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

Régis de La Bretèche, Gérald Tenenbaum

To cite this version:

Régis de La Bretèche, Gérald Tenenbaum. On strong and almost sure local limit theorems for a probabilistic model of the Dickman distribution. Lithuanian Mathematical Journal, In press, 10.1007/s10986-021-09529-6. hal-03167177

HAL Id: hal-03167177
https://hal.science/hal-03167177
Submitted on 11 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
On strong and almost sure local limit theorems for a probabilistic model of the Dickman distribution

Régis de la Bretèche & Gérard Tenenbaum

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

Abstract.

Let \( \{ Z_k \}_{k \geq 1} \) denote a sequence of independent Bernoulli random variables defined by 
\[ P(Z_k = 1) = 1/k \] \( = 1 - P(Z_k = 0) \) (\( k \geq 1 \)) and put \( T_n := \sum_{1 \leq k \leq n} k Z_k \). 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 \( \varrho(1) = 1 \) \((0 \leq u \leq 1)\). Improving on earlier work, we propose asymptotic formulae with remainders for the corresponding local and almost sure limit theorems, namely

\[
\sum_{m \geq 0} \left| P(T_n = m) - \frac{e^{-\gamma}}{n} \varrho\left(\frac{m}{n}\right) \right| = \frac{2 \log n}{\pi^2 n} \left\{ 1 + O\left(\frac{1}{\log^2 n}\right) \right\} \quad (n \to \infty),
\]
and

\[
(\forall u > 0) \sum_{n \leq N \atop \lfloor u n \rfloor} 1 = e^{-\gamma} \varrho(u) \log N + O\left((\log N)^{2/3+o(1)}\right) \quad \text{a.s.} \quad (N \to \infty),
\]

where \( \gamma \) 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 \( \varrho(1) = 1 \) \((0 \leq u \leq 1)\). It is known (see, e.g., [24; th. III.5.10]) that
\[ \int_0^\infty \varrho(u) \, du = e^\gamma, \]
where \( \gamma \) denotes Euler’s constant.

The Dickman distribution is defined as the law of a random variable \( D \) on \([0, \infty]\) with density
\[ \varrho_0(u) := e^{-\gamma} \varrho(u) \quad (u \geq 0). \]

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 an expository 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 expository 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 \( \sigma \)-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

$$
\sum_{n \geq 1} \prod_{1 \leq j \leq n} X_j,
$$

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 $\{Z_k\}_{k \geq 1}$ denotes a sequence of independent Bernoulli random variables such that $P(Z_k = 1) = 1/k = 1 - P(Z_k = 0)$ ($k \geq 1$) and if $T_n := \sum_{1 \leq k \leq n} kZ_k$, then $T_n/n$ converges weakly to $D$, viz.

$$
\lim_{n \to \infty} P(T_n \leq nu) = e^{-\gamma} \int_0^u \varrho(v) \, dv \quad (u \geq 0).
$$

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

$$(1.1) \quad v_n := \sum_{m \geq 0} \left| P(T_n = m) - \frac{e^{-\gamma}}{n} \varrho\left(\frac{m}{n}\right) \right| = o(1) \quad (n \to \infty).$$

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

$$(1.2) \quad v_n = \frac{2 \log n}{\pi^2 n} \left\{ 1 + O\left(\frac{1}{\log_2 n}\right) \right\} \quad (n \to \infty).$$

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

$$(1.3) \quad \sum_{m \geq 0} \left| P(Y_n = m) - \frac{e^{-\gamma}}{n} \varrho\left(\frac{m}{n}\right) \right| \ll \frac{1}{n} \quad (n \geq 1).$$

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.

$$
\text{d}_{TV}(T_n, Y_n) := \sum_{m \geq 0} \left| P(T_n = m) - P(Y_n = m) \right| = v_n + O\left(\frac{1}{n}\right)
$$

$$
= \frac{2 \log n}{\pi^2 n} \left\{ 1 + O\left(\frac{1}{\log_2 n}\right) \right\}.
$$

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

For $u > 0$, let $\varepsilon_n$ denote a non-negative sequence tending to 0 at infinity, and let $\{m_n\}_{n \geq 1}$ 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

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

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

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

\[
L_N(u) = \left\{ 1 + O \left( \eta_N + \frac{(\log N)^{1/2 + o(1)}}{(\log N)^{1/3}} \right) \right\} e^{-\gamma} \varrho(u) \log N,
\]

