engineering journals
Redefining Knowledge

uniform distribution theory

DOI: 10.66833/UDT-2026-0008
Unif. Distrib. Theory 21 (2026), no.1, 1-21

Distributions modulo one of slowly varying functions of primes

Arno Berger and Matthew Jamie Liu

ABSTRACT. For a common class of slowly varying functions f, this note studies the empirical distributions modulo one of sequences $(f(p_n))$ where p_n denotes the n-th prime number. By means of a Kantorovich metric for probability measures on the circle, simple explicit bounds are established for the deviation of these empirical distributions from their limit set, a circle consisting of all rotated copies of a single exponential distribution. Despite their simplicity the established bounds often are sharp. They complement known results in the literature and give rise to several intriguing quantitative questions.

Communicated by Friedrich Pillichshammer

1. Introduction

Given a smooth function $f : \mathbb{R}_+ \to \mathbb{R}$ assume that

$$\lim_{r\to\infty} rf'(r) = b \tag{1.0.1}$$

for some $b \in \mathbb{R}$. With p_n denoting the n-th prime for every $n \in \mathbb{N}$, this note presents an analysis of the distributions modulo one (mod 1, for short) of the sequence $(f(p_n))$. The main analytic tool is the compact space $\mathcal{P}_{\mathbb{T}}$ of all Borel probability measures on the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, metrized by the Kantorovich distance $d_{\mathbb{T}}$. To preview the main result, for every $N \in \mathbb{N}$ let $\nu_N \in \mathcal{P}_{\mathbb{T}}$ be the empirical distribution mod 1 of $f(p_1), f(p_2), \dots, f(p_N)$, or more formally $\nu_N = N^{-1}\sum_{n=1}^{N} \delta_{f(p_n)+\mathbb{Z}}$ with $\delta_{t+\mathbb{Z}} \in \mathcal{P}_{\mathbb{T}}$ denoting the Dirac measure concentrated at $t + \mathbb{Z}$. Moreover, for every $t \in \mathbb{R}$ let $\eta_{b,t} \in \mathcal{P}_{\mathbb{T}}$ be the distribution of $-bX + t + \mathbb{Z}$ where the real random variable X is standard exponential, that is, $\operatorname{Prob}\{X > r\} = e^{-r}$ for every $r \in \mathbb{R}_+$. Note that $\eta_{b,t+1} = \eta_{b,t}$ and $\eta_{0,t} = \delta_{t+\mathbb{Z}}$ for every $t \in \mathbb{R}$. Finally, fix an increasing smooth function $\phi : \mathbb{R}_+ \to \mathbb{R}$ so that

© 2026 Berger, A. and Liu, M. J., Published by Engineering Journals.
2020 Mathematics Subject Classification: 11K06, 11K31, 26A12, 28A33, 60B10.
Keywords: Slowly varying function, (uniform) distribution modulo one, empirical distribution, Kantorovich metric, prime number theorem.
Licensed under the Creative Commons BY-NC-ND 4.0 International Public License.

$\phi(r) = r\log r + r\log\log r - r + \mathcal{O}(r/\log r)$ as $r \to \infty$. With these ingredients, the main result of this note provides the following precise description of the distributions mod 1 of $(f(p_n))$, i.e., of the limit points of $\{\nu_N : N \in \mathbb{N}\}$ in $\mathcal{P}_{\mathbb{T}}$.

