Lith. Math. J. **61**,
n° 3 (2021), 301-311.

On strong and almost sure local limit theorems for a probabilistic model of the Dickman distribution

Régis de la Bretèche & Gérald Tenenbaum

*To the memory of Jonas Kubilius,
who stood on the bridge and invited us all.*

Abstract.

Let $\{Z_k\}_{k\geqslant 1}$ denote a sequence of independent Bernoulli random variables defined by $\mathbb{P}(Z_k=1)=1/k=1-\mathbb{P}(Z_k=0)$ $(k\geqslant 1)$ and put $T_n:=\sum_{1\leqslant k\leqslant n}kZ_k$. It is then known that T_n/n converges weakly to a real random variable D with density proportional to the Dickman function, defined by the delay-differential equation $u\varrho'(u)+\varrho(u-1)=0$ $(u>1)$ with initial condition $\varrho(u)=1$ $(0\leqslant u\leqslant 1)$. Improving on earlier work, we propose asymptotic formulae with remainders for the corresponding local and almost sure limit theorems, namely

$$\sum_{m\geqslant 0}\left|\mathbb{P}(T_n=m)-\frac{\mathrm{e}^{-\gamma}}{n}\varrho\Big(\frac{m}{n}\Big)\right|=\frac{2\log n}{\pi^2 n}\Big\{1+O\Big(\frac{1}{\log_2 n}\Big)\Big\}\qquad(n\to\infty),$$

and

$$(\forall u>0)\quad\sum_{\substack{n\leqslant N\\ T_n=\lfloor un\rfloor}}1=\mathrm{e}^{-\gamma}\varrho(u)\log N+O\Big((\log N)^{2/3+o(1)}\Big)\quad\text{a.s.}\quad(N\to\infty),$$

where γ denotes Euler's constant.

Keywords: almost sure limit theorems, almost sure local limit theorems, Dickman function, Dickman distribution, quickselect algorithm, friable integers, random permutations.

2020 Mathematics Subject Classification: primary 60F15; secondary: 11N25, 60C05.

1. Introduction and statement of results

Dickman's function is defined on $[0,\infty[$ as the continuous solution to the delay-differential equation $u\varrho'(u)+\varrho(u-1)=0$ $(u>1)$ with initial condition $\varrho(u)=1$ $(0\leqslant u\leqslant 1)$. It is known (see, e.g., [24; th. III.5.10]) that $\int_0^\infty\varrho(u)\,\mathrm{d}u=\mathrm{e}^\gamma$, where γ denotes Euler's constant. The Dickman distribution is defined as the law of a random variable D on $[0,\infty[$ with density

$$\varrho_0(u):=\mathrm{e}^{-\gamma}\varrho(u)\quad(u\geqslant 0).$$

This law appears in a large variety of mathematical topics, such as (the following list being non limiting):

Number theory, in the context of friable integers,$^{(1)}$ after the seminal paper of Dickman [10] : see [24] for a expositary account;

Random polynomials over finite fields : see, e.g., Car [6], Manstavičius [18], Arratia, Barbour & Tavaré [1], Knopfmacher & Manstavičius [17];

Random permutations: see in particular, Shepp & Lloyd [21], Kingman [16], Arratia, Barbour & Tavaré [3], Manstavičius & Petuchovas [19].

In number theory, the Dickman function also appears in Billingsley's model [5] for the vector distribution of large prime factors of integers (see [23] for an effective version) and in Kubilius' model$^{(2)}$: see Elliott [11], Arratia, Barbour & Tavaré [2], Tenenbaum [22], and [24; § III.6.5] for an expositary account.

---

1. That is integers free of large prime factors
2. A probabilistic model for the uniform probability defined on the set of the first N integers with σ-algebra comprising those events that can be defined by divisibility conditions involving solely small primes

A simple probabilistic description of D is provided by the almost surely convergent series

$$\sum_{n\geqslant 1}\prod_{1\leqslant j\leqslant n} X_j,$$

where the X_j are independent and uniform on $[0,1]$: see Goldie & Grübel [14], Fill & Huber [12], Devroye [8].

There is a vast bibliography on the various probabilistic models of the Dickman distribution: see, e.g., Chen & Hwang [7], Devroye & Fawzi [9], Pinsky [20].

In 2002, Hwang & Tsai [15] used a simple model to show that, suitably normalized, the cost of Hoare's quickselect algorithm converges weakly to D. This model may be described as follows: if $\{Z_k\}_{k\geqslant 1}$ denotes a sequence of independent Bernoulli random variables such that $\mathbb{P}(Z_k=1)=1/k=1-\mathbb{P}(Z_k=0)$ $(k\geqslant 1)$ and if $T_n:=\sum_{1\leqslant k\leqslant n} kZ_k$, then T_n/n converges weakly to D, viz.

$$\lim_{n\to\infty}\mathbb{P}(T_n\leqslant nu)=\mathrm{e}^{-\gamma}\int_0^u \varrho(v)\,\mathrm{d}v \qquad (u\geqslant 0).$$

A strong local limit theorem was then obtained by Giuliano, Szewczak & Weber [13], in the form

$$v_n:=\sum_{m\geqslant 0}\Big|\mathbb{P}(T_n=m)-\frac{\mathrm{e}^{-\gamma}}{n}\varrho\Big(\frac{m}{n}\Big)\Big|=o(1) \qquad (n\to\infty). \tag{1·1}$$

We propose a sharp estimate of the speed of convergence. Here and in the sequel, we let $\log_k$ denote the k-fold iterated logarithm.