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

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

\[
\delta_m := \left| \{ n \geq 1 : m_n = m \} \right| = o(\log m) \quad (m \to \infty),
\]

and assuming only that the sequence \( \{m_n\}_{n \geq 1} \) is non-decreasing. If \( \delta_m \ll (\log m)^{\alpha} \) with \( 0 \leq \alpha < 1 \), 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 \( \varrho(v) \ll v^{-\gamma} \), we first have

\[
\frac{1}{n} \sum_{m > M(n)} \varrho \left( \frac{m}{n} \right) \ll \frac{1}{n} \sum_{m > M(n)} e^{-m(\log_2 n)/2n} \ll \frac{1}{n}.
\]

Then, we have, for all \( y \geq 0 \)

\[
\sum_{m > M(n)} \mathbb{P}(T_n = m) \leq e^{-yM(n)} \mathbb{E}(e^{yT_n}) = e^{-yM(n)} \prod_{1 \leq k \leq n} \left( 1 + \frac{e^{ky} - 1}{k} \right).
\]

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

\[
\ll \exp \left\{ \int_0^1 \frac{e^{kv} - 1}{v} \, dv \right\} \ll n^{2/\log_2 n},
\]

hence the left-hand side of (2.1) is also \( \ll 1/n \), and we infer that

\[
v_n = \sum_{1 \leq m \leq M(n)} \left| \mathbb{P}(T_n = m) - \frac{e^{-\gamma}}{n} \varrho \left( \frac{m}{n} \right) \right| + O \left( \frac{1}{n} \right).
\]

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

\[
\hat{\varrho}_0(\tau) := \int_0^1 e^{iru} \varrho_0(u) \, du = e^{I(\tau)} \quad (\tau \in \mathbb{R}).
\]

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

$$
E(e^{i\tau T_n}) = \prod_{1 \leq k \leq n} \left(1 + \frac{e^{i\tau k} - 1}{k}\right)
$$

(2.4)

$$
= \exp \left\{ S_n(\tau) + U(\tau) + W_n(\tau) + O\left(\frac{\tau}{n(1 + n|\tau|)}\right) \right\},
$$

with

$$
S_n(\tau) := \sum_{1 \leq k \leq n} \frac{e^{i\tau k} - 1}{k}, \quad W_n(\tau) := \sum_{k > n} (\frac{e^{i\tau k} - 1}{2k^2}).
$$

$$
U(\tau) := \sum_{k \geq 1} \left\{ \log \left(1 + \frac{e^{i\tau k} - 1}{k}\right) - \frac{e^{i\tau k} - 1}{k} \right\}.
$$

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

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

with

$$
V_n(\tau) := \int_0^{i\tau} \left(\frac{1}{1 - e^{-v}} - \frac{1}{v}\right) dv = \int_0^1 (e^{in\tau v} - 1)g_\tau(v) dv,
$$

$$
g_\tau(v) := \frac{i\tau}{1 - e^{-i\tau v}} - \frac{1}{v} \quad (0 \leq v \leq 1).
$$

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

$$
V_n(\tau) = V(\tau) + \frac{a(\tau)e^{in\tau} - \frac{1}{n}}{n} + O\left(\frac{1}{n^2}\right),
$$

with

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

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

$$
W_n(\tau) = \sum_{k > n} \frac{e^{2ik\tau} - 2e^{ik\tau}}{2k^2} + \frac{1}{2n} + O\left(\frac{1}{n^2}\right) = \frac{1}{2n}\left\{ 1 + O\left(\frac{1}{1 + n \min(\tau, \pi - |\tau|)}\right) \right\}.
$$

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

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

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

$$
E(e^{i\tau T_n}) = \tilde{h}_0(\tau)\left\{ 1 + F(\tau) + W_n(\tau) + O\left(\frac{\tau}{1 + n^2\tau^2}\right) \right\} \quad (n \geq 1, |\tau| \leq \pi),
$$

(2.6)

with

$$
W_n(\tau) := \frac{a(\tau)\{1 + F(\tau)\}e^{in\tau}}{n} + O\left(\frac{1}{n + n^2 \min(|\tau|, \pi - |\tau|)}\right).
$$

(2.7)
It follows that, for $m \geq 1$,

$$\mathbb{P}(T_n = m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{E}(e^{i\tau T_n}) e^{-im\tau} d\tau$$