Theorem 1.1. *Assume that $f : \mathbb{R}_+ \to \mathbb{R}$ is smooth and satisfies* (1.0.1) *for some $b \in \mathbb{R}$. Then, for every sufficiently large $N \in \mathbb{N}$,*

$$d_{\mathbb{T}}(\nu_N, \eta_{b,f(p_N)}) \le \frac{1}{4\log N} + \frac{1}{N}\int_1^N \big|\phi(r)f'(\phi(r)) - b\big|\,\mathrm{d}r\,. \tag{1.0.2}$$

A proof of Theorem 1.1 is presented in Section 4 below. To appreciate the result, note first that the integrand in (1.0.2) tends to 0 as $r \to \infty$, and hence $\lim_{N\to\infty} d_{\mathbb{T}}(\nu_N, \eta_{b,f(p_N)}) = 0$. As $N \to \infty$, therefore, the (discrete) empirical distribution ν_N resembles the (absolutely continuous) distribution $\eta_{b,f(p_N)}$ ever more closely. Correspondingly, the sequence $(f(p_n))$ does not have a (unique) distribution mod 1, and hence in particular is *not* uniformly distributed; instead, it has a one-parameter family of limit distributions, namely $\{\eta_{b,s} : s \in [0,1[\}$ which may be interpreted as a circle in $\mathcal{P}_{\mathbb{T}}$; see Section 2 below. As $N \to \infty$, clearly ν_N approaches that circle at a $d_{\mathbb{T}}$-rate not slower than the RHS in (1.0.2). As discussed in some detail in Section 4 also, that rate often is sharp up to a multiplicative constant. For a simple preview example illustrating this sharpness, fix $c \in \mathbb{R}$ and let $f : \mathbb{R}_+ \to \mathbb{R}$ be smooth with

$$f(r) = \log r + c\log\log r \qquad \forall r \ge e\,. \tag{1.0.3}$$

For any such f, clearly $b = 1$, and the RHS in (1.0.2) is $\mathcal{O}(1/\log N)$ as $N \to \infty$. A careful analysis of the deviation (or signed measure) $\nu_N - \eta_{1,f(p_N)}$, outlined in Section 4, shows that if $c \ne -1$ then

$$\lim_{N\to\infty} \log N\, d_{\mathbb{T}}(\nu_N, \eta_{1,f(p_N)}) \in \mathbb{R}_+\,, \tag{1.0.4}$$

a fact established in [3, Thm. 5.5] for $c = 0$, whereas if $c = -1$ then

$$\lim_{N\to\infty} (\log N)^2\, d_{\mathbb{T}}(\nu_N, \eta_{1,f(p_N)}) \in \mathbb{R}_+\,.$$

For every f as in (1.0.3) with $c \ne -1$, therefore, the bound in (1.0.2) is sharp; only for $c = -1$ is it too pessimistic, by a factor $\log N$.

To put Theorem 1.1 in perspective, notice how it combines and complements the following three well-established aspects of (uniform) distribution theory:

(i) There exists an extensive literature on the distributions mod 1 of increasing sequences (x_n) in $\mathbb{R}$ with $x_n \to \infty$ but $x_{n+1} - x_n \to 0$; see, e.g., [2, 13, 14, 21, 24] and the many references therein. In particular, if

$$\lim_{n\to\infty} n(x_{n+1} - x_n) = b \tag{1.0.5}$$

for some $b \in \mathbb{R}_+$ then $(x_n) = (f(n))$ with an appropriate f satisfying (1.0.1), and $\lim_{r\to\infty} f(r)/\log r = b$. For such f, authors have studied the distributions mod 1 of (x_n), i.e., the limit points of $\{\tilde{\nu}_N : N \in \mathbb{N}\}$ in $\mathcal{P}_{\mathbb{T}}$ where, in analogy to ν_N introduced earlier, $\tilde{\nu}_N$ is the empirical distribution mod 1 of $f(1), f(2), \dots, f(n)$, or more formally $\tilde{\nu}_N = N^{-1} \sum_{n=1}^{N} \delta_{f(n)+\mathbb{Z}} \in \mathcal{P}_{\mathbb{T}}$. For instance, [2] proved that

$$\lim_{N\to\infty} \frac{N}{\sqrt{\log N}}\, d_{\mathbb{T}}(\tilde{\nu}_N, \eta_{b,f(N)}) \in \mathbb{R}_+\,, \tag{1.0.6}$$

provided that f is sufficiently close to $b \log$. Notice how in this situation $\tilde{\nu}_N$ approaches $\eta_{b,f(N)}$ much faster than ν_N approaches $\eta_{b,f(p_N)}$; see also FIGURE 1.