Theorem 1.1. *We have*

$$v_n=\frac{2\log n}{\pi^2 n}\Big\{1+O\Big(\frac{1}{\log_2 n}\Big)\Big\} \qquad (n\to\infty). \tag{1·2}$$

This estimate may be put in perspective with the following result of Manstavičius [18]. Let $\{X_k\}_{k\geqslant 1}$ denote a sequence of independent Poisson variables such that $\mathbb{E}(X_k)=1/k$, and put $Y_n:=\sum_{1\leqslant k\leqslant n} kX_k$. Then [18; cor. 2] readily yields the strong local limit theorem

$$\sum_{m\geqslant 0}\Big|\mathbb{P}(Y_n=m)-\frac{\mathrm{e}^{-\gamma}}{n}\varrho\Big(\frac{m}{n}\Big)\Big|\ll\frac{1}{n} \qquad (n\geqslant 1). \tag{1·3}$$

Thus, as may be expected, Poissonian approximations to the Bernoulli random variables Z_k provide a closer model of the Dickman distribution. As a byproduct of (1·2) and (1·3), we get an estimate of the total variation distance between T_n and Y_n, viz.

$$\begin{aligned} d_{TV}(T_n,Y_n)&:=\sum_{m\geqslant 0}|\mathbb{P}(T_n=m)-\mathbb{P}(Y_n=m)|=v_n+O\Big(\frac{1}{n}\Big)\\ &=\frac{2\log n}{\pi^2 n}\Big\{1+O\Big(\frac{1}{\log_2 n}\Big)\Big\}.\end{aligned}$$

We also point, without details, to a recent estimate of Bhattacharjee & Goldstein [4; th.1.1], which provides a bound $\leqslant 3/(4n)$ for a smooth Wasserstein-type distance between T_n/n and D.

For $u>0$, let ε_n denote a non-negative sequence tending to 0 at infinity, and let $\{m_n\}_{n\geqslant 1}$ denote a non-decreasing integer sequence such that $m_n=un+O(\varepsilon_n n)$ as $n\to\infty$. We may then define a sequence of random variables $\{L_N(u)\}_{N=1}^{\infty}$ by the formula

$$L_N(u):=\sum_{n\leqslant N,\,T_n=m_n} 1.$$

By a complicated proof resting on a general correlation inequality, an almost sure local limit theorem is established in [13] assuming furthermore that $\{m_n\}_{n\geqslant 1}$ is strictly increasing: for any $u\geqslant 1$, the asymptotic formula $L_N(u)\sim \mathrm{e}^{-\gamma}\varrho(u)\log N$ holds almost surely as $N\to\infty$.[3]

The following result, proved by a simple, direct method, provides an effective version.

Theorem 1.2. *Let* $u\geqslant 1$, $\varepsilon_n=o(1)$ *as* $n\to\infty$, *and let* $\{m_n\}_{n\geqslant 1}$ *denote a strictly increasing sequence of integers such that* $m_n=un+O(\varepsilon_n n)$ $(n\geqslant 1)$. *We have, almost surely,*

$$L_N(u)=\Big\{1+O\Big(\eta_N+\frac{(\log_2 N)^{1/2+o(1)}}{(\log N)^{1/3}}\Big)\Big\}\mathrm{e}^{-\gamma}\varrho(u)\log N, \tag{1·4}$$

where $\eta_N:=(1/\log N)\sum_{1\leqslant N}\varepsilon_n/n=o(1)$.

Furthermore, for any $u>0$, *the formula* $L_N(u)\sim\mathrm{e}^{-\gamma}\varrho(u)\log N$ *holds almost surely provided*

$$\vartheta_m:=|\{n\geqslant 1: m_n=m\}|=o(\log m)\qquad(m\to\infty),$$

and assuming only that the sequence $\{m_n\}_{n\geqslant 1}$ *is non-decreasing. If* $\vartheta_m\ll(\log m)^\alpha$ *with* $0\leqslant\alpha<1$, *the estimate* (1·4) *holds with remainder* $\ll\eta_N+1/(\log N)^{(1-\alpha)/3+o(1)}$.

We note that, for all $u>0$, the case $m_n:=\lfloor un\rfloor$ is covered by the second part of the statement with $\alpha=0$.

