 0 \le i \le n\}$ in $[0, 1] \times [0, 1]$ where we take $(a_0, b_0) = (0, 0)$, the origin. Let E be the set of directed edges $(a_i, b_i) \to (a_j, b_j)$ with $i \neq j$ and $(a_i, b_i) \preceq (a_j, b_j)$. Since $(0, 0) \preceq (a_i, b_i)$ for all $i = 1, \dots, n$, the edge set E contains all the directed edges $(a_0, b_0) \to (a_i, b_i)$ with $i \neq 0$. Let $\mathcal{G}$ be the collection of all graphs G with vertex set $G_V = \mathcal{V}$ and edge set $G_E \subseteq E$ such that for any $1 \le j \le n$, there exists a directed path from (a_0, b_0) to (a_j, b_j) with each edge in G_E. We define a MDST on $\mathcal{V}$ as any graph $T \in \mathcal{G}$ that minimizes $\sum_{e\in G_E} |e|$ where $|e|$ denotes the Euclidean length of the edge e. Clearly T is a tree and need not be unique.

Now let $\mathcal{P}$ be a random collection of n points uniformly and independently placed in the unit square $[0, 1]^2$ in $\mathbb{R}^2$. In this random setting, the MDST on the point set $\mathcal{V} = P \cup \{(0, 0)\}$ is uniquely defined almost surely, see [11]. By relabeling the points according to the size of their x-coordinate, without loss of generality, we may let the points in $\mathcal{P}$ be $(X_1, Y_1), \dots, (X_n, Y_n)$

where $Y_1, \dots, Y_n$ are independent $\mathcal{U}[0,1]$ random variables, and also independent of $X_1, \dots, X_n$, where $0 < X_1 < X_2 < \cdots < X_n < 1$ have the distribution of the order statistics generated from a sample of n independent $\mathcal{U}[0,1]$ variables.

Though the origin is the unique minimal point of $\mathcal{V}$, the usual set of interest is the collection of minimal points of $\mathcal{P}$, which has size at least one. For $i = 1, \dots, n$, observe that (X_i, Y_i) is a minimal point of $\mathcal{P}$ if and only if $Y_j > Y_i$ for all $j < i$. One much studied quantity in this context is the sum S_n of the αth powers of the Euclidean distances between the minimal points of the process and the origin for some $\alpha > 0$; the work [25] shows that S_n converges to $D_{2/\alpha}$ in distribution as n tends to infinity.

The lower record times $R_1, R_2, \dots$ of the height process $Y_1, \dots, Y_n$ are also studied, see [11], and are defined by letting $R_1 = 1$, and for $i > 1$ by

$$R_i = \begin{cases} \infty & \text{if } Y_j \ge Y_{R_{i-1}} \text{ for all } j > R_{i-1} \text{ or if } R_{i-1} \ge n, \\ \min\{j > R_{i-1} : Y_j < Y_{R_{i-1}}\} & \text{otherwise.} \end{cases}$$

In terms of these record times, the collection of the $k(n)$ minimal points inside the unit square is given by (X_{R_i}, Y_{R_i}) for $i = 1, \dots, k(n)$. We claim that the scaled sum of lower record times

$$W_n = \frac{1}{n} \sum_{i=1}^{k(n)} R_i \tag{2.2}$$

can be approximated by the Dickman distribution $\mathcal{D}$ in the Wasserstein-2 metric in (1.9) to within the bound specified by inequality (1.11) of Theorem 1.1. Indeed, for $1 \le j \le n$, letting $B_k = \mathbb{1}(k \in \{R_1, \dots, R_{k(n)}\})$ we have that $\sum_{i=1}^{k(n)} R_i = \sum_{k=1}^{n} k B_k$. As Lemma 2.1 of [11] shows that $B_1, \dots, B_n$ are independent with $B_k \sim \mathrm{Ber}(1/k)$ for $1 \le k \le n$, Theorem 1.1 yields the claimed bound for the Dickman approximation of (2.2).

We now present the proof of our first main result.

Proof of Theorem 1.1. Let W_n be as in (1.7) and take $\theta = 1$ in (2.1). Letting

$$W_n^{(k)} = W_n - \frac{Y_k}{n} B_k,$$

evaluating the first term on the right-hand side of (2.1) yields

$$\begin{aligned} E\big[W_n f'(W_n)\big] &= E\left[\frac{1}{n}\sum_{k=1}^n Y_k B_k f'(W_n)\right] = \frac{1}{n}\sum_{k=1}^n E\left[Y_k B_k f'\left(W_n^{(k)} + \frac{Y_k}{n} B_k\right)\right] \\ &= \frac{1}{n}\sum_{k=1}^n E\left[Y_k f'\left(W_n^{(k)} + \frac{Y_k}{n}\right)\right] P(B_k = 1) = \frac{1}{n}\sum_{k=1}^n E\left[\frac{Y_k}{k} f'\left(W_n^{(k)} + \frac{Y_k}{n}\right)\right]. \end{aligned}$$

The right-hand side of (2.1) is therefore the expectation of

$$\frac{1}{n}\sum_{k=1}^n \frac{Y_k}{k} f'\left(W_n^{(k)} + \frac{Y_k}{n}\right) - \int_0^1 f'(W_n + u)\,du$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{k=1}^{n}\frac{Y_k}{k}\left(f'\left(W_n^{(k)}+\frac{Y_k}{n}\right)-f'\left(W_n^{(k)}+\frac{k}{n}\right)\right) \\
&\quad + \frac{1}{n}\sum_{k=1}^{n}\left(\frac{Y_k}{k}f'\left(W_n^{(k)}+\frac{k}{n}\right)-f'\left(W_n^{(k)}+\frac{k}{n}\right)\right) \\
&\quad + \frac{1}{n}\sum_{k=1}^{n}\left(f'\left(W_n^{(k)}+\frac{k}{n}\right)-f'\left(W_n+\frac{k}{n}\right)\right) \\
&\quad + \left(\frac{1}{n}\sum_{k=1}^{n}f'\left(W_n+\frac{k}{n}\right)-\int_0^1 f'(W_n+u)\,du\right).
\end{aligned} \tag{2.3}
$$

Using that $f \in \mathcal{H}_{1,1/2}$, and hence in particular that $f'(\cdot)$ is Lipschitz, applying the Cauchy–Schwarz inequality to the first difference on the right-hand side of (2.3) we find that the expectation of that term is bounded by

$$
\frac{\|f''\|_\infty}{n^2}\sum_{k=1}^{n}E\left[\frac{|Y_k|}{k}|Y_k-k|\right] \le \frac{1}{2n^2}\sum_{k=1}^{n}\frac{1}{k}\sqrt{\left(\sigma_k^2+k^2\right)\sigma_k^2}.
$$

The expectation of the second difference is zero as $E[Y_k]=k$ and Y_k is independent of $W_n^{(k)}$. For the expectation of the third difference, noting that $E[Y_k B_k]=1$, we similarly obtain the bound

$$
\frac{\|f''\|_\infty}{n}\sum_{k=1}^{n}E\left|W_n^{(k)}-W_n\right| \le \frac{1}{2n}\sum_{k=1}^{n}E\left[\frac{Y_k}{n}B_k\right]=\frac{1}{2n}.
$$

Finally, for the fourth difference, applying that same bound on the second derivative of $f(\cdot)$, almost surely

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{k=1}^{n}f'\left(W_n+\frac{k}{n}\right)-\int_0^1 f'(W_n+u)\,du\right| \\
&\quad \le \sum_{k=1}^{n}\int_{\frac{k-1}{n}}^{\frac{k}{n}}\left|\left[f'(W_n+k/n)-f'(W_n+u)\right]\right|du \\
&\quad \le \frac{1}{2}\sum_{k=1}^{n}\int_{\frac{k-1}{n}}^{\frac{k}{n}}(k/n-u)\,du = \frac{1}{2}\left(\frac{1}{n^2}\sum_{k=1}^{n}k-\int_0^1 u\,du\right)=\frac{1}{4n}.
\end{aligned}
$$

Combining these three bounds yields, via (2.1) with $\theta=1$, that

$$
\left|E\left[h(W_n)\right]-E\left[h(D)\right]\right| \le \frac{3}{4n}+\frac{1}{2n^2}\sum_{k=1}^{n}\frac{1}{k}\sqrt{\left(\sigma_k^2+k^2\right)\sigma_k^2}.
$$

Taking the supremum over $\mathcal{H}_{1,1}$ and recalling the definition of the norm $d_{1,1}$ in (1.9) now yields the theorem. The final claim (1.11) holds as $\sigma_k^2 = 0$ when $Y_k = k$ a.s. $\square$

The proof of Theorem 1.2 proceeds along the same lines as that of Theorem 1.1, and can be found in the supplement [12].

2.2. Dickman approximation in number theory

Let $(p_k)_{k\geq 1}$ be an enumeration of the prime numbers in increasing order. Let $(X_k)_{k\geq 1}$ be a sequence of independent integer valued random variables and let

$$S_n = \frac{1}{\log(p_n)} \sum_{k=1}^{n} X_k \log(p_k) \qquad \text{for } n \geq 1. \tag{2.4}$$

Weak convergence of S_n to the Dickman distribution in the cases when the X_k's are distributed as geometric and Bernoulli variables is well known in probabilistic number theory, and [3] recently provided a rate of convergence in the Bernoulli case. We give bounds in a stronger metric and remove a logarithmic factor from their rate. We also prove such bounds when the X_k's are distributed as geometric or Poisson with parameters given by certain functions of p_k. For our results in this area, we rely heavily on the techniques in the proof of Lemma 2.3 of [3]; in particular, the identity (2.5) below, without remainder, is due to [3]. We begin with the following abstract theorem.

Theorem 2.1. *Let S be a non-negative random variable with finite variance such that for some constant μ and a random variable T satisfying $P(S+T=0)=0$,*

$$E\big[S\phi(S)\big] = \mu E\big[\phi(S+T)\big] + R_\phi \qquad \text{for all } \phi \in \mathrm{Lip}_{1/2}, \tag{2.5}$$

where the constant R_ϕ may depend on $\phi(\cdot)$. Then

$$d_{1,1}(S, D) \leq |\mu - 1| + \frac{1}{2} \inf_{(T,U)} E|T - U| + \sup_{\phi\in\mathrm{Lip}_{1/2}} |R_\phi|, \tag{2.6}$$

where D is a standard Dickman random variable, and the infimum is over all couplings (T, U) of T and $U \sim \mathcal{U}[0, 1]$ constructed on the same space as S, with U independent of S.

Remark 2.1. We note the connection between the relation in (2.5) and size biasing, where for a non-negative random variable S with finite mean μ, we say S^s has the S-size biased distribution when

$$E\big[S\phi(S)\big] = \mu E\big[\phi\big(S^s\big)\big]$$

for all functions $\phi(\cdot)$ for which these expectations exist. In particular, when R_ϕ in (2.5) is zero for all $\phi \in \mathrm{Lip}_{1/2}$, we obtain that $S^s =_d S + T$; for an application which requires the remainder, see

Lemma 2.2. Additionally, Section 4.3 of [6] shows that the standard Dickman D is the unique non-negative solution to the distributional equality $W^s =_d W + U$, where U is $\mathcal{U}[0,1]$, and independent of W. Hence, the error term comparing T and U in Theorem 2.1 is natural.