(ii) For sequences (x_n) as in (i), among others, the behaviour of $\tilde{\nu}_N$ as $N \to \infty$ is usually studied utilizing distribution functions; see, e.g., [14, 21]. Specifically, one considers non-decreasing functions with $0 \le F^- \le F^+ \le 1$, given by

$$F^-(s) = \liminf_{N\to\infty} F_{\tilde{\nu}_N}(s)\,, \quad F^+(s) = \limsup_{N\to\infty} F_{\tilde{\nu}_N}(s) \qquad \forall s \in [0,1[\,,$$

and referred to as the *lower* and *upper distribution function* mod 1 of (x_n), respectively. Though straightforward, usage of $F^{\pm}$ has several unavoidable shortcomings. For instance, $F^{\pm}$ may not be right-continuous, let alone be the distribution function of $\nu^{\pm} \in \mathcal{P}_{\mathbb{T}}$, say; even if it is, $\nu^{\pm}$ may fail to be a limit point for $(\tilde{\nu}_N)$. Moreover, if $F(s) = \lim_{j\to\infty} F_{\tilde{\nu}_{N_j}}(s)$ for almost all $s \in [0,1[$ and some increasing sequence (N_j) in $\mathbb{N}$, then clearly $F^- \le F \le F^+$, but not every (non-decreasing and right-continuous) function with the latter property is a distribution function of (x_n). In addition, one may encounter outright artifacts, e.g., for $(x_n) = (\log\log n)$ where $b = 0$ in (1.0.5) and $F(s) \equiv c$ is a distribution function for *every* $0 \le c \le 1$; when interpreted as probability measures on $\mathbb{T}$, all these functions are but different manifestations of $\delta_{0+\mathbb{Z}}$.

(iii) A sequence of considerable (historical as well as present) interest is $(x_n) = (\log p_n)$. Accordingly, distributional properties of many slowly varying sequences obtained from primes have been studied extensively [12, 17, 18, 19, 22]. Notice, however, that (1.0.5) fails for $(x_n) = (\log p_n)$, as

$$\liminf_{n\to\infty} n(\log p_{n+1} - \log p_n) = 0\,, \quad \limsup_{n\to\infty} n(\log p_{n+1} - \log p_n) = \infty\,. \tag{1.0.7}$$

This is indicative of the somewhat irregular growth of (p_n). Still, by the prime number theorem,

$$\log p_n = \log n + \log\log n + \mathcal{O}\left(\frac{\log\log n}{\log n}\right) \quad \text{as } n \to \infty\,,$$

so it is plausible that the distributions mod 1 of $(\log p_n)$ may be the same as those of $(f(n))$ with f as in (1.0.3) and $c = 1$. Utilizing quite different approaches, [3, 18, 19] all demonstrated that this indeed is the case; see also Figure 1.