2. Proof of Theorem 1.1

Let c be a large constant and put $M(n):=cn(\log n)/(\log_2 3n)$. We first show that the contribution to v_n of those $m>M(n)$ is negligible. Indeed, since $\varrho(v)\ll v^{-v}$, we first have

$$\frac{1}{n}\sum_{m>M(n)}\varrho\Big(\frac{m}{n}\Big)\ll\frac{1}{n}\sum_{m>M(n)}\mathrm{e}^{-m(\log_2 n)/2n}\ll\frac{1}{n}\cdot$$

Then, we have, for all $y\geqslant 0$

$$\sum_{m>M(n)}\mathbb{P}(T_n=m)\leqslant\mathrm{e}^{-yM(n)}\mathbb{E}(\mathrm{e}^{yT_n})=\mathrm{e}^{-yM(n)}\prod_{1\leqslant k\leqslant n}\Big(1+\frac{\mathrm{e}^{ky}-1}{k}\Big). \tag{2·1}$$

Selecting $y=(\log_2 n)/n$, we see that the last product is

$$\ll\exp\Big\{\int_0^n\frac{\mathrm{e}^{yv}-1}{v}\,\mathrm{d}v\Big\}\ll n^{2/\log_2 n},$$

hence the left-hand side of (2·1) is also $\ll 1/n$, and we infer that

$$v_n=\sum_{1\leqslant m\leqslant M(n)}\Big|\mathbb{P}(T_n=m)-\frac{\mathrm{e}^{-\gamma}}{n}\varrho\Big(\frac{m}{n}\Big)\Big|+O\Big(\frac{1}{n}\Big). \tag{2·2}$$

Recall the definition $\varrho_0(u):=\mathrm{e}^{-\gamma}\varrho(u)$ $(u\in\mathbb{R})$ and let $I(s):=\int_0^1(\mathrm{e}^{vs}-1)\,\mathrm{d}v/v$ $(s\in\mathbb{C})$. From [24; th. III.5.10], we know that

$$\widehat{\varrho_0}(\tau):=\int_{\mathbb{R}}\mathrm{e}^{i\tau u}\varrho_0(u)\,\mathrm{d}u=\mathrm{e}^{I(i\tau)}\qquad(\tau\in\mathbb{R}). \tag{2·3}$$

Next, for $|\tau|<\pi$, we have

$$\begin{aligned}\mathbb{E}\big(\mathrm{e}^{i\tau T_n}\big)&=\prod_{1\leqslant k\leqslant n}\Big(1+\frac{\mathrm{e}^{i\tau k}-1}{k}\Big)\\&=\exp\Big\{S_n(\tau)+U(\tau)+W_n(\tau)+O\Big(\frac{\tau}{n(1+n|\tau|)}\Big)\Big\},\end{aligned} \tag{2·4}$$

---

[3] The authors of [13] state that this almost sure asymptotic formula holds for all $u>0$. However, the requirement that $\{m_n\}_{n=1}^\infty$ should be strictly increasing is incompatible with the assumption $m_n\sim un$ if $u<1$.

with

$$S_n(\tau) := \sum_{1\leqslant k\leqslant n} \frac{\mathrm{e}^{i\tau k}-1}{k}, \quad W_n(\tau) := \sum_{k>n} \frac{(\mathrm{e}^{i\tau k}-1)^2}{2k^2}$$
$$U(\tau) := \sum_{k\geqslant 1} \Big\{ \log\Big(1+\frac{\mathrm{e}^{i\tau k}-1}{k}\Big) - \frac{\mathrm{e}^{i\tau k}-1}{k}\Big\}.$$

For $|\tau| < 2\pi$, we may write

$$S_n(\tau) = \sum_{1\leqslant k\leqslant n} \int_0^{i\tau} \mathrm{e}^{kv}\,\mathrm{d}v = \int_0^{i\tau} \frac{\mathrm{e}^{nv}-1}{1-\mathrm{e}^{-v}}\,\mathrm{d}v = \int_0^n \frac{\mathrm{e}^{i\tau v}-1}{v}\,\mathrm{d}v + V_n(\tau)$$

with

$$V_n(\tau) := \int_0^{i\tau} (\mathrm{e}^{nv}-1)\Big(\frac{1}{1-\mathrm{e}^{-v}} - \frac{1}{v}\Big)\,\mathrm{d}v = \int_0^1 (\mathrm{e}^{in\tau v}-1) g_\tau(v)\,\mathrm{d}v,$$
$$g_\tau(v) := \frac{i\tau}{1-\mathrm{e}^{-i\tau v}} - \frac{1}{v} \quad (0 \leqslant v \leqslant 1).$$

Since $g_\tau(v)$ is twice continuously differentiable on $[0,1]$, partial integration yields

$$V_n(\tau) = V(\tau) + \frac{a(\tau)\mathrm{e}^{in\tau} - \frac{1}{2}}{n} + O\Big(\frac{1}{n^2}\Big) \qquad (|\tau| \leqslant \pi),$$

with

$$V(\tau) := -\int_0^1 g_\tau(v)\,\mathrm{d}v, \quad a(\tau) := \frac{g_\tau(1)}{i\tau} = \frac{1}{1-\mathrm{e}^{-i\tau}} - \frac{1}{i\tau},$$

We have $W_n(\tau) \ll \tau$ if $|\tau| \leqslant 1/n$. When $1/n \leqslant |\tau| \leqslant \pi$, we have

$$\text{(2·5)} \quad W_n(\tau) = \sum_{k>n} \frac{\mathrm{e}^{2ik\tau} - 2\mathrm{e}^{ik\tau}}{2k^2} + \frac{1}{2n} + O\Big(\frac{1}{n^2}\Big) = \frac{1}{2n}\Big\{1 + O\Big(\frac{1}{1+n\min(|\tau|, \pi-|\tau|)}\Big)\Big\}.$$

by Abel's summation. This estimate is hence also valid for $|\tau| \leqslant 1/n$, and so we deduce that

$$V_n(\tau) + W_n(\tau) = V(\tau) + \frac{a(\tau)\mathrm{e}^{in\tau}}{n} + O\Big(\frac{\tau}{n(1+n|\tau|)} + \frac{1}{n+n^2\min(|\tau|, \pi-|\tau|)}\Big).$$