Proof of Theorem 2.1. We first show that the set of couplings over which the infimum is taken in (2.6) is non-empty. Note that the case when S is identically zero is trivial since one can take $\mu = 0$, $T = 0$ and $R_\phi = 0$ for all $\phi \in \mathrm{Lip}_{1/2}$. For a nontrivial S, let $\mu = E[S]$, and let S^s and U be constructed on the same space as S, independently of S, with S^s having the S-size biased distribution and $U \sim \mathcal{U}[0,1]$. Then setting $T = S^s - S$ identity (2.5) is satisfied with $R_\phi = 0$ for all $\phi \in \mathrm{Lip}_{1/2}$, and the pair (T, U) satisfies the conditions required of the infimum in the theorem.

Invoking Theorem 4.9 with $\theta = 1$, for any given $h \in \mathcal{H}_{1,1}$ there exists a function $f(\cdot)$ satisfying $\|f'\|_{(0,\infty)} \le 1$ and $\|f''\|_{(0,\infty)} \le 1/2$ such that

$$E\big[h(S)\big] - E\big[h(D)\big] = E\big[Sf'(S) + f(S) - f(S+1)\big].$$

Now consider μ and T satisfying (2.5) with (T, U) constructed on the same space as S, with $U \sim \mathcal{U}[0,1]$ and independent of S. Then, using that $P(S + T = 0) = 0$ and the mean value theorem for the second inequality and recalling definitions (1.21) and (1.22), we obtain

$$\begin{aligned}
&\big|E\big[h(S)\big] - E\big[h(D)\big]\big| \\
&\quad = \big|E\big[Sf'(S) - f'(S+U)\big]\big| = \big|E\big[\mu f'(S+T) - f'(S+U) + R_{f'}\big]\big| \\
&\quad \le \big|E\big[\mu f'(S+T) - f'(S+T)\big]\big| + \big|E\big[f'(S+T) - f'(S+U)\big]\big| + |R_{f'}| \\
&\quad \le \big\|f'\big\|_{(0,\infty)}|\mu - 1| + \big\|f''\big\|_{(0,\infty)} E|T-U| + |R_{f'}| \le |\mu - 1| + \frac{1}{2}E|T-U| + |R_{f'}|.
\end{aligned}$$

Now taking the infimum on the right hand side over all couplings (T, U) satisfying the conditions of the theorem yields

$$\big|E\big[h(S)\big] - E\big[h(D)\big]\big| \le |\mu - 1| + \frac{1}{2}\inf_{(T,U)} E|T-U| + |R_{f'_h}|,$$

where we have written $f = f_h$ to emphasize the dependence of $f(\cdot)$ on $h(\cdot)$. Taking supremum over $h \in \mathcal{H}_{1,1}$ first on the right, and then on the left now yields the result upon applying definition (1.9). □

Now we will demonstrate a few applications of Theorem 2.1. In all these examples the conditions that the variance of S is finite and that $S + T > 0$ almost surely are straightforward to check, and will not be mentioned further. For $n \ge 1$, let Ω_n denote the set of integers with no prime factor larger than p_n, and let Π_n be the distribution on Ω_n with mass function

$$\Pi_n(m) = \frac{1}{\pi_n m} \qquad \text{for } m \in \Omega_n,$$

where $\pi_n = \sum_{m \in \Omega_n} 1/m$ is the normalizing factor. One can check, see, for example, Proposition 1 in [26], that $M_n = \prod_{k=1}^n p_k^{X_k}$ has distribution Π_n, where $X_k \sim \mathrm{Geom}(1 - 1/p_k)$ are independent for $1 \le k \le n$; we remind the reader that we write $X \sim \mathrm{Geom}(p)$ when $P(X = m) = (1-p)^m p$ for $m \ge 0$. For $n \ge 1$, the random variable S_n as in (2.4) is therefore given by

$$S_n = \frac{1}{\log(p_n)} \sum_{k=1}^{n} X_k \log(p_k) = \frac{\log M_n}{\log(p_n)}. \tag{2.7}$$

Taking the mean, we find

$$\mu_n = E[S_n] = \frac{1}{\log(p_n)} \sum_{k=1}^{n} \frac{\log(p_k)}{p_k - 1}. \tag{2.8}$$

Now define the random variable I taking values in $\{1, \dots, n\}$, and independent of S_n, with mass function

$$P(I = k) = \frac{\log(p_k)}{(p_k - 1)\log(p_n)\mu_n} \qquad \text{for } k \in \{1, \dots, n\}. \tag{2.9}$$

The next lemma closely follows the arguments in Lemmas 3 and 5 of [3], which considered only weighted Bernoulli sums and obtained an $O(\log\log n/\log n)$ bound. In Lemmas 2.2, 2.3 and 2.4 below, we consider weighted sums of geometric, Bernoulli and Poisson random variables respectively, and obtain an $O(1/\log n)$ bound; see the supplement [12] for detailed proofs.

Lemma 2.2. *Let S_n be as in (2.7) with $X_1, \dots, X_n$ independent with $X_k \sim \mathrm{Geom}(1 - 1/p_k)$, μ_n as in (2.8), I with distribution given in (2.9) and independent of S_n and*

$$T_n = \frac{\log(p_I)}{\log(p_n)} \quad \text{and} \quad R_{n,\phi} = \frac{1}{\log(p_n)} \sum_{k=1}^{n} \frac{\log(p_k)}{p_k - 1} E\left[X_k\left(\phi\left(S_n + \frac{\log(p_k)}{\log(p_n)}\right) - \phi(S_n)\right)\right].$$

Then

$$E\big[S_n\phi(S_n)\big] = \mu_n E\big[\phi(S_n + T_n)\big] + R_{n,\phi} \qquad \text{for all } \phi \in \mathrm{Lip}_{1/2}.$$

Moreover

$$\sup_{\phi \in \mathrm{Lip}_{1/2}} |R_{n,\phi}| = O\left(\frac{1}{\log^2 n}\right) \quad \text{and} \quad \mu_n - 1 = O\left(\frac{1}{\log n}\right), \tag{2.10}$$

and there exists a coupling between $U \sim \mathcal{U}[0,1]$ and T_n with U independent of S_n, such that

$$E|T_n - U| = O\left(\frac{1}{\log n}\right).$$

Proof of Theorem 1.3. The upper bound follows directly from Theorem 2.1 upon invoking Lemma 2.2. We prove that the order of the bound is unimprovable by lower bounding the distance between S_n and D in the metric (1.9) by $|Eh(S_n) - Eh(D)| = |\mu_n - 1|$ for the choice $h(x) = x$, which is a member of $\mathcal{H}_{1,1}$, and showing that the order of that difference in (2.10) is optimal. For details, see [12]. □

For our next example, for $n \geq 1$ let Ω'_n denote the set of square-free integers whose largest prime factor is less than or equal to p_n and let Π'_n denote the distribution on Ω'_n with mass function

$$\Pi'_n(m) = \frac{1}{\pi'_n m} \qquad \text{for } m \in \Omega'_n,$$

where $\pi'_n = \sum_{m \in \Omega'_n} 1/m$ is the normalizing factor. We again consider S_n as in (2.7), here for $M_n = \prod_{k=1}^n p_k^{X_k}$ where $X_k \sim \mathrm{Ber}(1/(1+p_k))$ are independent for $1 \leq k \leq n$. One can check, see, for example, [13], that $M_n \sim \Pi'_n$. Following [3], let

$$\mu_n = E[S_n] = \frac{1}{\log(p_n)} \sum_{k=1}^n \frac{\log(p_k)}{1+p_k}. \tag{2.11}$$

The following lemma combines Lemmas 3 and 5 of [3]. By following tightly the same lines of argument in [3] the bounds, we obtain in (2.14) and (2.15) are $O(1/\log n)$ whereas [3] claims only the order $O(\log\log n/\log n)$.

Lemma 2.3. *Let S_n be as in (2.7) with $X_1, \ldots, X_n$ independent with $X_k \sim \mathrm{Ber}(1/(1+p_k))$. With μ_n as given in (2.11), let the random variable I take values in $\{1, \ldots, n\}$ with mass function*

$$P(I = k) = \frac{\log(p_k)}{(1+p_k)\log(p_n)\mu_n} \qquad \text{for } k \in \{1, \ldots, n\},$$

and be independent of $X_1, \ldots, X_n$. For

$$T_n = \frac{\log(p_I)}{\log(p_n)} - \frac{X_I \log(p_I)}{\log(p_n)}, \tag{2.12}$$

we have

$$E\big[S_n \phi(S_n)\big] = \mu_n E\big[\phi(S_n + T_n)\big] \qquad \text{for all } \phi \in \mathrm{Lip}_{1/2}. \tag{2.13}$$

Moreover,

$$\mu_n - 1 = O\left(\frac{1}{\log n}\right) \quad \text{and} \quad E\left|\frac{X_I \log(p_I)}{\log(p_n)}\right| = O\left(\frac{1}{\log^2 n}\right), \tag{2.14}$$

and there exists a coupling between a random variable $U \sim \mathcal{U}[0,1]$ and I with U independent of S_n such that

$$E\left|U - \frac{\log(p_I)}{\log(p_n)}\right| = O\left(\frac{1}{\log n}\right). \tag{2.15}$$

Proof of Theorem 1.4. The upper bound follows directly from Theorem 2.1 upon invoking Lemma 2.3 which gives that $R_\phi = 0$ for all $\phi \in \mathrm{Lip}_{1/2}$, and noting that with T_n and U as in (2.12) and (2.15) respectively,

$$E|T_n - U| \le E\left|\frac{X_I \log(p_I)}{\log(p_n)}\right| + E\left|U - \frac{\log(p_I)}{\log(p_n)}\right| = O\left(\frac{1}{\log n}\right),$$

by using (2.14) and (2.15) on these two terms, respectively. For a proof of the optimality of the order of the bound, see [12]. □

We also prove that these types of convergence results hold for S_n given in (2.7) when $X_k \sim \mathrm{Poi}(\lambda_k)$, $k \ge 1$ for certain sequences of positive real numbers $(\lambda_k)_{k\ge1}$. Here we take μ_n equal to the mean of S_n,

$$\mu_n = \frac{1}{\log(p_n)} \sum_{k=1}^{n} \lambda_k \log(p_k) \quad \text{and}$$
$$P(I = k) = \frac{\lambda_k \log(p_k)}{\log(p_n)\mu_n} \qquad \text{for } k \in \{1, \dots, n\}, \tag{2.16}$$

with I independent of S_n. Under this framework, we have the following construction of a variable having the size bias distribution of S_n.