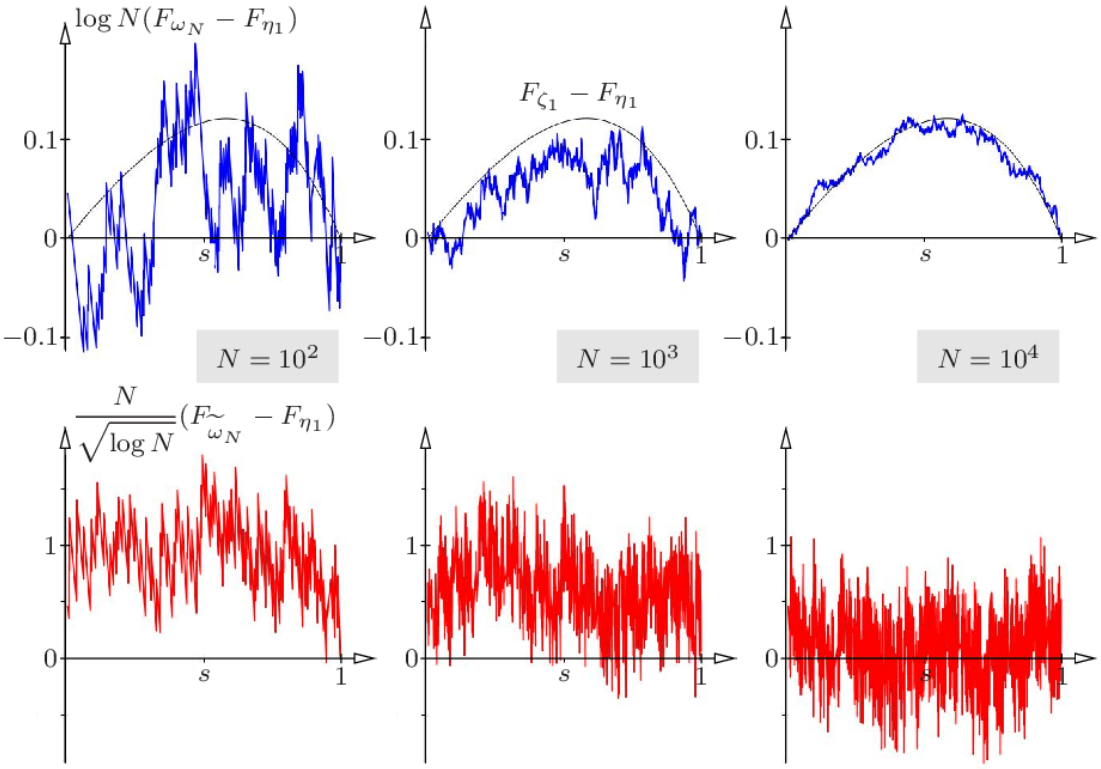


Figure 1: With $\omega_N = \nu_N \circ R^{-1}_{-\log p_N}$ and $\tilde{\omega}_N = \tilde{\nu}_N \circ R^{-1}_{-\log N}$ denoting the empirical distribution mod 1 of $\left\{\log\frac{p_n}{p_N} : 1 \le n \le N\right\}$ and $\left\{\log\frac{n}{N} : 1 \le n \le N\right\}$, respectively, the sequences (ω_N) and $(\tilde{\omega}_N)$ both converge to the "wrapped" exponential distribution $\eta_1 = \eta_{1,0}$, but do so at the very different rates $(\log N)^{-1}$ (see (1.0.4) for $c = 0$; top) and $N^{-1}\sqrt{\log N}$ (see (1.0.6) for $b = 1$; bottom). While $\omega_N \to \eta_1$ takes place along the distinguished "direction" $\zeta_1 - \eta_1$, with $\zeta_1 \in \mathcal{P}_{\mathbb{T}}$ defined properly in Section 2, the convergence $\tilde{\omega}_N \to \eta_1$ appears to be "unbiased".

In a nutshell, then, Theorem 1.1 combines aspects (i) and (iii) by more generally considering $(x_n) = (f(p_n))$ for all f satisfying (1.0.1). As far as the authors have been able to ascertain, the only previous studies directly pertinent to this setup are [18, 19] which establish upper bounds on differences between distribution functions. These bounds tend to be quite complicated and rather rough. Thus, while [18, 19] did inspire the present note, the reader may appreciate the relative simplicity and widespread

sharpness of Theorem 1.1, not least in light of the shortcomings inherent to distribution functions noted in (ii).