Put $F(\tau) := \mathrm{e}^{U(\tau)+V(\tau)} - 1$, so that $F(0) = 0$ and F may be analytically continued to the disc $\{z \in \mathbb{C} : |z| < 2\pi\}$. We finally get

$$\text{(2·6)} \qquad \mathbb{E}(\mathrm{e}^{i\tau T_n}) = \widehat{\varrho}_0(n\tau)\Big\{1 + F(\tau) + W_n^*(\tau) + O\Big(\frac{\tau}{1+n^2\tau^2}\Big)\Big\} \qquad (n \geqslant 1,\ |\tau| \leqslant \pi),$$

with

$$\text{(2·7)} \qquad W_n^*(\tau) := \frac{a(\tau)\{1+F(\tau)\}\mathrm{e}^{in\tau}}{n} + O\Big(\frac{1}{n+n^2\min(|\tau|, \pi-|\tau|)}\Big).$$

It follows that, for $m \geqslant 1$,

$$\text{(2·8)} \qquad \begin{aligned} \mathbb{P}(T_n = m) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbb{E}(\mathrm{e}^{i\tau T_n})\mathrm{e}^{-im\tau}\,\mathrm{d}\tau \\ &= \frac{1}{2\pi n}\int_{-n\pi}^{n\pi} \widehat{\varrho}_0(\tau)\mathrm{e}^{-i\tau m/n}\Big\{1 + F\Big(\frac{\tau}{n}\Big) + W_n^*\Big(\frac{\tau}{n}\Big) + O\Big(\frac{\tau}{n(1+\tau^2)}\Big)\Big\}\,\mathrm{d}\tau. \end{aligned}$$

By (2·3) and, say, [24; lemma III.5.9], we have

$$\text{(2·9)} \qquad \widehat{\varrho}_0(\tau) = \frac{-1}{i\tau} + O\Big(\frac{1}{\tau(1+|\tau|)}\Big) \quad (\tau \neq 0), \qquad \widehat{\varrho}_0(\tau) \asymp \frac{1}{1+|\tau|} \quad (\tau \in \mathbb{R}).$$

Therefore, the error term of (2·8) contributes $\ll 1/n^2$ to the right-hand side. Summing over $m \leqslant M(n)$, we obtain that the corresponding contribution to the right-hand side of (2·2) is $\ll (\log n)/(n \log_2 n)$, in accordance with (1·2).

We first evaluate

$$\frac{1}{2\pi n}\int_{-n\pi}^{n\pi} \widehat{\varrho_0}(\tau)\mathrm{e}^{-i\tau m/n}\,\mathrm{d}\tau \qquad \big(n \geqslant 1,\ 1 \leqslant m \leqslant M(n)\big) \tag{2·10}$$

by extending the integration range to $\mathbb{R}$ and inserting the first estimate (2·9) to bound the integral over $\mathbb{R} \smallsetminus [-\pi n, \pi n]$. This yields

$$\begin{aligned}\frac{1}{2\pi n}\int_{-n\pi}^{n\pi} \widehat{\varrho_0}(\tau)\mathrm{e}^{-i\tau m/n}\,\mathrm{d}\tau - \frac{\varrho_0(m/n)}{n} &= \frac{-1}{\pi n}\int_{n\pi}^{\infty} \frac{\sin(\tau m/n)}{\tau}\,\mathrm{d}\tau + O\Big(\frac{1}{n^2}\Big)\\ &= \frac{(-1)^{m+1}}{\pi^2 mn} + O\Big(\frac{1}{m^2 n} + \frac{1}{n^2}\Big).\end{aligned} \tag{2·11}$$

In order to estimate the contributions from F and W_n^* to the main term of (2·8), we use the more precise formula

$$\widehat{\varrho_0}(\tau) = \frac{i}{\tau} - \frac{\mathrm{e}^{i\tau}}{\tau^2} + O\Big(\frac{1}{\tau^3}\Big) \qquad (|\tau| \geqslant 1). \tag{2·12}$$

Writing $F(\tau) = \tau G(\tau)$, we indeed deduce from by (2·12) that

$$\begin{aligned}\int_{-n\pi}^{n\pi} \widehat{\varrho_0}(\tau)\mathrm{e}^{-i\tau m/n} F\Big(\frac{\tau}{n}\Big)\,\mathrm{d}\tau &= \int_{-n\pi}^{n\pi} \tau\widehat{\varrho_0}(\tau)\frac{\mathrm{e}^{-i\tau m/n}}{n} G\Big(\frac{\tau}{n}\Big)\,\mathrm{d}\tau\\ &= i\int_{I_n} \mathrm{e}^{-i\tau m} G(\tau)\Big(1 + i\frac{\mathrm{e}^{i\tau n}}{n\tau}\Big)\,\mathrm{d}\tau + O\Big(\frac{1}{n}\Big),\end{aligned}$$