Lemma 2.4. *For a sequence of positive real numbers $(\lambda_k)_{1\le k\le n}$ and independent random variables $X_1, \dots, X_n$ with $X_k \sim \mathrm{Poi}(\lambda_k)$, let*

$$S_n = \frac{1}{\log(p_n)} \sum_{k=1}^{n} X_k \log(p_k).$$

For μ_n as in (2.16) and $T_n = \log(p_I)/\log(p_n)$, where I is distributed as in (2.16) and is independent of S_n, we have

$$E\big[S_n\phi(S_n)\big] = \mu_n E\big[\phi(S_n + T_n)\big] \qquad \textit{for all } \phi \in \mathrm{Lip}_{1/2}.$$

The proof of this result can be found in the supplement [12], and is a direct consequence of the well known method for size biasing a sum of independent non-negative variables with finite mean, see, for example, [14].

We now present two applications of Lemma 2.4 with notation and assumptions as there.

Example 2.1. Let $\lambda_k = 1/(1 + p_k)$. As the mean of the X_k variables are the same here as in Lemma 2.3, μ_n and the distribution of I also correspond. Taking $U \sim \mathcal{U}[0, 1]$ independent of S_n, and coupling I and U similarly as in Lemma 2.3, we have that

$$|\mu_n - 1| = O\left(\frac{1}{\log n}\right) \quad \text{and} \quad E\left|U - \frac{\log(p_I)}{\log(p_n)}\right| = O\left(\frac{1}{\log n}\right).$$

Now, by Theorem 2.1 and Lemma 2.4, we obtain

$$d_{1,1}(S_n, D) \le \frac{C}{\log n}$$

for some universal constant C. One may show that the order of this bound is optimal by arguing as in the proof of Theorem 1.3.

Example 2.2. Let $p_0 = 1$ and and $\lambda_k = 1 - \log(p_{k-1})/\log(p_k)$ for $k \ge 1$. Then clearly $\mu_n = 1$ in (2.16). Now to obtain a coupling (T_n, U), we take $U \sim \mathcal{U}[0,1]$ independent of S_n, and define

$$I = k \qquad \text{if } \frac{\log(p_{k-1})}{\log(p_n)} \le U < \frac{\log(p_k)}{\log(p_n)} \text{ for } 1 \le k \le n.$$

Then by construction we have

$$P(I = k) = \frac{\lambda_k \log(p_k)}{\log(p_n)\mu_n} \qquad \text{for } 1 \le k \le n.$$

Conditioning on I, we have

$$E|T_n - U| = \sum_{k=1}^{n} P(I = k) E\left(\left|\frac{\log(p_k)}{\log(p_n)} - U\right| \middle| I = k\right) \le \sum_{k=1}^{n} P(I = k)\left|\frac{\log(p_{k-1})}{\log(p_n)} - \frac{\log(p_k)}{\log(p_n)}\right|.$$

Now using that $p_k/p_{k-1} \le 2$ by Bertrand's postulate (see, e.g., [29]) for all $k \ge 1$, we obtain

$$E|T_n - U| \le \frac{\log(2)}{\log(p_n)}.$$

Hence from Theorem 2.1 with $\mu_n = 1$ and $R_\phi = 0$ for all $\phi \in \mathrm{Lip}_{1/2}$, we have

$$d_{1,1}(S_n, D) \le \frac{\log(2)}{2\log(p_n)} \le \frac{C}{\log n}$$

for some universal constant C.

Following the distribution of a draft of this manuscript, [5] pointed out that the approach in [4] may be used to obtain bounds in the Wasserstein-1 metric for some results in this section.

3. Perpetuities and the $\mathcal{D}_{\theta,s}$ family, simulations and distributional bounds

In this section, we develop the extension of the generalized Dickman distribution to the $\mathcal{D}_{\theta,s}$ family for $\theta > 0$ and a function $s : [0,\infty) \to [0,\infty)$. As detailed in the Introduction, the recursion (1.2) associated with the $\mathcal{D}_\theta$ family can be interpreted as giving the successive values of a

Vervaat perpetuity under the assumption that the utility function is the identity. More generally, with utility function $s(\cdot)$, one obtains the recursion

$$s(W_{n+1}) = U_n^{1/\theta} s(W_n + 1) \qquad \text{for } n \geq 0, \tag{3.1}$$

where $U_n, n \geq 0$ are independent and have the $\mathcal{U}[0,1]$ distribution, U_n is independent of W_n, and W_0 has some given initial distribution. In Section 3.1, under Condition 3.1 below on $s(\cdot)$, we prove Theorem 3.3 that shows that the distributional fixed points $\mathcal{D}_{\theta,s}$ of (3.1) exist and are unique. When $s(\cdot)$ is invertible, at it is under Condition 3.1 below, we may write (3.1) as

$$W_{n+1} = s^{-1}\bigl(U_n^{1/\theta} s(W_n + 1)\bigr) \qquad \text{for } n \geq 0. \tag{3.2}$$

In Section 3.2, we provide distributional bounds for approximation of the $\mathcal{D}_{\theta,s}$ distribution. Using direct coupling, Corollary 3.1 gives a bound on how well the utility $s(W_n)$ in (3.1) approximates the utility of its limit $D_{\theta,s}$. Next, Theorem 3.4 extends the main Wasserstein bound (1.3) of [20] to

$$d_1(W, D_{\theta,s}) \leq (1-\rho)^{-1} d_1\bigl(W^*, W\bigr) \qquad \text{where } W^* =_d s^{-1}\bigl(U^{1/\theta} s(W+1)\bigr) \tag{3.3}$$

for $U \sim \mathcal{U}[0,1]$, independent of W. The constant ρ is defined in (3.11) as a uniform bound on an integral involving (θ, s) given by (3.10). However, [8] shows that this quantity can be interpreted in terms of the Markov chain (3.2) and its properties connected to those of its transition operator $(Ph)(x) = E[h(s^{-1}(U^{1/\theta} s(x+1)))]$ in this, and some more general, cases. In particular, for $h \in \mathrm{Lip}_1$, ρ is a bound on the essential supremum norm of the derivative of the transition operator. Though linear stochastic recursions are ubiquitous and are well known to be highly tractable, this special class of Markov chains, despite its non-linear transitions, seems also amenable to deeper analysis.

We apply the inequality (3.3) in Corollary 3.2 to obtain a bound on the Wasserstein distance between the iterates W_n of (3.2) and $D_{\theta,s}$. Finally in Section 3.3, we give a few examples of some new distributions that arise as a result of utility functions that appear in the economics literature.

3.1. Existence and uniqueness of the $\mathcal{D}_{\theta,s}$ distribution

In the following, we use the terms increasing and decreasing in the non-strict sense. Let $\leq_{\mathrm{st}}$ denote inequality between random variables in the stochastic order. The proofs of all the claims in this subsection can be found in the supplement [12].

Lemma 3.1. *Let $\theta > 0$ and $s : [0,\infty) \to [0,\infty)$ satisfy*

$$s(x+1) \leq s(x) + 1 \qquad \textit{for all } x \geq 0, \tag{3.4}$$

let W_0 be a given non-negative random variable and let $\{W_n, n \geq 1\}$ be generated by recursion (3.1). Then

$$s(W_{n+1}) \leq U_n^{1/\theta}\bigl(s(W_n) + 1\bigr) \qquad \textit{for all } n \geq 0. \tag{3.5}$$

If in addition $s(W_0) \le_{\mathrm{st}} D_\theta$, *then*

$$s(W_n) \le_{\mathrm{st}} D_\theta \qquad \text{for all } n \ge 0. \tag{3.6}$$

Theorem 3.3, showing the existence and uniqueness of the fixed point $\mathcal{D}_{\theta,s}$ to (1.15), requires the following condition to hold on the utility function $s(\cdot)$.

Condition 3.1. The function $s : [0, \infty) \to [0, \infty)$ is continuous, strictly increasing with $s(0) = 0$ and $s(1) = 1$, and satisfies

$$s(x + 1) \le s(x) + 1 \qquad \text{for all } x \ge 0 \tag{3.7}$$

and

$$\bigl|s(x + 1) - s(y + 1)\bigr| \le \bigl|s(x) - s(y)\bigr| \qquad \text{for all } x, y \ge 0. \tag{3.8}$$

The following result, proven by constructing a direct coupling, shows that choice of the starting distribution in (3.1) has vanishing effect asymptotically as measured in the d_1 Wasserstein norm.

Lemma 3.2. *Let* $\theta > 0$ *and Condition* 3.1 *be in force. Let* W_0 *and* V_0 *be given non-negative random variables such that the means of* $s(W_0)$ *and* $s(V_0)$ *are finite. For* $n \ge 1$ *let* $s(V_n)$ *and* $s(W_n)$ *have distributions as specified in* (3.1). *Then* $s(W_n)$ *and* $s(V_n)$ *have finite mean for all* $n \ge 0$, *and*

$$d_1\bigl(s(W_n), s(V_n)\bigr) \le \left(\frac{\theta}{\theta + 1}\right)^n d_1\bigl(s(W_0), s(V_0)\bigr) \qquad \text{for all } n \ge 0. \tag{3.9}$$

Define the generalized inverse of an increasing function $s : [0, \infty) \to [0, \infty)$ as

$$s^-(x) = \inf\bigl\{y : s(y) \ge x\bigr\}$$

with the convention that $\inf \varnothing = \infty$. In particular, for X a random variable we consider $s^-(X)$ as a random variable taking values in the extended real line. When writing the stochastic order relation $V \le_{\mathrm{st}} W$ between two extended valued random variables, we mean that $P(V \ge t) \le P(W \ge t)$ holds for all t in the extended real line. Note that $s^-(\cdot)$ and $s^{-1}(\cdot)$ coincide on the range of $s(\cdot)$ when $s(\cdot)$ is continuous and strictly increasing.

Theorem 3.3. *Let* $\theta > 0$ *and* $s(\cdot)$ *satisfy Condition* 3.1. *Then there exists a unique distribution* $\mathcal{D}_{\theta,s}$ *for a random variable* $D_{\theta,s}$ *such that* $s(D_{\theta,s})$ *has finite mean and satisfies* $D_{\theta,s} =_d D^*_{\theta,s}$, *with* $D^*_{\theta,s}$ *given by* (1.15). *In addition*, $D_{\theta,s} \le_{\mathrm{st}} s^-(D_\theta)$.

3.2. Distributional bounds for $\mathcal{D}_{\theta,s}$ approximation and simulations

In this section, we study the accuracy of recursive methods to approximately sample from the $\mathcal{D}_{\theta,s}$ family, starting with the following simple corollary to Lemma 3.2 that gives a bound on how well the utility $s(W_n)$, satisfying the recursion (3.1), approximates the long term utility of the fixed point.