The remainder of this note is organized as follows: Section 2 briefly reviews the required analytic tools pertaining to Theorem 1.1, notably the metric $d_{\mathbb{T}}$, empirical distribution ν_N, and family of distributions $(\eta_{b,t})$. Section 3 establishes an auxiliary upper bound on $d_{\mathbb{T}}(\tilde{\nu}_N, \eta_{b,f(N)})$ that supersedes [3, Thm. 4.1] and may be of independent interest. Finally, Section 4 provides a proof of Theorem 1.1 and discusses aspects of its sharpness, as alluded to above.

2. Analytic preliminaries

Throughout, the sets of positive integers, integers, positive real, and real numbers, with their usual arithmetic, order and topology, are denoted by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}_+$, and $\mathbb{R}$ respectively; for convenience, let $\mathbb{I} = \{r \in \mathbb{R}_+ : r \geq 1\}$. For every $t \in \mathbb{R}$, denote by $\lfloor t \rfloor$ be the largest integer not larger than t, and by $\langle t \rangle = t - \lfloor t \rfloor$ the fractional part of t. Recall that $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is a compact metrizable space and Abelian group. For all $t, s \in \mathbb{R}$, let $d_{\mathbb{T}}(t,s) = \min_{k\in\mathbb{Z}} |t - s + k|$; when restricted to $\mathbb{T} \times \mathbb{T}$, this semi-metric yields a metric that induces the topology of $\mathbb{T}$. Geometrically, $\mathbb{T}$ is a (counter-clockwise oriented) circle, with $d_{\mathbb{T}}(x,y)$ being the normalized length of a shortest arc that joins $x, y \in \mathbb{T}$. For every $x \in \mathbb{T}$, $t \in \mathbb{R}$ let $R_t(x) = x + t$. Geometrically, R_t is the (counter-clockwise) rotation of $\mathbb{T}$ by $2\pi t$, and hence an isometry of $\mathbb{T}$.

Denote by $C(\mathbb{T})$ the Banach space of all continuous functions $h : \mathbb{T} \to \mathbb{R}$ with the usual $\|\cdot\|_\infty$-norm, and by $\mathcal{P}_{\mathbb{T}}$ the set of all Borel probability measures on $\mathbb{T}$, endowed with the topology of weak convergence. With each $\mu \in \mathcal{P}_{\mathbb{T}}$ associate its distribution function $F_\mu : \mathbb{R} \to \mathbb{R}$ given by

$$F_\mu(t) = \mu(\{s + \mathbb{Z} : 0 \leq s \leq \langle t \rangle\}) + \lfloor t \rfloor \qquad \forall t \in \mathbb{R}\,.$$

Thus F_μ is non-decreasing and right-continuous. Also, $t \mapsto F_\mu(t) - t$ is 1-periodic, and $F_\mu(0) \geq 0$ as well as $\lim_{t\uparrow 1} F_\mu(t) = 1$. Every function with these properties is the distribution function of one and only one $\mu \in \mathcal{P}_{\mathbb{T}}$. Examples of elements of $\mathcal{P}_{\mathbb{T}}$ relevant for this note include the Dirac measure δ_x at $x \in \mathbb{T}$ for which $F_{\delta_x}(t) \equiv \lfloor t - \langle x \rangle \rfloor + 1$; the Lebesgue measure $\lambda_{\mathbb{T}}$ for which $F_{\lambda_{\mathbb{T}}}(t) \equiv t$; and two one-parameter families (η_b), (ζ_b) for which for every $b \in \mathbb{R} \setminus \{0\}$,

$$F_{\eta_b}(s) = \frac{e^{s/b} - 1}{e^{1/b} - 1}\,, \quad F_{\zeta_b}(s) = \frac{e^{s/b}(1 + b - s) - b}{b(e^{1/b} - 1)} + \frac{e^{s/b} - e^{1/b}}{b(e^{1/b} - 1)^2} \qquad \forall s \in [0,1]\,.$$

In probabilistic terms, η_b and ζ_b simply are the distributions on $\mathbb{T}$ of $-bX_1 + \mathbb{Z}$ and $-b(X_1 + X_2) + \mathbb{Z}$, respectively, where the real random variables X_1, X_2 are independent and standard exponential.