with $I_n := [-\pi,\pi] \smallsetminus [-1/n, 1/n]$. A standard computation furnishes $G(\pi) - G(-\pi) = -2/\pi$.[4] Integrating by parts, we get

$$\begin{aligned} i\int_{I_n} \mathrm{e}^{-i\tau m} G(\tau)\,\mathrm{d}\tau &= 2\frac{(-1)^{m+1}}{\pi m} + \frac{1}{m}\int_{-\pi}^{\pi} \mathrm{e}^{-i\tau m} G'(\tau)\,\mathrm{d}\tau + O\Big(\frac{1}{n}\Big)\\ &= 2\frac{(-1)^{m+1}}{\pi m} + O\Big(\frac{1}{m^2} + \frac{1}{n}\Big),\end{aligned}$$

and, similarly,

$$\frac{-1}{n}\int_{I_n} \mathrm{e}^{-i\tau m} G(\tau)\frac{\mathrm{e}^{i\tau n}}{\tau}\,\mathrm{d}\tau = \frac{-1}{n}\int_{I_n} \frac{G(\tau) - G(0)}{\tau}\mathrm{e}^{i\tau(n-m)}\,\mathrm{d}\tau + O\Big(\frac{1}{n}\Big) \ll \frac{1}{n}\cdot$$

We can thus state that, for $n \geqslant 1$, $1 \leqslant m \leqslant M(n)$, we have

$$\frac{1}{2\pi n}\int_{-n\pi}^{n\pi} \widehat{\varrho_0}(\tau)\mathrm{e}^{-i\tau m/n} F\Big(\frac{\tau}{n}\Big)\,\mathrm{d}\tau = \frac{(-1)^{m+1}}{\pi^2 mn} + O\Big(\frac{1}{m^2 n} + \frac{1}{n^2}\Big). \tag{2·13}$$

It remains to estimate the contribution to (2·8) involving $W_n^*(\tau/n)$. The error term of (2·7) clearly contributes $\ll 1/n^2$. Arguing as for the proof of (2·11), we finally show that the contribution to (2·8) arising from $a(\tau/n)(1 + F(\tau/n))\mathrm{e}^{i\tau}/n$ is also $\ll 1/n^2$.

The above estimates and (2·11) furnish together

$$\mathbb{P}(T_n = m) - \frac{1}{n}\varrho_0\Big(\frac{m}{n}\Big) = (-1)^{m+1}\frac{2}{\pi^2 mn} + O\Big(\frac{1}{m^2 n} + \frac{1}{n^2}\Big) \qquad \big(1 \leqslant m \leqslant M(n)\big). \tag{2·14}$$

Summing over $m \leqslant M(n)$ provides the required estimate (1·2).

---

[4] $V(\pm\pi) = \log(\pi/2) \mp \frac{1}{2}i\pi$, $\mathrm{e}^{U(\pm\pi)} = -2\mathrm{e}^{\gamma}$, $F(\pm\pi) = \mp i\pi\mathrm{e}^{\gamma} - 1$.

3. Proof of Theorem 1.2

We first note that (2·14) yields

$$\mathbb{P}(T_n=m)=\frac{1}{n}\varrho_0\Big(\frac{m}{n}\Big)+O\Big(\frac{1}{n^2}\Big)\qquad(m\asymp n\geqslant 1).\tag{3·1}$$

Hence, writing L_N for $L_N(u)$ here and throughout,

$$\mathbb{E}(L_N)=\sum_{n\leqslant N}\frac{1}{n}\varrho_0\Big(\frac{m_n}{n}\Big)+O(1)=\varrho_0(u)\log N+O(\eta_N\log N+1)\qquad(N\geqslant 1).\tag{3·2}$$

The stated result will follow from an estimate of the variance $\mathbb{V}(L_N)$. We have

$$\mathbb{E}(L_N^2)=\mathbb{E}(L_N)+2\sum_{1\leqslant\nu<n\leqslant N}\mathbb{P}(T_\nu=m_\nu)\alpha_{\nu n},\tag{3·3}$$

with

$$\alpha_{\nu n}:=\mathbb{P}\big(T_n-T_\nu=m_n-m_\nu\big)\qquad(1\leqslant\nu<n\leqslant N).$$

Let us initially assume that $\{m_n\}_{n\geqslant 1}$ is strictly increasing and hence that $u\geqslant 1$. We note right away that, for large n and $\nu>(u+1)n/(u+2)$, we have $\alpha_{\nu n}=0$, since the corresponding event is then impossible: either $T_n-T_\nu>\nu>(u+1)(n-\nu)>m_n-m_\nu$ or $T_n-T_\nu=0\neq m_n-m_\nu$. Therefore, we may assume $\nu\leqslant n(u+1)/(u+2)$ in the sequel.

By (2·4), we have, writing $\varphi_j(\tau):=\mathbb{E}\big(\mathrm{e}^{i\tau T_j}\big)$ $(j\geqslant 1)$,

$$\alpha_{\nu n}=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\varphi_n(\tau)}{\varphi_\nu(\tau)}\mathrm{e}^{-i(m_n-m_\nu)\tau}\,\mathrm{d}\tau=\beta_{\nu n}+\Delta_{\nu n},\tag{3·4}$$

where $\beta_{\nu n}$ denotes the contribution of the interval $|\tau|\leqslant 1/2\nu$, and $\Delta_{\nu n}$ that of the complementary range $1/2\nu<|\tau|\leqslant\pi$.