(2.8)

$$= \frac{1}{2\pi n} \int_{-\pi}^{\pi} \hat{\theta}_0(\tau) e^{-i\tau n/m} \left\{ 1 + F_{\tau \pi/n} + W_{\tau \pi/n} + O\left(\frac{\tau}{n(1 + \tau^2)}\right) \right\} d\tau.$$  

By (2.3) and, say, [24; lemma III.5.9], we have

$$\hat{\theta}_0(\tau) = -\frac{1}{i\tau} + O\left(\frac{1}{\tau(1 + |\tau|)}\right) \quad (\tau \neq 0), \quad \hat{\theta}_0(\tau) \asymp \frac{1}{1 + |\tau|} \quad (\tau \in \mathbb{R}).$$

Therefore, the error term of (2.8) contributes $\ll 1/n^2$ to the right-hand side. Summing over $m \leq M(n)$, 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

$$\frac{1}{2\pi n} \int_{-\pi}^{\pi} \hat{\theta}_0(\tau) e^{-i\tau n/m} d\tau \quad (n \geq 1, 1 \leq m \leq M(n))$$

(2.10)

by extending the integration range to $\mathbb{R}$ and inserting the first estimate (2.9) to bound the integral over $\mathbb{R} \setminus [-\pi, \pi]$. This yields

$$\frac{1}{2\pi n} \int_{-\pi}^{\pi} \hat{\theta}_0(\tau) e^{-i\tau n/m} d\tau = \frac{\theta_0(m/n)}{n} = \frac{1}{\pi n} \int_{-\pi}^{\pi} \sin(\tau m/n) \frac{1}{\tau} d\tau + O\left(\frac{1}{n^2}\right) = \frac{(-1)^{m+1}}{\pi^2 mn} + O\left(\frac{1}{m^2 n} + \frac{1}{n^2}\right).$$

In order to estimate the contributions from $F$ and $W_n^*$ to the main term of (2.8), we use the more precise formula

$$\hat{\theta}_0(\tau) = \frac{i}{\tau} - \frac{e^{i\tau}}{\tau^2} + O\left(\frac{1}{\tau^3}\right) \quad (|\tau| \geq 1).$$

(2.12)

Writing $F(\tau) = \tau G(\tau)$, we indeed deduce from by (2.12) that

$$\int_{-\pi}^{\pi} \hat{\theta}_0(\tau) e^{-i\tau \pi/n} F\left(\frac{\pi}{n}\right) d\tau = \int_{-\pi}^{\pi} \tau \hat{\theta}_0(\tau) e^{-i\tau \pi/n} G\left(\frac{\pi}{n}\right) d\tau$$

$$= i \int_{I_n} e^{-i\tau \pi} G(\tau) \left(1 + \frac{e^{i\tau \pi}}{n\tau}\right) d\tau + O\left(\frac{1}{n}\right),$$

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

$$i \int_{I_n} e^{-i\tau \pi} G(\tau) d\tau = 2\frac{(-1)^{m+1}}{\pi m} + \frac{1}{m} \int_{-\pi}^{\pi} e^{-i\tau m} G'(\tau) d\tau + O\left(\frac{1}{n}\right)$$

$$= 2\frac{(-1)^{m+1}}{\pi m} + O\left(\frac{1}{m^2} + \frac{1}{n}\right),$$

and, similarly,

$$-\frac{1}{n} \int_{I_n} e^{-i\tau \pi} G(\tau) e^{i\tau \pi} d\tau = -\frac{1}{n} \int_{I_n} G(\tau) - G(0) e^{i\tau (n-m)} d\tau + O\left(\frac{1}{n}\right) \ll \frac{1}{n}.$$

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


We can thus state that, for $n \geq 1$, $1 \leq m \leq M(n)$, we have
\begin{equation}
(2.13) \quad \frac{1}{2\pi n} \int_{-n\pi}^{n\pi} \vartheta_0(\tau) e^{-i\tau m/n} F\left(\frac{\tau}{n}\right) d\tau = \frac{(-1)^{m+1}}{\pi^2 mn} + O\left(\frac{1}{m^2 n} + \frac{1}{n^2}\right).
\end{equation}

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 $a(\tau/n)(1 + F(\tau/n))e^{i\tau}/n$ is also $\ll 1/n^2$.

The above estimates and (2.11) furnish together
\begin{equation}
(2.14) \quad \mathbb{P}(T_n = m) = \frac{1}{n} \vartheta_0\left(\frac{m}{n}\right) + O\left(\frac{1}{n^2}\right) \quad (1 \leq m \leq M(n)).
\end{equation}

Summing over $m \leq M(n)$ provides the required estimate (1.2).