Corollary 3.1. *Let* $\theta > 0$ *and Condition* 3.1 *be in force. Then* $s(W_n)$ *given by* (3.1) *satisfies*

$$d_1\big(s(W_n), s(D_{\theta,s})\big) \le \left(\frac{\theta}{\theta+1}\right)^n d_1\big(s(W_0), s(D_{\theta,s})\big) \qquad \text{for all } n \ge 0.$$

Proof. The result follows from (3.9) of Lemma 3.2 by taking $V_0 =_d D_{\theta,s}$ and noting that $D_{\theta,s}$ is fixed by the transformation (3.3) so that $s(V_n) =_d s(D_{\theta,s})$ for all n. □

Corollary 3.1 depends on the direct coupling used to prove Lemma 3.2, which constructs the variables $s(W_n)$ and $s(V_n)$ on the same space. Theorem 3.4 below gives a bound for when a non-negative random variable W is used to approximate the distribution of $D_{\theta,s}$. Though direct coupling can still be used to obtain bounds such as those in Theorem 3.4 for the $\mathcal{D}_\theta$ family, doing so is no longer possible for the more general $\mathcal{D}_{\theta,s}$ family as iterates of (3.2) can no longer be written explicitly when $s(\cdot)$ is non-linear. Theorem 3.4 below provides a Wasserstein bound between $D_{\theta,s}$ and W assuming certain natural conditions on the function $s(\cdot)$.

For $\theta > 0$, suppressed in the notation, and $x > 0$ such that $s'(x)$ exists, let

$$I(x) = \frac{\theta s'(x)}{s^{\theta+1}(x)} \int_0^x s^\theta(v)\,dv. \tag{3.10}$$

For $S \subset [0,\infty)$, we say a function $f : [0,\infty) \to [0,\infty)$ is locally absolutely continuous on S if it is absolutely continuous when restricted to any compact sub-interval of S. Unless otherwise stated, locally absolutely continuity will mean over the domain of $f(\cdot)$.

Theorem 3.4. *Let* $\theta > 0$ *and* $s : [0,\infty) \to [0,\infty)$ *satisfying Condition* 3.1 *be locally absolutely continuous on* $[0,\infty)$ *and such that* $E[D_{\theta,s}] < \infty$. *With* $I(\cdot)$ *as in* (3.10), *if there exists* $\rho \in [0,1)$ *such that*

$$\|I\|_\infty \le \rho, \tag{3.11}$$

then for any non-negative random variable W *with finite mean*,

$$d_1(W, D_{\theta,s}) \le (1-\rho)^{-1} d_1\big(W^*, W\big). \tag{3.12}$$

In the special case $s(x) = x$, $\|I\|_\infty = \theta/(\theta+1) \in [0,1)$, *and one may take* ρ *equal to this value.*

Remark 3.1. Note that $E[s^{-1}(D_\theta)] < \infty$ implies $E[D_{\theta,s}] < \infty$ as $D_{\theta,s} \le_{\mathrm{st}} s^{-1}(D_\theta)$ by Theorem 3.3.

Remark 3.2. By a simple argument, similar to the one in Section 3 of [20], for $\theta > 0$ and $s : [0,\infty) \to [0,\infty)$ satisfying Condition 3.1, (3.15) below and $E[D_{\theta,s}] < \infty$, for any non-negative random variable W with finite mean, we have

$$d_1(W, D_{\theta,s}) \le (1+\theta) d_1\big(W^*, W\big)$$

so that (3.12) holds with $\rho = \theta/(\theta+1)$.

The use of Stein's method in Theorem 3.4 does not require that $s(\cdot)$ satisfy (3.15) but does need $s(\cdot)$ to be locally absolutely continuous. In addition, the alternative approach in [20] has no scope for improvement in terms of finding the best constant ρ; Example 3.2 presents a case where taking $\rho = \theta/(\theta+1)$ is not optimal. Theorem 3.7 below gives a verifiable criteria by which one can show when the canonical choice $\rho = \theta/(\theta+1)$ is not improvable.

We will prove Theorem 3.4 using Stein's method in Section 4. Here, we provide the following corollary applicable for the simulation of $\mathcal{D}_{\theta,s}$ distributed random variables. Note that when $s(\cdot)$ is strictly increasing and continuous, for W independent of $U \sim \mathcal{U}[0,1]$ the transform W^* as given by (1.15) satisfies

$$W^* =_d s^{-1}\bigl(U^{1/\theta} s(W+1)\bigr) \le W+1. \tag{3.13}$$

Corollary 3.2. *Let* $s:[0,\infty) \to [0,\infty)$ *be as in Theorem 3.4 and let* $\{W_n, n \ge 1\}$ *be generated by (3.2) with* W_0 *non-negative and* $EW_0 < \infty$, *independent of* $\{U_n, n \ge 0\}$. *If* $\rho \in [0,1)$ *exists satisfying (3.11), then*

$$d_1(W_n, D_{\theta,s}) \le (1-\rho)^{-1} d_1(W_{n+1}, W_n). \tag{3.14}$$

Moreover, if $s(\cdot)$ *satisfies*

$$\bigl|s^{-1}\bigl(as(x)\bigr) - s^{-1}\bigl(as(y)\bigr)\bigr| \le a|x-y| \qquad \text{for } a \in [0,1] \text{ and } x, y \ge 1, \tag{3.15}$$

then

$$d_1(W_n, D_{\theta,s}) \le (1-\rho)^{-1}\left(\frac{\theta}{\theta+1}\right)^n d_1(W_1, W_0). \tag{3.16}$$

When $W_0 = 0$,

$$d_1(W_n, D_{\theta,s}) \le (1-\rho)^{-1}\left(\frac{\theta}{\theta+1}\right)^n E\bigl[s^{-1}\bigl(U^{1/\theta}\bigr)\bigr], \tag{3.17}$$

and in the particular the case of the generalized Dickman $\mathcal{D}_\theta$ *family,*

$$d_1(W_n, D_\theta) \le \theta\left(\frac{\theta}{\theta+1}\right)^n. \tag{3.18}$$

Proof. Identity (3.2), the inequality in (3.13) and induction show that $W_n \le W_0 + n$, and hence $EW_n < \infty$, for all $n \ge 0$. Inequality (3.14) now follows from Theorem 3.4 noting from (1.15) that $W_n^* =_d W_{n+1}$ for all $n \ge 0$.

To show (3.16), recalling that the bound (1.6) is achieved for real valued random variables, for every $n \ge 1$ we may construct W'_{n-1} and V'_n independent of U_n such that $W'_{n-1} =_d W_{n-1}$, $V'_n =_d W_n$ and $E|V'_n - W'_{n-1}| = d_1(W_n, W_{n-1})$. Now letting

$$W''_n = s^{-1}\bigl(U_n^{1/\theta} s\bigl(W'_{n-1}+1\bigr)\bigr) \quad \text{and} \quad V''_{n+1} = s^{-1}\bigl(U_n^{1/\theta} s\bigl(V'_n+1\bigr)\bigr)$$

we have $W_n'' =_d W_n$ and $V_{n+1}'' =_d W_{n+1}$. Thus, using (1.6) followed by (3.15) we have

$$\begin{aligned}d_1(W_{n+1}, W_n) &\le E\big|V_{n+1}'' - W_n''\big| \\ &= E\big|s^{-1}\big(U_n^{1/\theta} s\big(V_n' + 1\big)\big) - s^{-1}\big(U_n^{1/\theta} s\big(W_{n-1}' + 1\big)\big)\big| \\ &\le E\big[U_n^{1/\theta}\big|V_n' - W_{n-1}'\big|\big] = \frac{\theta}{\theta+1} d_1(W_n, W_{n-1}).\end{aligned}$$

Induction now yields

$$d_1(W_{n+1}, W_n) \le \left(\frac{\theta}{\theta+1}\right)^n d_1(W_1, W_0)$$

and applying (3.14) we obtain (3.16).

Inequality (3.17) now follows from (3.16) noting in this case, using $s(1) = 1$, that $(W_0, W_1) = (0, s^{-1}(U_0^{1/\theta}))$, and (3.18) is now achieved from (3.17) by taking ρ to be $\theta/(\theta + 1)$, as provided by Theorem 3.4 when $s(x) = x$. □

In the remainder of this subsection, in Lemma 3.6 we present some general and easily verifiable conditions on $s(\cdot)$ for the satisfaction of (3.15), and in Theorem 3.7 ones under which the integral bound $\|I\|_\infty \le \rho$ in (3.11) holds with $\rho \in [0, 1)$. Lastly we show our bounds are equivalent to what can be obtained by a direct coupling method in the cases where the latter is available.

Condition 3.2. The function $s : [0, \infty) \to [0, \infty)$ is continuous at 0, strictly increasing with $s(0) = 0$ and $s(1) = 1$, and concave.

Lemma 3.5. *If a function $f : [0, \infty) \to [0, \infty)$ is increasing, continuous at 0 and locally absolutely continuous on $(0, \infty)$, then it is locally absolutely continuous on its domain.*

The proof of Lemma 3.5 is straightforward, see [12] for details.

Lemma 3.6. *If $s : [0, \infty) \to [0, \infty)$ satisfies Condition 3.2, then it is locally absolutely continuous on $[0, \infty)$, satisfies Condition 3.1 and*

$$\big|s^{-1}\big(as(y)\big) - s^{-1}\big(as(x)\big)\big| \le a|y - x| \qquad \text{for all } x, y \ge 0 \text{ and } a \in [0, 1]. \tag{3.19}$$

Proof. First, since $s(\cdot)$ is concave, it is locally absolutely continuous on $(0, \infty)$. Thus, by Lemma 3.5, $s(\cdot)$ is locally absolutely continuous on its domain. Next we show $s(\cdot)$ is subadditive, that is, that

$$s(x + y) \le s(x) + s(y) \qquad \text{for } x, y \ge 0. \tag{3.20}$$

Taking $x, y \ge 0$, we may assume both x and y are non-zero as (3.20) is trivial otherwise since $s(0) = 0$. By concavity,

$$\frac{y}{x+y}s(0) + \frac{x}{x+y}s(x+y) \le s(x) \quad \text{and} \quad \frac{x}{x+y}s(0) + \frac{y}{x+y}s(x+y) \le s(y).$$

Since $s(0) = 0$, adding these two inequalities yield (3.20). Taking $y = 1$ and using $s(1) = 1$ we obtain (3.7). Next, the local absolute continuity and concavity of $s(\cdot)$ on $[0, \infty)$ imply that it is almost everywhere differentiable on this domain, with $s'(\cdot)$ decreasing almost everywhere. Thus for $x \ge y \ge 0$, we have

$$\begin{aligned} s(x+1) - s(x) &= \int_x^{x+1} s'(u)\,du \le \int_x^{x+1} s'(u+y-x)\,du \\ &= \int_y^{y+1} s'(u)\,du = s(y+1) - s(y), \end{aligned}$$

which together with the fact that $s(\cdot)$ is increasing implies (3.8). Hence, $s(\cdot)$ satisfies Condition 3.1.