Recall that $\mathcal{P}_{\mathbb{T}}$ is compact and metrizable, with $\lim_{n\to\infty} \mu_n = \mu$ in $\mathcal{P}_{\mathbb{T}}$ precisely if $\lim_{n\to\infty} \int_{\mathbb{T}} h \,\mathrm{d}\mu_n = \int_{\mathbb{T}} h \,\mathrm{d}\mu$ for every $h \in C(\mathbb{T})$. A familiar metric metrizing $\mathcal{P}_{\mathbb{T}}$,

henceforth also denoted by $d_{\mathbb{T}}$ for simplicity and often referred to as the *Kantorovich* (or 1-*Wasserstein*) *metric*, is given by

$$d_{\mathbb{T}}(\mu,\nu) = \sup_{h\in \mathrm{Lip}_1(\mathbb{T})} \int_{\mathbb{T}} h \,\mathrm{d}(\mu-\nu) \qquad \forall \mu,\nu\in\mathcal{P}_{\mathbb{T}}\,,$$

where $\mathrm{Lip}_1(\mathbb{T}) = \{h\in C(\mathbb{T}) : |h(x)-h(y)| \le d_{\mathbb{T}}(x,y) \ \ \forall x,y\in\mathbb{T}\}$. Note that $x\mapsto \delta_x$ yields an isometric embedding of $\mathbb{T}$ into $\mathcal{P}_{\mathbb{T}}$. An easy-to-use explicit formula for $d_{\mathbb{T}}$ is

$$d_{\mathbb{T}}(\mu,\nu) = \min_{t\in\mathbb{R}} \int_0^1 |F_\mu(s) - F_\nu(s) - t|\,\mathrm{d}s \qquad \forall \mu,\nu\in\mathcal{P}_{\mathbb{T}}\,; \tag{2.0.1}$$

see [3, 8, 10]. As an application of (2.0.1), notice that $d_{\mathbb{T}}(\mu,\lambda_{\mathbb{T}}) \le \frac{1}{4}$ for every $\mu\in\mathcal{P}_{\mathbb{T}}$, with equality if and only if $\mu=\delta_x$ for some $x\in\mathbb{T}$. Moreover,

$$d_{\mathbb{T}}(\eta_b,\lambda_{\mathbb{T}}) = |b|\log\Big(\cosh\frac{1}{4b}\Big) < \frac{1}{4}\,, \quad d_{\mathbb{T}}(\eta_b,\delta_{0+\mathbb{Z}}) = b\tanh\frac{1}{4b} < \frac{1}{4}$$

for all $b\in\mathbb{R}\setminus\{0\}$, which shows that $\lim_{|b|\to\infty}\eta_b = \lambda_{\mathbb{T}}$ and $\lim_{b\to 0}\eta_b = \delta_{0+\mathbb{Z}}$; for convenience, let $\eta_0 = \delta_{0+\mathbb{Z}}$.

Given $\mu\in\mathcal{P}_{\mathbb{T}}$ and any (Borel measurable) map $S:\mathbb{T}\to\mathbb{T}$, recall that $\mu\circ S^{-1}$, the *push-forward* of μ under S, is the unique $\nu\in\mathcal{P}$ with $\nu(B) = \mu(S^{-1}(B))$ for every Borel set $B\subset\mathbb{T}$. For example, $\delta_x\circ R_t^{-1} = \delta_{x+t}$ and $\lambda_{\mathbb{T}}\circ R_t^{-1} = \lambda_{\mathbb{T}}$ for all $x\in\mathbb{T}$, $t\in\mathbb{R}$. As well, with $\eta_{b,t}$ as in the Introduction,