3. Proof of Theorem 1.2

We first note that (2.14) yields
\begin{equation}
(3.1) \quad \mathbb{P}(T_n = m) = \frac{1}{n} \vartheta_0\left(\frac{m}{n}\right) + O\left(\frac{1}{n^2}\right) \quad (m \asymp n \geq 1).
\end{equation}

Hence, writing $L_n$ for $L_N(u)$ here and throughout,
\begin{equation}
(3.2) \quad \mathbb{E}(L_N) = \sum_{n \leq N} \frac{1}{n} \vartheta_0\left(\frac{m_n}{n}\right) + O(1) = \vartheta_0(u) \log N + O(\eta_N \log N + 1) \quad (N \geq 1).
\end{equation}

The stated result will follow from an estimate of the variance $\mathbb{V}(L_N)$. We have
\begin{equation}
(3.3) \quad \mathbb{E}(L_N^2) = \mathbb{E}(L_N) + 2 \sum_{1 \leq \nu < n \leq N} \mathbb{P}(T_\nu = m_\nu) \alpha_{\nu n},
\end{equation}

with
\begin{equation}
\alpha_{\nu n} := \mathbb{P}(T_n - T_\nu = m_n - m_\nu) \quad (1 \leq \nu < n \leq N).
\end{equation}

Let us initially assume that $\{m_n\}_{n \geq 1}$ is strictly increasing and hence that $u \geq 1$. 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 $\nu \leq n(u + 1)/(u + 2)$ in the sequel.

By (2.4), we have, writing $\varphi_j(\tau) := \mathbb{E}(e^{i\tau T_j})$ ($j \geq 1$),
\begin{equation}
(3.4) \quad \alpha_{\nu n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\varphi_n(\tau)}{\varphi_\nu(\tau)} e^{-i(m_n - m_\nu)\tau} d\tau = \beta_{\nu n} + \Delta_{\nu n},
\end{equation}

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

Now we may derive from (2.6) that
\begin{equation}
\beta_{\nu n} = \frac{1}{2\pi n} \int_{-n/2\nu}^{n/2\nu} \vartheta_0(\tau) e^{-i(m_n - m_\nu)\tau/n} d\tau.
\end{equation}

Invoking (2.9) and (2.6)-(2.7) with $\nu$ in place of $n$ to estimate the contribution of the error term, we get
\begin{equation}
(3.5) \quad \beta_{\nu n} = \beta^*_{\nu n} + O\left(\frac{\log n}{n^2}\right).
\end{equation}
Carrying back into (3·6) we used the bounds and note right away that $S_{\nu}$ since the assumption that $\nu \leq (u+1)n/(u+2)$ and observing that 

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

we may write

$$\frac{1 + F(\tau/n)}{\varphi(\tau/n)} = 1 + \sum_{j \geq 1} \mu_{\nu j} \left(\frac{\tau}{n}\right)^j \quad (|\tau| \leq n/2\nu)$$

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

$$\frac{1}{2\pi n} \sum_{j \geq 1} \frac{\mu_{\nu j}}{n^j} \int_{-n/2\nu}^{n/2\nu} \left\{ \tau^j e^{i\tau(m_n - m_\nu)/n} + O\left(\frac{\tau^{j-1}}{1 + |\tau|}\right) \right\} d\tau$$

$$\ll \frac{1}{n} \sum_{j \geq 1} \mu_{\nu j} \int_{-1/2\nu}^{1/2\nu} \left\{ \nu^{j-1} e^{i\nu(m_n - m_\nu)} + O\left(\nu^{j-1} \frac{1 + |\nu|}{1 + |\nu|}\right) \right\} d\nu$$

$$\ll \frac{\nu}{n(n - \nu)} + \frac{\nu}{n^2} + \frac{\nu \log(2n/\nu)}{n^2} \ll \frac{\nu \log(2n/\nu)}{n^2},$$

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

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

$$\beta_{\nu n}^* = \frac{1}{n} \left[ G_0\left(\frac{m_n - m_\nu}{n}\right) + O\left(\frac{\nu \log(2n/\nu)}{n^2}\right) \right] \quad (1 \leq \nu < n).$$