Lastly, we show that $s(\cdot)$ satisfies (3.19). Since $s(0) = 0$ the inequality is trivially satisfied for $a = 0$, so fix some $a \in (0, 1]$. Again as the result is trivial otherwise, we may take $x \ne y$; without loss, let $0 \le x < y$. The inverse function $r(\cdot) = s^{-1}(\cdot)$ is continuous at zero and convex on the range S of $s(\cdot)$, a possibly unbounded convex subset $[0, \infty)$ that includes the origin. Letting $u = s(x)$ and $v = s(y)$, as $s(\cdot)$, and hence $r(\cdot)$, are strictly increasing and $x \ne y$, inequality (3.19) may be written

$$r(av) - r(au) \le a\big(r(v) - r(u)\big) \quad \text{or equivalently} \quad \frac{r(av) - r(au)}{av - au} \le \frac{r(v) - r(u)}{v - u}, \tag{3.21}$$

where all arguments of $r(\cdot)$ in (3.21) lie in S, it being a convex set containing $\{0, u, v\}$.

The second inequality in (3.21) follows from the following slightly more general one that any convex function $r : [0, \infty) \to [0, \infty)$ which is continuous at 0 satisfies by virtue of its local absolute continuity and a.e. derivative $r'(\cdot)$ being increasing: if (u_1, v_1) and (u_2, v_2) are such that $u_1 \ne v_1$, $u_1 \le u_2$ and $v_1 \le v_2$, and all these values lie in the range of $r(\cdot)$, then

$$\begin{aligned} \frac{r(v_1) - r(u_1)}{v_1 - u_1} &= \frac{1}{v_1 - u_1}\int_{u_1}^{v_1} r'(w)\,dw = \int_0^1 r'\big(u_1 + (v_1 - u_1)w\big)\,dw \\ &\le \int_0^1 r'\big(u_2 + (v_2 - u_2)w\big)\,dw = \frac{1}{v_2 - u_2}\int_{u_2}^{v_2} r'(w)\,dw = \frac{r(v_2) - r(u_2)}{v_2 - u_2}, \end{aligned}$$

as one easily has that $u_1 + (v_1 - u_1)w \le u_2 + (v_2 - u_2)w$ for all $w \in [0, 1]$. □

When the function $s(\cdot)$ is nice enough, we can actually say more about the constant ρ in (3.11) of Theorem 3.4.

Theorem 3.7. *Assume that $\theta > 0$ and $s : [0, \infty) \to [0, \infty)$ is concave and continuous at 0. Then with $I(x)$ as given in (3.10),*

$$\|I\|_\infty \le \frac{\theta}{\theta+1}. \tag{3.22}$$

If moreover $s(\cdot)$ is strictly increasing with $s(0) = 0$ and $\lim_{n\to\infty} s'(x_n) < \infty$ for some sequence of distinct real numbers $x_n \downarrow 0$ in the domain of $s'(\cdot)$, then

$$\|I\|_\infty = \frac{\theta}{\theta+1}. \tag{3.23}$$

Proof. Since $s(\cdot)$ is concave and continuous at 0, it is locally absolutely continuous with $s'(\cdot)$ decreasing almost everywhere on $[0, \infty)$. Since $u^{\theta+1}$ is Lipschitz on any compact interval, by composition, $s^{\theta+1}(\cdot)$ is absolutely continuous on $[0, x]$ for any $x \ge 0$, and thus for almost every x,

$$\begin{aligned}\frac{(\theta+1)I(x)}{\theta} &= \frac{(\theta+1)s'(x)}{s^{\theta+1}(x)} \int_0^x s^\theta(v)\,dv \le \frac{1}{s^{\theta+1}(x)} \int_0^x (\theta+1)s^\theta(v)s'(v)\,dv \\ &= \frac{s^{\theta+1}(x) - s^{\theta+1}(0)}{s^{\theta+1}(x)} \le 1,\end{aligned}$$

proving (3.22).

To prove the second claim, first note that $0 < \lim_{n\to\infty} s'(x_n) < \infty$, the existence of the limit and second inequality holding by assumption, and the first inequality holding as $s(\cdot)$ is strictly increasing and $s'(\cdot)$ is decreasing almost everywhere.

Thus, in the second equality using a version of the Stolz–Cesàro theorem [36] adapted to accommodate $s^{\theta+1}(x_n)$ decreasing to zero,

$$\begin{aligned}\lim_{n\to\infty} I(x_n) &= \theta \lim_{n\to\infty} s'(x_n) \lim_{n\to\infty} \frac{\int_0^{x_n} s^\theta(v)\,dv}{s^{\theta+1}(x_n)} \\ &= \theta \lim_{n\to\infty} s'(x_n) \lim_{n\to\infty} \frac{\int_{x_{n+1}}^{x_n} s^\theta(v)\,dv}{s^{\theta+1}(x_n) - s^{\theta+1}(x_{n+1})} \\ &= \theta \lim_{n\to\infty} s'(x_n) \lim_{n\to\infty} \frac{\int_{x_{n+1}}^{x_n} s^\theta(v)\,dv}{(\theta+1)\int_{x_{n+1}}^{x_n} s^\theta(v)s'(v)\,dv} \\ &= \frac{\theta}{\theta+1} \lim_{n\to\infty} s'(x_n) \lim_{n\to\infty} \frac{1}{s'(x_n)} = \frac{\theta}{\theta+1},\end{aligned}$$

where the penultimate equality follows from the fact that

$$\lim_{n\to\infty} \frac{1}{s'(x_n)} = \lim_{n\to\infty} \frac{1}{s'(x_{n+1})} \le \lim_{n\to\infty} \frac{\int_{x_{n+1}}^{x_n} s^\theta(v)\,dv}{\int_{x_{n+1}}^{x_n} s^\theta(v)s'(v)\,dv} \le \lim_{n\to\infty} \frac{1}{s'(x_n)}$$

and hence

$$\|I\|_\infty \geq \frac{\theta}{\theta+1},$$

which together with (3.22) proves (3.23). □

The bound (3.18) of Corollary 3.2 is obtained by specializing results for the $\mathcal{D}_{\theta,s}$ family, proven using the tools of Stein's method, to the case where $s(x) = x$. For this special case, letting $V_j = U_j^{1/\theta}$ for $j \geq 0$, the iterates of the recursion (3.2), starting at $W_0 = 0$, can be written explicitly as

$$W_n = \sum_{k=0}^{n-1} \prod_{j=k}^{n-1} V_j,$$

allowing one to obtain bounds using direct coupling. Interestingly, the results obtained by both methods agree, as seen as follows. First, we show

$$W_n =_d Y_n \qquad \text{where } Y_n = \sum_{k=0}^{n-1} \prod_{j=0}^{k} V_j, \quad \text{and} \quad Y_\infty \sim D_\theta \qquad \text{where } Y_\infty = \sum_{k=0}^{\infty} \prod_{j=0}^{k} V_j.$$

The first claim is true since for every $n \geq 1$,

$$(V_0, \dots, V_{n-1}) =_d (V_{n-1}, \dots, V_0).$$

For the second claim, note that the limit Y_∞ exists almost everywhere and has finite mean by monotone convergence. Now using definition (1.1), with $U_{-1} \sim \mathcal{U}[0,1]$ independent of $U_0, U_1, \dots$ and setting $V_{-1} = U_{-1}^{1/\theta}$, we have

$$\begin{aligned}
Y_\infty^* = U_{-1}^{1/\theta}(Y_\infty + 1) &= V_{-1}\left(\sum_{k=0}^{\infty} \prod_{j=0}^{k} V_j + 1\right) \\
&= \sum_{k=0}^{\infty} \prod_{j=-1}^{k} V_j + V_{-1} = \sum_{k=-1}^{\infty} \prod_{j=-1}^{k} V_j \\
&= \sum_{k=0}^{\infty} \prod_{j=0}^{k} V_{j-1} =_d \sum_{k=0}^{\infty} \prod_{j=0}^{k} V_j = Y_\infty.
\end{aligned}$$

Hence $Y_\infty \sim \mathcal{D}_\theta$. As (Y_n, Y_∞) is a coupling of a variable with the W_n distribution to one with the $\mathcal{D}_\theta$ distribution, by (1.6) we obtain

$$d_1(W_n, D_\theta) = d_1(Y_n, Y_\infty) \leq E|Y_\infty - Y_n| = E\left(\sum_{k=n}^{\infty} \prod_{j=0}^{k} V_j\right)$$

$$= \sum_{k=n}^{\infty} \left(\frac{\theta}{\theta+1} \right)^{k+1} = \theta \left(\frac{\theta}{\theta+1} \right)^n,$$

in agreement with (3.18).

3.3. Examples

We now consider three new distributions that arise as special cases of the $\mathcal{D}_{\theta,s}$ family. Expected Utility (EU) theory has long been considered as an acceptable paradigm for decision making under uncertainty by researchers in both economics and finance, see e.g. [17]. To obtain tractable solutions to many problems in economics, one often restricts the EU criterion to a certain class of utility functions, which includes in particular the ones in Examples 3.1 and 3.3. In these two examples, we apply the bounds provided in Corollary 3.2 for the simulation of the limiting distributions these functions give rise to via the recursion (3.2) with say, $W_0 = 0$. For each example, we will verify Condition 3.2, implying Condition 3.1 by Lemma 3.6, and hence existence and uniqueness of $D_{\theta,s}$.

Example 3.1. The exponential utility function $u(x) = 1 - e^{-\alpha x}$ is the only model, up to linear transformations, exhibiting constant absolute risk aversion (CARA), see [17]. Since utility is unique up to linear transformations, we consider its scaled version

$$s_\alpha(x) = \frac{1 - e^{-\alpha x}}{1 - e^{-\alpha}} \qquad \text{for } x \geq 0$$

characterized by a parameter $\alpha > 0$. Clearly $s_\alpha(\cdot)$ is continuous at 0, strictly increasing with $s_\alpha(0) = 0$ and $s_\alpha(1) = 1$ and concave. Since $\lim_{x\downarrow 0} s'_\alpha(x) = \alpha(1-e^{-\alpha})^{-1} \in (0, \infty)$, for all $\theta > 0$, by (3.23) of Theorem 3.7, one can take ρ to be $\theta/(\theta + 1)$ and not strictly smaller, and (3.17) of Corollary 3.2 yields

$$d_1(W_n, D_{\theta,s_\alpha}) \leq \theta \left(\frac{\theta}{\theta+1} \right)^{n-1} \qquad \text{for all } n \geq 0,$$

using that $0 \leq s_\alpha^{-1}(U^{1/\theta}) \leq s_\alpha^{-1}(1) = 1$ almost surely.