$$\eta_{b,t} = \eta_b\circ R_t^{-1} \qquad \forall b,t\in\mathbb{R}\,,$$

and hence in particular $\eta_{b,k} = \eta_b$ for all $k\in\mathbb{Z}$. Note that $\mu\mapsto\mu\circ R_t^{-1}$ yields an isometry of $\mathcal{P}_{\mathbb{T}}$. A straightforward calculation shows that for every $b\in\mathbb{R}\setminus\{0\}$,

$$4|b|\log\Big(\cosh\frac{1}{4b}\Big)\, d_{\mathbb{T}}(t,s) \le d_{\mathbb{T}}(\eta_{b,t},\eta_{b,s}) \le \Big|\tanh\frac{1}{4b}\Big|\, d_{\mathbb{T}}(t,s) \quad \forall t,s\in\mathbb{R}\,, \tag{2.0.2}$$

where the left and right inequality is readily seen to be sharp letting $d_{\mathbb{T}}(t,s) = \frac{1}{2}$ and $d_{\mathbb{T}}(t,s)\to 0$, respectively.

Given any sequence (x_n) in $\mathbb{R}$, for every $N\in\mathbb{N}$ denote by $\mu_N[x_n]\in\mathcal{P}_{\mathbb{T}}$ the empirical distribution mod 1 of $x_1, x_2, \dots, x_N$, that is,

$$\mu_N[x_n] = \frac{1}{N}\sum_{n=1}^N \delta_{x_n+\mathbb{Z}} \qquad \forall N\in\mathbb{N}\,.$$

In other words, $\mu_N[x_n]$ is the unique $\mu\in\mathcal{P}_{\mathbb{T}}$ with $\int_{\mathbb{T}} h\,\mathrm{d}\mu = \frac{1}{N}\sum_{n=1}^N h(x_n+\mathbb{Z})$ for every $h\in C(\mathbb{T})$. Whenever (x_n) is clear from the context, $\mu_N[x_n]$ may simply be written as μ_N. Let $\mathcal{M}[x_n]$ be the set of all limit points of $\{\mu_N : N\in\mathbb{N}\}$ in $\mathcal{P}_{\mathbb{T}}$. It is well known that $\mathcal{M}[x_n]$ is non-empty, compact, and connected, and every subset of $\mathcal{P}_{\mathbb{T}}$ with these three properties equals $\mathcal{M}[x_n]$ for some (increasing) sequence (x_n) in $\mathbb{R}$; see, e.g., [3, Sec. 3], [14, Sec. 2.4] or [21, Sec. 2.2]. For the sequences considered in this note,

typically $\mathcal{M}[x_n]$ is very small. An extreme situation occurs if $\mathcal{M}[x_n] = \{\mu\}$ for some $\mu \in \mathcal{P}_{\mathbb{T}}$, that is, if $\lim_{N\to\infty} \mu_N = \mu$, in which case (x_n) is said to *have distribution* μ mod 1. The most classical case is that $\mathcal{M}[x_n] = \{\lambda_{\mathbb{T}}\}$, referred to as (x_n) being *uniformly distributed* (*u.d.*) mod 1; see, e.g., [13, 14]. Although for most sequences (x_n) studied herein the set $\mathcal{M}[x_n]$ is *not* a singleton, it nonetheless is very simple: In all situations of interest, $\mathcal{M}[x_n]$ turns out to equal $\{\eta_{b,s} : s \in [0,1[\}$ for some $b \in \mathbb{R}$ and hence may be interpreted as a circle in $\mathcal{P}_{\mathbb{T}}$; see [3, Sec. 2].