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

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

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

$$\frac{\varphi_\nu(\tau)}{\varphi(\tau)} = \frac{\nu}{n} \prod_{\nu < k \leq n} \left(1 + \frac{e^{ik\tau}}{k-1}\right) = \frac{\nu}{n} \exp \left\{ S_{\nu n}(\tau) + O\left(\frac{1}{\nu + |\tau|\nu^2} + \frac{1}{\nu + (\pi - |\tau|)\nu^2}\right) \right\},$$

where we used the bounds

$$\sum_{\nu < k \leq n} \frac{e^{ik\tau}}{k(k-1)} \ll \frac{1}{\nu + |\tau|\nu^2}, \quad \sum_{\nu < k \leq n} \frac{e^{ik\tau}}{(k-1)^2} \ll \frac{1}{\nu + |\tau|\nu^2} + \frac{1}{\nu + (\pi - |\tau|)\nu^2},$$

$$\sum_{\nu < k \leq n} \frac{e^{ikj\tau}}{(k-1)^j} \ll \frac{1}{\nu^{j-1}} (j > 2).$$

Carrying back into (3·4), we obtain

$$\Delta_{\nu n} = \frac{\nu}{\pi n} \int_{1/2\nu}^{\pi} e^{S_{\nu n}(\tau) - i(m_n - m_\nu)\tau} d\tau + O\left(\frac{\log 2\nu}{n\nu}\right).$$
Now observe that $|S'_{v \nu}(\tau)| \leq \pi/|\tau| \leq \frac{1}{2}(n - \nu) \leq \frac{1}{2}(m_n - m_\nu)$ if, say $\nu \leq n/15$ 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 $1/2\nu < |\tau| \leq 10(u + 1)/\nu$ to the same integral is trivially $\ll 1/\nu$, we finally get

$$\Delta_{v \nu} \ll \frac{\log 2 \nu}{n^2} + \frac{\nu}{n^2} \quad (1 \leq \nu \leq (u + 1)n/(u + 2)).$$

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

$$\alpha_{v \nu} = \frac{1}{n} \varrho_0 \left( \frac{m_{v}}{n} \right) + O \left( \frac{\log(2\nu \nu)}{n^2} + \frac{\log 2 \nu}{n^2} \right) \quad (1 \leq \nu < n).$$

Carrying back into (3-3) and applying (3-1) for the pair $(\nu, m_\nu)$ yields

$$\mathbb{E}(L_N^2) - \mathbb{E}(L_N) = 2 \sum_{1 \leq \nu < n \leq N} \left\{ \varrho_0 \left( \frac{m_{\nu}}{n} \right) \varrho_0 \left( \frac{m_{v}}{\nu} \right) \frac{1}{\nu n} + O \left( \frac{\log(2\nu \nu)}{n^2} + \frac{\log 2 \nu}{n^2} \right) \right\},$$

and hence

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

Selecting $N = N_k := 2^k$ for $k \geq 1$, 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 $\{m_n\}_{n \geq 1}$. Accordingly, we fix $u > 0$. By hypothesis, for some integer $q = q_N \geq 2$ such that $q_N = o\left( \log N \right)$ as $N \rightarrow \infty$, we have $m_n > m_\nu$ whenever $n - \nu \geq q_N$.

Put

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

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

$$\mathbb{E}(L_N(u; a)) = \sum_{\substack{n \leq N \\cap \\nu \equiv a \mod q \\cap \nu T_n = m_n}} \frac{1}{n} \varrho_0 \left( \frac{m_{a}}{n} \right) + O \left( \frac{1}{n^2} \right),$$

and, by (3-10),

$$\mathbb{V}(L_N(u; a)) - \mathbb{E}(L_N(u; a)) \ll \sum_{\substack{1 \leq \nu < n \leq N \\cap \nu \equiv a \mod q \\cap \nu T_n = m_n}} \left\{ \frac{\log 2 \nu}{n^2} + \frac{\log 2 \nu}{n^2} \right\} \ll \sum_{\substack{q \leq n \leq N \\cap \nu \equiv a \mod q \\cap \nu T_n = m_n}} \left\{ \frac{1}{n^2} + \frac{\log 2 a}{a^2 n} \right\} \ll \log N \left\{ \frac{1}{q} + \frac{\log 2 a}{a^2} \right\}.$$
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. Bhattacherjee & 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 $F_q[X]$, *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, *Math. 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, *Combinatorics, Probability and Computing* **11**, (2002) 353–371.
[14] 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.
[15] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Graduate Studies in Mathematics **163**, Amer. Math. Soc. 2015.

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 Elie Cartan
Université de Lorraine
BP 70239
54506 Vandœuvre-lès-Nancy Cedex
France
gerard.tenenbaum@univ-lorraine.fr