Letting $W_\alpha \sim D_{\theta,s_\alpha}$ it is easy to verify that

$$s_\alpha(W_\alpha) =_d U^{1/\theta} s_\alpha(W_\alpha + 1) = U^{1/\theta}\big(1 + e^{-\alpha} s_\alpha(W_\alpha)\big).$$

Using this identity, that Theorem 3.3 gives $0 \leq s_\alpha(W_\alpha) \leq_{\text{st}} D_\theta$ for all $\alpha > 0$, and that $\lim_{\alpha\downarrow 0} s_\alpha(x) = x$ for all $x \geq 0$ one can show that W_α converges to D_θ as $\alpha \downarrow 0$. Hence, now setting $s_0(x) = x$, the family of models $D_{\theta,s_\alpha}, \alpha \geq 0$ is parameterized by a tuneable values of $\alpha \geq 0$ whose value may be chosen depending on a desired level of risk aversion, including the canonical $\alpha = 0$ case where utility is linear.

Example 3.2. Here we show how standard Vervaat perpetuity models can be seen to assume an implicit concave utility function, and how uncertainty in these utilities can be accommodated using the new families we introduce. Indeed, letting $\theta = 1$ in (1.14) and then $s_\theta(x) = x^\theta, \theta \in (0,1]$, it is easy to see that $\mathcal{D}_{1,s_\theta} = \mathcal{D}_\theta$. To model situations where these utilities are themselves subject to uncertainty, we may let A be a random variable supported in $(0,1]$ and consider the mixture $s(x) = E[s_A(x)]$.

More formally, for some $0 < a \le 1$, let μ be a probability measure on the interval $(0, a]$, and define

$$s(x) = \int_0^a s_\alpha(x)\, d\mu(\alpha).$$

Since $0 < a \le 1$, each $s_\alpha(\cdot)$ is concave and satisfies Condition 3.2 and hence so does $s(\cdot)$. By (3.22) of Theorem 3.7, for the family $\mathcal{D}_{\theta,s}$ one can take $\rho = \theta/(\theta+1)$.

Fix $l > 0$. For $x \ge l$, note that $\partial x^\alpha/\partial x = \alpha x^{\alpha-1} \le \alpha l^{\alpha-1}$ which is bounded and hence μ-integrable on $[0, a]$. Thus by dominated convergence, since $l > 0$ is arbitrary, we obtain

$$s'(x) = \int_0^a \frac{\partial x^\alpha}{\partial x}\, d\mu(\alpha) = \int_0^a \alpha x^{\alpha-1}\, d\mu(\alpha) \qquad \text{for all } x > 0. \tag{3.24}$$

Now note that for $a < 1$, $\lim_{x\downarrow 0} s'(x)$ diverges to infinity, and hence (3.23) of Theorem 3.7 cannot be invoked. We show, in fact, that one may obtain a bound better than $\theta/(\theta+1)$ in this case.

Taking $\theta = 1$ and computing $I(x)$ directly from (3.10), using (3.24) for the first equality and Fubini's theorem for the second, we have

$$\begin{aligned}
I(x) &= \frac{[\int_0^a \alpha x^{\alpha-1}\, d\mu(\alpha)][\int_0^x \int_0^a v^\alpha\, d\mu(\alpha)\, dv]}{[\int_0^a x^\alpha\, d\mu(\alpha)]^2} = \frac{[\int_0^a \alpha x^{\alpha-1}\, d\mu(\alpha)][\int_0^a \frac{x^{\alpha+1}}{\alpha+1}\, d\mu(\alpha)]}{[\int_0^a x^\alpha\, d\mu(\alpha)]^2} \\
&= \frac{[\int_0^a \int_0^a \frac{\alpha}{\beta+1} x^{\alpha+\beta}\, d\mu(\alpha)\, d\mu(\beta)]}{[\int_0^a x^\alpha\, d\mu(\alpha)]^2} = \frac{[\int_0^a \int_0^a \frac{1}{2}(\frac{\alpha}{\beta+1} + \frac{\beta}{\alpha+1}) x^{\alpha+\beta}\, d\mu(\alpha)\, d\mu(\beta)]}{\int_0^a \int_0^a x^{\alpha+\beta}\, d\mu(\alpha)\, d\mu(\beta)} \\
&\le \sup_{\alpha,\beta\in[0,a]} \frac{1}{2}\left(\frac{\alpha}{\beta+1} + \frac{\beta}{\alpha+1}\right).
\end{aligned}$$

Taking $0 \le \alpha \le \beta \le a$, the reverse case being handled similarly, using the simple fact that

$$(\beta - \alpha)^2 \le \beta - \alpha \qquad \text{for } 0 \le \alpha \le \beta \le 1$$

shows that for $0 \le \alpha \le \beta \le a$,

$$\frac{\alpha}{\beta+1} + \frac{\beta}{\alpha+1} \le \frac{2\beta}{\beta+1} \le \frac{2a}{a+1}$$

and hence one can take $\rho = a/(a+1)$. Note that when $a = 1/2$, say, we obtain the upper bound $\rho = 1/3$, whereas the bound (3.22) of Theorem 3.7 gives $1/2$ when $\theta = 1$. Taking μ to be unit mass at 1 yields $\rho = 1/2$ which recovers the bound on ρ for the standard Dickman derived in [20], and as given in Theorem 3.4, for the value $\theta = 1$.

Example 3.3. The logarithm $u(x) = \log x$ is another commonly used utility function as it exhibits constant relative risk aversion (CRRA) which often simplifies many problems encountered in macroeconomics and finance, see [17]. Applying a shift to make it non-negative, let

$$s(x) = \log(x+1)/\log 2 \qquad \text{for } x \geq 0.$$

Clearly $s(\cdot)$ satisfies Condition 3.2. To apply Corollary 3.2 it remains to compute an upper bound ρ on the integral in (3.10). Now since $\lim_{x\downarrow 0} s'(x) < \infty$, by (3.23) of Theorem 3.7, we may take $\rho = \theta/(\theta+1)$. Noting $s^{-1}(x) = 2^x - 1$, simulating from this distribution by the recursion

$$W_{n+1} = (W_n + 2)^{U_n^{1/\theta}} - 1 \qquad \text{for } n \geq 1 \text{ with initial value } W_0 = 0,$$

inequality (3.17) of Corollary 3.2 yields

$$d_1(W_n, D_{\theta,s}) \leq \theta \left(\frac{\theta}{\theta+1} \right)^{n-1} \qquad \text{for all } n \geq 0,$$

using that $0 \leq s^{-1}(U^{1/\theta}) = 2^{U^{1/\theta}} - 1 \leq 1$ almost surely.

4. Smoothness bounds

In this section, we turn to proving Theorem 4.7 from which Theorem 3.4 readily follows. We develop the necessary tools building on [19]. For notational simplicity, in this section given (θ, s), let

$$t(x) = s^\theta(x) \qquad \text{for all } x \geq 0. \tag{4.1}$$

Throughout this section $t : [0, \infty) \to [0, \infty)$ will be strictly increasing and hence almost everywhere differentiable by Lebesgue's Theorem, see, for example, Section 6.2 of [32], inducing the measure ν satisfying $d\nu/dv = t'(v)$ on $[0, \infty)$, where v is Lebesgue measure. For $h \in L^1([0, a], \nu)$ for some $a > 0$, define the averaging operator

$$A_x h = \frac{1}{t(x)} \int_0^x h(v) t'(v)\, dv \qquad \text{for } x \in (0, a] \quad \text{and} \quad A_0 h = h(0)\mathbb{1}\big(t(0) = 0\big). \tag{4.2}$$

We first need to state several lemmas before proving Theorem 4.7. Proofs of Lemmas 4.1, 4.2 and 4.4 can be found in the supplement [12], and the remainder of the results required, Lemmas 4.3, 4.5, 4.6 and Theorem 4.7, are straightforward generalizations of results in [19]; hence we omit the proofs.

Lemma 4.1. *Let* $t : [0, \infty) \to [0, \infty)$ *be a strictly increasing function. If* $h \in L^1([0, a], \nu)$ *for some* $a > 0$, *then*

$$f(x) = A_x h \qquad \textit{satisfies } \frac{t(x)}{t'(x)} f'(x) + f(x) = h(x) \textit{ a.e. on } (0, a]. \tag{4.3}$$

Conversely, if in addition $t(\cdot)$ is locally absolutely continuous on $[0,\infty)$ with $t(0)=0$, and $f \in \bigcup_{\alpha\geq 0}\mathrm{Lip}_\alpha$, then the function $h(\cdot)$ as given by the right-hand side of (4.3) *is in $L^1([0,a],\nu)$ for all $a>0$ and*

$$f(x)=A_x h \qquad \text{for all } x\in(0,\infty). \tag{4.4}$$

Lemma 4.2. *Let $t:[0,\infty)\to[0,\infty)$ be given by $t^{1/\theta}(\cdot)=s(\cdot)$ for $s(\cdot)$ a strictly increasing locally absolutely continuous function on $[0,\infty)$ with $s(0)=0$. Then $t(\cdot)$ is also locally absolutely continuous on $[0,\infty)$. Moreover, for W a non-negative random variable and W^* with distribution as in* (1.15), *for $h\in\bigcap_{a\in S}L^1([0,a],\nu)$ where S is the support of $W+1$,*

$$E\big[h\big(W^*\big)\big]=E[A_{W+1}h] \tag{4.5}$$

whenever either expectation above exists, and letting $f(x)=A_x h$ for all $x\in S$,

$$E\left[\frac{t(W^*)}{t'(W^*)}f'\big(W^*\big)+f\big(W^*\big)\right]=E\big[f(W+1)\big], \tag{4.6}$$

when the expectation of either side exists.

For an a.e. differentiable function $f(\cdot)$, let

$$\mathbb{D}_t f(x)=\frac{t(x)}{t'(x)}f'(x)+f(x)-f(x+1). \tag{4.7}$$

Note that if $f(x)=A_x g$ for some $g(\cdot)$, then under the conditions of Lemma 4.1, by (4.3) we may write (4.7) as

$$\mathbb{D}_t f(x)=g(x)-A_{x+1}g \qquad \text{almost everywhere.} \tag{4.8}$$

Condition 3.1 is assumed in some of the following statements to assure that the distribution of $D_{\theta,s}$ exists uniquely. The proof of the next lemma is omitted, as it follows using Lemmas 4.1 and 4.2, similar to the proof of Lemma 3.2 in [19].

Lemma 4.3. *Let $\theta>0$ and $s(\cdot)$ satisfy Condition* 3.1. *If $s(\cdot)$ is locally absolutely continuous on $[0,\infty)$, then,*

$$E\big[h(D_{\theta,s})\big]=E[A_{D_{\theta,s}+1}h] \quad \text{and} \quad E\big[\mathbb{D}_t f(D_{\theta,s})\big]=0,$$

for all $h(\cdot)\in\bigcap_{a\in(0,\infty)}L^1([0,a],\nu)$ and $f(\cdot)\in\bigcup_{\alpha\geq 0}\mathrm{Lip}_\alpha$ for which $E[\mathbb{D}_t f(D_{\theta,s})]$ exists, respectively.