Now we may derive from (2·6) that

$$\beta_{\nu n}=\frac{1}{2\pi n}\int_{-n/2\nu}^{n/2\nu}\frac{\widehat{\varrho}_0(\tau)}{\varphi_\nu(\tau/n)}\Big\{1+F\Big(\frac{\tau}{n}\Big)+O\Big(\frac{1}{n}\Big)\Big\}\mathrm{e}^{-i(m_n-m_\nu)\tau/n}\,\mathrm{d}\tau.$$

Invoking (2·9) and (2·6)-(2·7) with ν in place of n to estimate the contribution of the error term, we get

$$\beta_{\nu n}=\beta^*_{\nu n}+O\bigg(\frac{\log n}{n^2}\bigg)\tag{3·5}$$

with

$$\beta^*_{\nu n}:=\frac{1}{2\pi n}\int_{-n/2\nu}^{n/2\nu}\frac{\widehat{\varrho}_0(\tau)}{\varphi_\nu(\tau/n)}\Big(1+F\Big(\frac{\tau}{n}\Big)\Big)\mathrm{e}^{-i(m_n-m_\nu)\tau/n}\,\mathrm{d}\tau.\tag{3·6}$$

Note that the remainder term of (3·5) in turn contributes $\ll 1$ to (3·3).

Still assuming that $\nu\leqslant(u+1)n/(u+2)$ and observing that

$$\psi_\nu(z):=\prod_{k\leqslant\nu}\{1+(\mathrm{e}^{kz}-1)/k\}^{-1}\ll 1\qquad(z\in\mathbb{C},\ |z|\leqslant 2/3\nu),$$

we may write

$$\frac{1+F(\tau/n)}{\varphi_\nu(\tau/n)}=1+\sum_{j\geqslant 1}\mu_{\nu j}\Big(\frac{\tau}{n}\Big)^j\qquad(|\tau|\leqslant n/2\nu)$$

with $\mu_{\nu j} \ll (3\nu/2)^j$ $(j \geqslant 1)$. Hence, in view of (2·9), the contribution of the series to (3·6) is

$$\begin{aligned}\frac{1}{2\pi n}\sum_{j\geqslant 1}\frac{\mu_{\nu j}}{n^j}&\int_{-n/2\nu}^{n/2\nu}\Big\{\tau^{j-1}\mathrm{e}^{i\tau(m_n-m_\nu)/n}+O\Big(\frac{\tau^{j-1}}{1+|\tau|}\Big)\Big\}\,\mathrm{d}\tau\\ &\ll\frac{1}{n}\sum_{j\geqslant 1}\mu_{\nu j}\int_{-1/2\nu}^{1/2\nu}\Big\{v^{j-1}\mathrm{e}^{iv(m_n-m_\nu)}+O\Big(\frac{v^{j-1}}{1+n|v|}\Big)\Big\}\,\mathrm{d}v\\ &\ll\frac{\nu}{n(n-\nu)}+\frac{\nu}{n^2}+\frac{\nu\log(2n/\nu)}{n^2}\ll\frac{\nu\log(2n/\nu)}{n^2},\end{aligned}$$

since the assumption that $\{m_n\}_{n=1}^{\infty}$ is strictly increasing implies $m_n-m_\nu\geqslant n-\nu$.

Arguing as in the proof of (2·10) to evaluate the main term, we eventually get,

$$\beta^*_{\nu n}=\frac{1}{n}\varrho_0\Big(\frac{m_n-m_\nu}{n}\Big)+O\Big(\frac{\nu\log(2n/\nu)}{n^2}\Big)\qquad(1\leqslant\nu<n). \tag{3·7}$$

We now consider $\Delta_{\nu n}$. Let

$$S_{\nu n}(\tau):=\sum_{\nu<k\leqslant n}\frac{\mathrm{e}^{ik\tau}}{k},$$

and note right away that $S_{\nu n}(\tau)$ is bounded for $1/2\nu<|\tau|\leqslant\pi$. For $\nu\geqslant 1$, we have

$$\frac{\varphi_n(\tau)}{\varphi_\nu(\tau)}=\frac{\nu}{n}\prod_{\nu<k\leqslant n}\Big(1+\frac{\mathrm{e}^{ik\tau}}{k-1}\Big)=\frac{\nu}{n}\exp\Big\{S_{\nu n}(\tau)+O\Big(\frac{1}{\nu+|\tau|\nu^2}+\frac{1}{\nu+(\pi-|\tau|)\nu^2}\Big)\Big\},$$

where we used the bounds

$$\begin{aligned}&\sum_{\nu<k\leqslant n}\frac{\mathrm{e}^{ik\tau}}{k(k-1)}\ll\frac{1}{\nu+|\tau|\nu^2},\qquad\sum_{\nu<k\leqslant n}\frac{\mathrm{e}^{2ik\tau}}{(k-1)^2}\ll\frac{1}{\nu+|\tau|\nu^2}+\frac{1}{\nu+(\pi-|\tau|)\nu^2},\\ &\sum_{\nu<k\leqslant n}\frac{\mathrm{e}^{ikj\tau}}{(k-1)^j}\ll\frac{1}{\nu^{j-1}}\quad(j>2).\end{aligned}$$