The second claim of the lemma and (4.7) suggest the Stein equation

$$\frac{t(x)}{t'(x)}f'(x)+f(x)-f(x+1)=h(x)-E\big[h(D_{\theta,s})\big], \tag{4.9}$$

which via (4.8) may be rewritten as

$$g(x) - A_{x+1}g = h(x) - E\big[h(D_{\theta,s})\big] \tag{4.10}$$

whenever $g(\cdot)$ is such that $A_x g$ exists for all x and $f(x) = A_x g$.

To prove Theorem 3.4, we first need to identify a set of broad sufficient conditions on $t(\cdot)$ under which we can find a nice solution $g(\cdot)$ to (4.10) when $h \in \mathrm{Lip}_{1,0}$, where, suppressing dependence on θ and $s(\cdot)$ for notational simplicity, for $\alpha > 0$, we let

$$\mathrm{Lip}_{\alpha,0} = \big\{h : [0,\infty) \to \mathbb{R} : h \in \mathrm{Lip}_\alpha,\, E\big[h(D_{\theta,s})\big] = 0\big\}. \tag{4.11}$$

We note that the integral $I(x)$ in (3.10) can be written as the one appearing in (4.13) below when $t(x) = s^\theta(x)$ as in (4.1). Also note that by Lemma 4.2, if $s(\cdot)$ is strictly increasing with $s(0) = 0$, locally absolutely continuity of one of $s(\cdot)$ and $t(\cdot)$ implies that of the other. Hence, given that either one is locally absolutely continuous on $[0,\infty)$, as any continuous function $h : [0,\infty) \to \mathbb{R}$ is bounded on $[0,a]$ for all $a \geq 0$, we have $h \in \bigcap_{a>0} L^1([0,a],\nu)$. As the integrability of $h(\cdot)$ can thus be easily verified, it will not be given further mention.

Lemma 4.4. *Let $t : [0,\infty) \to [0,\infty)$ be a strictly increasing and locally absolutely continuous function on $[0,\infty)$. If $h(\cdot)$ is absolutely continuous on $[0,a]$ for some $a > 0$ with a.e. derivative $h'(\cdot)$, then with $A_x h$ as in (4.2),*

$$(A_x h)' = \frac{t'(x)}{t^2(x)} \int_0^x h'(u)t(u)\,du \qquad \textit{a.e. on } x \in (0,a]. \tag{4.12}$$

If there exists some $\rho \in [0,\infty)$ such that

$$\operatorname*{ess\,sup}_{x>0} I(x) \leq \rho \qquad \textit{where } I(x) = \frac{t'(x)}{t^2(x)} \int_0^x t(u)\,du, \tag{4.13}$$

then $A_x h \in \mathrm{Lip}_{\alpha\rho}$ on $[0,\infty)$ whenever $h \in \mathrm{Lip}_\alpha$ for some $\alpha \geq 0$.

Remark 4.1. If $\theta > 0$ and $t(\cdot)$ is given by $t(\cdot) = s^\theta(\cdot)$ for $s(\cdot)$ concave and continuous at zero, then $\|I\|_\infty \leq \theta/(\theta+1)$ by Theorem 3.7. Hence, $\rho \in [0,1)$ always exists for such choices of $t(\cdot)$.

Lemmas 4.5, 4.6 and Theorem 4.7 generalize Lemmas 3.5, 3.6 and Theorem 3.1 in [19] for the generalized Dickman; their proofs follow closely those in [19] and hence are omitted.

Lemma 4.5. *Let $\theta > 0$ and $s(\cdot)$ satisfy Condition 3.1. Moreover assume that $\mu = E[D_{\theta,s}]$ exists. Then with $\mathrm{Lip}_{\alpha,0}$ as in (4.11), for any $\alpha > 0$,*

$$\sup\nolimits_{h\in \mathrm{Lip}_{\alpha,0}} \big|h(0)\big| = \alpha\mu. \tag{4.14}$$

To define iterates of the averaging operator on a function $h(\cdot)$, let $A^0_{x+1}h = h(x)$ and

$$A^n_{x+1} = A_{x+1}\big(A^{n-1}_{\bullet+1}\big) \qquad \text{for } n \geq 1,$$

and for a class of functions $\mathcal{H}$ let

$$A^n_{x+1}(\mathcal{H}) = \{A^n_{x+1}h : h \in \mathcal{H}\} \qquad \text{for } n \geq 0.$$

Lemma 4.6. *Let $s(\cdot)$ satisfy Condition 3.1 and be locally absolutely continuous on $[0,\infty)$. If there exists $\rho \in [0,\infty)$ such that (4.13) holds, then for all $\theta > 0$, $\alpha \geq 0$ and $n \geq 0$,*

$$A^n_{x+1}(\mathrm{Lip}_{\alpha,0}) \subset \mathrm{Lip}_{\alpha\rho^n,0}.$$

In the following, by replacing $h(x)$ by $h(x) - E[h(D_{\theta,s})]$, when handling the Stein equations (4.9) and (4.10), without loss of generality we may assume that $E[h(D_{\theta,s})] = 0$.

For a given function $h \in \mathrm{Lip}_{\alpha,0}$ for some $\alpha \geq 0$, let

$$\begin{aligned} h^{(\star k)}(x) &= A^k_{x+1}h \qquad \text{for } k \geq 0, \qquad g(x) = \sum_{k\geq 0} h^{(\star k)}(x) \quad \text{and} \\ g_n(x) &= \sum_{k=0}^{n} h^{(\star k)}(x). \end{aligned} \tag{4.15}$$

Also recall definition (1.22) that for any $a \geq 0$ and function $f(\cdot)$, $\|f\|_{[0,a]} = \sup_{x\in[0,a]} |f(x)|$.

Theorem 4.7. *Let $s(\cdot)$ satisfy Condition 3.1 and be locally absolutely continuous on $[0,\infty)$. Further assume that $\mu = E[D_{\theta,s}]$ exists. If there exists $\rho \in [0,1)$ such that (4.13) holds, then for all $a \geq 0$ and $h \in \mathrm{Lip}_{1,0}$ we have*

$$\left\|h^{(\star k)}\right\|_{[0,a]} \leq (\mu + a)\rho^k, \tag{4.16}$$

$g_n \in \mathrm{Lip}_{(1-\rho^{n+1})/(1-\rho)}$ and $g(\cdot)$ given by (4.15) is a $\mathrm{Lip}_{1/(1-\rho)}$ solution to (4.10).

Proof of Theorem 3.4. The proof follows by arguing as in the proof of Theorem 1.3 of [19], with the final claim obtained by applying Theorem 3.7 to $s(x) = x$; we omit the details. $\square$

In the remainder of this section, we specialize to the case of the generalized Dickman distribution where for some $\theta > 0$ we have $t(x) = x^\theta$, $dv/dv = \theta v^{\theta-1}$ and the Stein equation (4.9) becomes

$$(x/\theta)f'(x) + f(x) - f(x+1) = h(x) - E\big[h(D_\theta)\big]. \tag{4.17}$$

Note that the function $s(x) = x$ trivially satisfies Condition 3.1. For notational simplicity, in what follows, let $\rho_i = \theta/(\theta + i)$ for $i \in \{1,2\}$.

Lemma 4.8. *For non-negative α and β, let $\mathcal{H}_{\alpha,\beta}$ be as in (1.10). For every $\theta > 0$, if $h \in \mathcal{H}_{\alpha,\beta}$ then $A_x h \in C^2[(0,\infty)]$ and both $A_x h$ and $A_{x+1}h$ are elements of $\mathcal{H}_{\alpha\rho_1,\beta\rho_2}$.*

Proof. Take $h \in \mathcal{H}_{\alpha,\beta}$. Since $h \in \mathrm{Lip}_\alpha$, by Lemmas 4.6 and 4.4, $h(\cdot)$ is ν-integrable on any interval of the form $[0, a]$ for all $a > 0$, $A_x h \in \mathrm{Lip}_{\alpha\rho_1}$ and

$$(A_x h)' = \frac{\theta}{x^{\theta+1}} \int_0^x h'(v) v^\theta \, dv \qquad \text{for } x > 0.$$

Taking another derivative, we obtain

$$(A_x h)'' = \frac{\theta}{x^{\theta+1}} \left[h'(x) x^\theta - \frac{\theta+1}{x} \int_0^x h'(v) v^\theta \, dv \right] \qquad \text{for } x > 0.$$

As $h' \in \mathrm{Lip}_\beta$, the function $A_x h$ is twice continuously differentiable on $(0, \infty)$ proving the first claim. Since

$$x^\theta = \frac{\theta+1}{x} \int_0^x v^\theta \, dv$$

we have

$$(A_x h)'' = \frac{\theta(\theta+1)}{x^{\theta+2}} \left[\int_0^x \big(h'(x) - h'(v)\big) v^\theta \, dv \right].$$

Taking absolute value and using that $h' \in \mathrm{Lip}_\beta$ now yields

$$\begin{aligned} \big|(A_x h)''\big| &\le \frac{\theta(\theta+1)}{x^{\theta+2}} \left[\int_0^x \big|h'(x) - h'(v)\big| v^\theta \, dv \right] \\ &\le \frac{\beta\theta(\theta+1)}{x^{\theta+2}} \left[\int_0^x (x - v) v^\theta \, dv \right] = \frac{\beta\theta(\theta+1)}{x^{\theta+2}} \frac{x^{\theta+2}}{(\theta+1)(\theta+2)} = \frac{\beta\theta}{\theta+2} = \beta\rho_2. \end{aligned}$$

Since both $A_x h$ and $(A_x h)'$ are continuous at 0 and belong in $C^1[(0, \infty)]$, we obtain $A_x h \in \mathcal{H}_{\alpha\rho_1,\beta\rho_2}$. The final claim is a consequence of the fact that $A_{x+1} h$ is a left shift of $A_x h$. □

Theorem 4.9. *For every $\theta > 0$ and $h \in \mathcal{H}_{1,1}$, there exists a solution $f \in \mathcal{H}_{\theta,\theta/2}$ to (4.17) with $\|f'\|_{(0,\infty)} \le \theta$ and $\|f''\|_{(0,\infty)} \le \theta/2$.*

Proof. Take $h \in \mathcal{H}_{1,1}$. By replacing $h(\cdot)$ by $h - E[h(D_\theta)]$ we may assume $E[h(D_\theta)] = 0$. Clearly $s(x) = x$ satisfies Condition 3.1 and $E[D_\theta] = \theta$ (see, e.g., [15]). Also, by Theorem 3.4, $\rho = \rho_1$ satisfies (4.13). For $h \in \mathrm{Lip}_{1,0}$, Theorem 4.7 shows that $g(\cdot)$ given by (4.15) is a $\mathrm{Lip}_{1/(1-\rho_1)}$ solution to (4.10). Since $g(\cdot)$ is Lipschitz, we have $g \in \bigcap_{a>0} L^1([0, a], \nu)$ and hence $f(x) = A_x g$ is a solution to (4.17) by the equivalence of (4.9) and (4.10). Now for $a > 0$, for any function $h \in L^1([0, a], \nu)$,

$$\|A_\bullet h\|_{[0,a]} = \sup_{x\in[0,a]} |A_x h| \le \sup_{x\in[0,a]} \frac{1}{x^\theta} \int_0^x \big|h(v)\big| \theta v^{\theta-1} \, dv \le \|h\|_{[0,a]}. \tag{4.18}$$