Carrying back into (3·4), we obtain

$$\Delta_{\nu n}=\frac{\nu}{\pi n}\Re\int_{1/2\nu}^{\pi}\mathrm{e}^{S_{\nu n}(\tau)-i(m_n-m_\nu)\tau}\,\mathrm{d}\tau+O\Big(\frac{\log 2\nu}{n\nu}\Big). \tag{3·8}$$

Now observe that $|S'_{\nu n}(\tau)|\leqslant\pi/|\tau|\leqslant\frac{1}{2}(n-\nu)\leqslant\frac{1}{2}(m_n-m_\nu)$ if, say $\nu\leqslant n/15$ or $|\tau|>10(u+1)/\nu$. Hence, on this assumption, a standard estimate on trigonometric integrals such as [25; Lemma 4.2] furnishes the bound $\ll 1/(m_n-m_\nu)\ll 1/(n-\nu)\ll 1/n$ for the last integral. Since, in the case $\nu>n/15$, the contribution of the range $1/2\nu<|\tau|\leqslant 10(u+1)/\nu$ to the same integral is trivially $\ll 1/\nu$, we finally get

$$\Delta_{\nu n}\ll\frac{\log 2\nu}{n\nu}+\frac{\nu}{n^2}\qquad\big(1\leqslant\nu\leqslant(u+1)n/(u+2)\big). \tag{3·9}$$

Gathering our estimates and using the fact that ϱ is Lipschitz on $[0,\infty[$, we obtain

$$\alpha_{\nu n}=\frac{1}{n}\varrho_0\Big(\frac{m_n}{n}\Big)+O\Big(\frac{\nu\log(2n/\nu)}{n^2}+\frac{\log 2\nu}{n\nu}\Big)\quad(1\leqslant\nu<n). \tag{3·10}$$

Carrying back into (3·3) and applying (3·1) for the pair (ν,m_ν) yields

$$\mathbb{E}(L_N^2)-\mathbb{E}(L_N)=2\sum_{1\leqslant\nu<n\leqslant N}\Big\{\varrho_0\Big(\frac{m_n}{n}\Big)\varrho_0\Big(\frac{m_\nu}{\nu}\Big)\frac{1}{\nu n}+O\Big(\frac{\log(2n/\nu)}{n^2}+\frac{\log 2\nu}{n\nu^2}\Big)\Big\},$$

and hence

$$\mathbb{V}(L_N) \ll \log N.$$

Selecting $N = N_k := 2^{k^3}$ for $k \geqslant 1$, we deduce from the Borel-Cantelli lemma that, given any $\varepsilon > 0$, the estimate

$$L_{N_k} - \mathbb{E}(L_{N_k}) \ll k^2(\log 2k)^{1/2+\varepsilon}$$

holds almost surely. In view of (3·2), this implies the stated result since L_N is a non-decreasing function of N.

We next consider the case of a non-decreasing sequence $\{m_n\}_{n\geqslant 1}$. Accordingly, we fix $u > 0$. By hypothesis, for some integer $q = q_N \geqslant 2$ such that $q_N = o\big(\log N\big)$ as $N \to \infty$, we have $m_n > m_\nu$ whenever $n - \nu \geqslant q_N$.

Put

$$L_N(u;a) := \sum_{\substack{n\leqslant N\\ n\equiv a \,(\mathrm{mod}\, q)\\ T_n=m_n}} 1 \quad (1 \leqslant a \leqslant q).$$

By (3·1), we have, for all $a \in [1, q]$,

$$\mathbb{E}\big(L_N(u;a)\big) = \sum_{\substack{n\leqslant N\\ n\equiv a \,(\mathrm{mod}\, q)}} \frac{1}{n}\varrho_0\Big(\frac{m_n}{n}\Big) + O\Big(\frac{1}{a^2}\Big),$$

and, by (3·10),

$$\begin{aligned}\mathbb{V}\big(L_N(u;a)\big) - \mathbb{E}\big(L_N(u;a)\big) &\ll \sum_{\substack{1\leqslant\nu<n\leqslant N\\ \nu,n\equiv a \,(\mathrm{mod}\, q)}} \Big\{\frac{\log 2n/\nu}{n^2} + \frac{\log 2\nu}{n\nu^2}\Big\}\\ &\ll \sum_{\substack{q<n\leqslant N\\ n\equiv a \,(\mathrm{mod}\, q)}} \Big\{\frac{1}{nq} + \frac{\log 2a}{a^2 n}\Big\} \ll \frac{\log N}{q}\Big\{\frac{1}{q} + \frac{\log 2a}{a^2}\Big\}.\end{aligned}$$

Summing over $a \in [1, q]$, we get

$$\mathbb{V}\big(L_N(u)\big) \leqslant q \sum_{1\leqslant a\leqslant q} \mathbb{V}\big(L_N(u;a)\big) \ll q\mathbb{E}(L_N(u)) + \log N \ll q \log N.$$

This is all needed.

Acknowledgment. The authors express warm thanks to Eugenijus Manstavičius for precise and useful comments on a preliminary version of the manuscript.