Let

$$g_n(x) = \sum_{k=0}^{n} h^{(\star k)}(x) \quad \text{and} \quad f_n(x) = A_x g_n.$$

Since $g_n \in \mathrm{Lip}_{(1-\rho^{n+1})/(1-\rho)}$ by Theorem 4.7, it is ν-integrable over $[0, a]$. Now using (4.18), the triangle inequality and (4.16) of Theorem 4.7, noting $E[D_\theta] = \theta$, we have

$$\begin{aligned}\|f - f_n\|_{[0,a]} &= \|A_\bullet g - A_\bullet g_n\|_{[0,a]} \le \|g - g_n\|_{[0,a]} \\ &\le \sup_{x\in[0,a]} \sum_{k\ge n+1} \|h^{(\star k)}\|_{[0,a]} \le (\theta + a) \sum_{k\ge n+1} \rho_1^k = (\theta + a)\frac{\rho_1^{n+1}}{1-\rho_1}.\end{aligned}$$

Letting $n \to \infty$, we obtain $f(x) = \sum_{n\ge 0} A_x h^{(\star n)}$. Lemma 4.8 and induction imply that $A_x h^{(\star n)} \in C^2[(0, \infty)]$ and $A_x h^{(\star n)} \in \mathcal{H}_{\rho_1^{n+1},\rho_2^{n+1}}$ for all $n \ge 0$, and hence

$$\|(A_x h^{(\star n)})'\|_{(0,\infty)} \le \rho_1^{n+1} \quad \text{and} \quad \|(A_x h^{(\star n)})''\|_{(0,\infty)} \le \rho_2^{n+1}. \tag{4.19}$$

Thus, for any $a > 0$, on the interval $(0, a]$, $f_n'(x) = \sum_{k=0}^{n} (A_x h^{(\star k)})'$ and $f_n''(x) = \sum_{k=0}^{n} (A_x h^{(\star k)})''$ converge uniformly to the corresponding infinite sums respectively, noting that by (4.19), the infinite sums are absolutely summable. Thus we obtain (see, e.g., Theorem 7.17 in [33])

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad \text{and} \quad f''(x) = \lim_{n\to\infty} f_n''(x) \qquad \text{for all } x \in [0, a].$$

Hence, again using (4.19), with $\|\cdot\|_{(0,\infty)}$ the supremum norm defined as in (1.22),

$$\|f'\|_{(0,\infty)} \le \sum_{n\ge 0} \rho_1^{n+1} = \frac{\rho_1}{1-\rho_1} = \theta \quad \text{and} \quad \|f''\|_{(0,\infty)} \le \sum_{n\ge 0} \rho_2^{n+1} = \frac{\rho_2}{1-\rho_2} = \frac{\theta}{2}.$$

Finally, since $f(\cdot)$ and $f'(\cdot)$ are differentiable everywhere on $(0, \infty)$ with bounded derivative, they are absolutely continuous on $(0, \infty)$. Also both $f(\cdot)$ and $f'(\cdot)$ are continuous at 0 since by definition, $f(0) = A_0 g = g(0) = \lim_{x\downarrow 0} f(x)$ and $f'(0) = \lim_{x\downarrow 0} f'(x)$. Now noting that if a function is absolutely continuous on $(0, \infty)$ with bounded derivative and continuous at 0, then it is Lipschitz, we obtain that $f \in \mathcal{H}_{\theta,\theta/2}$. □

Remark 4.2. The reasoning in the proof of Theorem 4.9 holds in greater generality in $t(\cdot)$, and only specifically depends on the form $t(x) = x^\theta$ when invoking Lemma 4.8.

Remark 4.3. In contrast to the bound $\|f''\|_\infty \le 2\|h'\|_\infty$ (see, e.g., (2.12) of [14]) for the solution of Stein equation in the normal case, one cannot uniformly bound the second derivatives of the solutions $f(\cdot)$ of (4.17) in Theorem 4.9 assuming only a Lipschitz condition on the test functions $h(\cdot)$ in a class $\mathcal{H}$; see [12] for details.

Remark 4.4. Shortly after a draft of this manuscript was posted, as a special case of their work on infinitely divisible laws, Arras and Houdré proved smoothness bounds in [1] for a solution to the standard Dickman Stein equation of the form

$$xt(x) - \int_0^1 t(x+u)\,du = h(x) - Eh(D); \tag{4.20}$$

this equation corresponds to (4.17) upon identifying $t(\cdot)$ and $f'(\cdot)$. Lemma 5.2 in [1] shows that when $h(\cdot)$ is in the class $\mathcal{H} = \{h : \|h\|_\infty \le 1, \|h'\|_\infty \le 1, h'(\cdot) \text{ is continuous}\}$ then there exists a solution $t(\cdot)$ to (4.20) with $\|t'\|_\infty \le 1$. The proof of Theorem 2.1 requires a uniform bound on $f'(\cdot)$ over $(0, \infty)$ to control the coefficient of $|\mu - 1|$ in (2.6). As no such bound is provided in [1], in the case $\mu = 1$ one can argue as for Theorem 2.1 to produce a version of it for the metric induced by $\mathcal{H}$. As neither class $\mathcal{H}$ nor $\mathcal{H}_{1,1}$ in (1.10) contains the other, the first class requiring the test functions to be uniformly bounded, and the second requiring their derivatives to be Lipschitz, the resulting metrics they induce are incomparable.

Acknowledgements

The authors would like to thank two anonymous reviewers for their valuable comments. The second author was partially supported by NSA grant H98230-15-1-0250.

Supplementary Material

Supplement to: Dickman approximation in simulation, summations and perpetuities (DOI: 10.3150/18-BEJ1070SUPP; .pdf). This self contained article (also available at https://arxiv.org/abs/1706.08192) gives detailed proofs of results that were required, but not provided, in the current manuscript.

References

[1] Arras, B. and Houdré, C. (2017). On Stein's method for infinitely divisible laws with finite first moment. Available at https://arxiv.org/abs/1712.10051.

[2] Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2016). Distances between probability distributions via characteristic functions and biasing. Available at https://arxiv.org/abs/1605.06819v1.

[3] Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2017). A new approach to the Stein–Tikhomirov method: With applications to the second Wiener chaos and Dickman convergence. Available at https://arxiv.org/abs/1605.06819.

[4] Arratia, R. (2002). On the amount of dependence in the prime factorization of a uniform random integer. In *Contemporary Combinatorics*. *Bolyai Soc*. *Math*. *Stud*. **10** 29–91. Budapest: János Bolyai Math. Soc. MR1919568

[5] Arratia, R. (2017). Personal communication.

[6] Arratia, R., Barbour, A.D. and Tavaré, S. (2003). *Logarithmic Combinatorial Structures*: *A Probabilistic Approach*. Zürich: European Mathematical Society (EMS). MR2032426

[7] Barbour, A.D. and Nietlispach, B. (2011). Approximation by the Dickman distribution and quasi-logarithmic combinatorial structures. *Electron. J. Probab.* **16** 880–902. MR2793242

[8] Baxendale, P. (2017). Personal communication.

[9] Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scientiarum Imperiales Petropolitanae* **5** 175–192.

[10] Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica* **22** 23–36. MR0059834

[11] Bhatt, A.G. and Roy, R. (2004). On a random directed spanning tree. *Adv. in Appl. Probab.* **36** 19–42. MR2035772

[12] Bhattacharjee, C. and Goldstein, L. (2018). Supplement to "Dickman approximation in simulation, summations and perpetuities." DOI:10.3150/18-BEJ1070SUPP.

[13] Cellarosi, F. and Sinai, Y.G. (2013). Non-standard limit theorems in number theory. In *Prokhorov and Contemporary Probability Theory*. *Springer Proc. Math. Stat.* **33** 197–213. Heidelberg: Springer. MR3070473

[14] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. Heidelberg: Springer. MR2732624

[15] Devroye, L. and Fawzi, O. (2010). Simulating the Dickman distribution. *Statist. Probab. Lett.* **80** 242–247. MR2575452

[16] Dickman, K. (1930). On the frequency of numbers containing prime factors of a certain relative magnitude. *Ark. Mat. Astron. Fys.* **22** 1–14.

[17] Eeckhoudt, L., Gollier, C. and Schlesinger, H. (2005). *Economic and Financial Decisions Under Risk*. Princeton: Princeton Univ. Press.

[18] Fill, J.A. and Huber, M.L. (2010). Perfect simulation of Vervaat perpetuities. *Electron. J. Probab.* **15** 96–109. MR2587562

[19] Goldstein, L. (2017). Non asymptotic distributional bounds for the Dickman approximation of the running time of the Quickselect algorithm. Available at https://arxiv.org/abs/1703.00505v1.

[20] Goldstein, L. (2018). Non asymptotic distributional bounds for the Dickman approximation of the running time of the quickselect algorithm. *Electron. J. Probab.* **23** 1–13.

[21] 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

[22] Hwang, H.-K. and Tsai, T.-H. (2002). Quickselect and the Dickman function. *Combin. Probab. Comput.* **11** 353–371. MR1918722

[23] 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

[24] Peköz, E.A., Röllin, A. and Ross, N. (2013). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23** 1188–1218. MR3076682

[25] Penrose, M.D. and Wade, A.R. (2004). Random minimal directed spanning trees and Dickman-type distributions. *Adv. in Appl. Probab.* **36** 691–714. MR2079909

[26] Pinsky, R. (2016). A natural probabilistic model on the integers and its relation to Dickman-type distributions and Buchstab's function. Available at https://arxiv.org/abs/1606.02965.

[27] Pinsky, R. (2016). On the strange domain of attraction to generalized Dickman distributions for sums of independent random variables. Available at https://arxiv.org/abs/1611.07207.

[28] Rachev, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models*. Chichester: Wiley. MR1105086

[29] Ramanujan, S. (1919). A proof of Bertrand's postulate. *J. Indian Math. Soc.* (*N.S.*) **11** 27.

[30] Rényi, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand No.* **8** 7–13. MR0286162

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

[32] Royden, H.L. and Fitzpatrick, P. (1988). *Real Analysis*, 3rd ed. New York: Macmillan Publishing Company. MR1013117
[33] Rudin, W. (1964). *Principles of Mathematical Analysis*, *Vol*. 3. New York: McGraw-Hill. MR0166310
[34] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. 583–602. MR0402873
[35] Stein, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **7**. Hayward, CA: IMS. MR0882007
[36] Stolz, O. (1885). *Vorlesungen über Allgemeine Arithmetik*: *Nach Den Neueren Ansichten* (*Vol*. 1). Leipzig: BG Teubner.
[37] Vervaat, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv*. *in Appl*. *Probab*. **11** 750–783. MR0544194

Received June 2017 and revised August 2018