References

[1] R. Arratia, A.D. Barbour, S. Tavaré, On random polynomials over finite fields, *Math. Proc. Cambridge Philos. Soc.* **114** (1993), no. 2, 347–368.

[2] R. Arratia, A.D. Barbour, S. Tavaré, The Poisson-Dirichlet distribution and the scale-invariant Poisson process, *Combin. Probab. Comput.* **8** (1999), no. 5, 407–416.

[3] R. Arratia, A.D. Barbour, S. Tavaré, Limits of logarithmic combinatorial structures, *Ann. Probab.* **28** (2000), no. 4, 1620–1644.

[4] C. Bhattacharjee & L. Goldstein, Dickman approximation in simulation, summations and perpetuities, *Bernoulli* **25** (2019), no. 4A, 2758–2792.

[5] P. Billingsley, On the distribution of large prime factors, *Period. Math. Hungar.* **2** (1972), 283–289.

[6] M. Car, Théorèmes de densité dans $\mathbb{F}_q[X]$, *Acta Arith.* **48** (1987), no. 2, 145–165.

[7] W.-M. Chen & H.-K. Hwang, Analysis in distribution of two randomized algorithms for finding the maximum in a broadcast communication model, *J. Algorithms* **46** (2003), no. 2, 140–177.

[8] L. Devroye, Simulating perpetuities, *Meth. Comput. Appl. Probab.* **3** (2001), 97–115.

[9] L. Devroye & O. Fawzi, Simulating the Dickman distribution, *Statist. Probab. Lett.* **80** (2010), no. 3-4, 242–247.

[10] K. Dickman, On the frequency of numbers containing primes of a certain relative magnitude, *Ark. Mat. Astr. Fys.* **22** (1930), 1–14.

[11] P.D.T.A. Elliott, Probabilistic number theory : mean value theorems, Grundlehren der Math. Wiss. 239; Probabilistic number theory : central limit theorems, *ibid.* 240, Springer-Verlag, New York, Berlin, Heidelberg 1979, 1980.

[12] J.A. Fill & M.L. Huber, Perfect simulation of Vervaat perpetuities *Electron. J. Probab.* **15** (2010), no. 4, 96–109.

[13] R. Giuliano, Z.S. Szewczak & M.J.G. Weber, Almost sure local limit theorem for the Dickman distribution, *Period. Math. Hungar.* **76**, no. 2 (2018), 155–197.

[14] C.M. Goldie & R. Grübel, Perpetuities with thin tails, *Adv. in Appl. Probab.* **28** (1996), no. 2, 463–480.

[15] H.K. Hwang & T.-H. Tsai, Quickselect and the Dickman Function, *Combinatorics, Probability and Computing* **11**, (2002) 353–371.

[16] J.F.C. Kingman, The population structure associated with the Ewens sampling formula, *Theoretical Population Biology* **11** (1977), 274–283.

[17] A. Knopfmakher & E. Manstavičius, On the largest degree of an irreducible factor of a polynomial in $\mathbb{F}_q[X]$ (Russian), *Liet. Mat. Rink.* **37** (1997), no. 1, 50–60; translation in: *Lithuanian Math. J.* **37** (1997), no. 1, 38–45.

[18] E. Manstavičius, Remarks on elements of semigroups that are free of large prime factors (Russian), *Liet. Mat. Rink.* **32** (1992), no. 4, 512–525; translation in: *Lithuanian Math. J.* **32** (1992), no. 4, 400–409 (1993).

[19] E. Manstavičius & R. Petuchovas, Local probabilities and total variation distance for random permutations, *Ramanujan J.* **43** (2017), no. 3, 679–696.

[20] R. G. Pinsky, A natural probabilistic model on the integers and its relation to Dickman-type distributions and Buchstab's function, in: *Probability and analysis in interacting physical systems*, 267–294, Springer Proc. Math. Stat. 283, Springer, Cham, 2019.

[21] L.A. Shepp & S.P. Lloyd, Ordered cycle lengths in a random permutation, *Trans. Amer. Math. Soc.* **121** (1966), 340–357.

[22] G. Tenenbaum, Crible d'Ératosthène et modèle de Kubilius, in: K. Győry, H. Iwaniec, J. Urbanowicz (eds.), *Number Theory in Progress*, Proceedings of the conference in honor of Andrzej Schinzel, Zakopane, Poland 1997, 1099–1129, Walter de Gruyter, Berlin, New York.

[23] G. Tenenbaum, A rate estimate in Billingsley's theorem for the size distribution of large prime factors, *Quart. J. Math. (Oxford)* **51** (2000), 385–403.

[24] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Graduate Studies in Mathematics 163, Amer. Math. Soc. 2015.

[25] E.C. Titchmarsh, *The theory of the Riemann zeta-function*, Oxford University Press, 1951.

Régis de la Bretèche
Université de Paris, Sorbonne Université, CNRS,
Institut de Math. de Jussieu-Paris Rive Gauche,
F-75013 Paris,
France
`regis.de-la-breteche@imj-prg.fr`

Gérald Tenenbaum
Institut Élie Cartan
Université de Lorraine
BP 70239
54506 Vandœuvre-lès-Nancy Cedex
France
`gerald.tenenbaum@univ-lorraine.fr`