Remark 2.1. This note uses only the metric $d_{\mathbb{T}}$ on $\mathcal{P}_{\mathbb{T}}$. As discussed in [3, 7, 10, 20], relatively few precise quantitative results regarding $d_{\mathbb{T}}$ are recorded in the literature. By contrast, a vast literature exists utilizing the metric d_∞ on $\mathcal{P}_{\mathbb{T}}$, given by $d_\infty(\mu,\nu) = \|F_\mu - F_\nu\|_\infty$. Note that $(\mathcal{P}_{\mathbb{T}}, d_\infty)$ is complete but not separable, let alone compact, its topology being much finer. Clearly, $d_{\mathbb{T}} \le d_\infty$, but unlike the former, d_∞ is not induced in a natural way by a metric on $\mathbb{T}$. Correspondingly, d_∞ is not invariant under $\mu \mapsto \mu \circ R_t^{-1}$. Although the rates of convergence may differ greatly, $d_\infty(\mu_n,\mu) \to 0$ precisely if $d_{\mathbb{T}}(\mu_n,\mu) \to 0$, provided that $\mu \in \mathcal{P}_{\mathbb{T}}$ has no atoms. For instance, a sequence (x_n) in $\mathbb{R}$ is u.d. mod 1 if and only if $D_N^*(x_n) := d_\infty(\mu_N[x_n], \lambda_{\mathbb{T}}) \to 0$ as $N \to \infty$. The quantity $D_N^*(x_n)$, the *star-discrepancy* of (x_n), is studied extensively in uniform distribution theory; see, e.g., [14, Ch. 2].

3. An auxiliary upper bound

Throughout this note, $\mathcal{F}$ denotes the (linear) space of all smooth functions $f : \mathbb{I} \to \mathbb{R}$ with the property that $\lim_{r\to\infty} rf'(r)$ exists (in $\mathbb{R}$). For every $f \in \mathcal{F}$, let $b_f = \lim_{r\to\infty} rf'(r)$ and $\Delta_f(r) = rf'(r) - b_f$, so $\Delta_f : \mathbb{I} \to \mathbb{R}$ is smooth, with $\lim_{r\to\infty} \Delta_f(r) = 0$. For convenience, let $\mathcal{F}_0 = \{f \in \mathcal{F} : b_f = 0\}$. Plainly, every $f \in \mathcal{F} \setminus \mathcal{F}_0$ is slowly varying at infinity and, in a way, differs but little from $b_f \log$. Now, if $f(r) = b\log r$ for some $b \in \mathbb{R}\setminus\{0\}$ and all large r then it is well known, and straightforward to see directly, that $\mu_N[f(n)]$ tends to $\eta_{b,f(N)}$ as $N \to \infty$; see, e.g., [13], [14, Sec. 1.7] or [21, Ex. 1]. It is plausible to expect that this simple asymptotic behaviour of $\mu_N[f(n)]$ persists for all $f \in \mathcal{F}$. The following quantitative result, a mild strengthening of [3, Thm. 4.1], confirms that this indeed is the case. To state it neatly, for every $f \in \mathcal{F}$ and $N \in \mathbb{N}$ let

$$D_f(N) = d_{\mathbb{T}}(\mu_N[f(n)], \eta_{b_f, f(N)}) = d_{\mathbb{T}}(\mu_N[f(n) - f(N)], \eta_{b_f})\,;$$

notice that if $b_f = 0$ then simply $\eta_{b_f,f(N)} = \delta_{f(N)+\mathbb{Z}}$ and $\eta_{b_f} = \delta_{0+\mathbb{Z}}$. With this, the result provides an *upper* bound on $D_f(N)$ for large N; in particular, it shows that $\lim_{N\to\infty} D_f(N) = 0$ for every $f \in \mathcal{F}$.

Theorem 3.1. *Assume that $f \in \mathcal{F}$. Then there exists $N_f \in \mathbb{N}$ so that*

$$D_f(N) \le \frac{1}{4N} + |b_f|\frac{\log N}{N} + \frac{1}{N}\int_1^N |\Delta_f(r)|\,\mathrm{d}r \qquad \forall N \ge N_f\,. \tag{3.0.1}$$

Proof. Fix $f \in \mathcal{F}$. For convenience, write b_f, Δ_f simply as b, Δ respectively, and for every $N \in \mathbb{N}$ let $\mu_N = \mu_N[f(n)]$ as well as $\omega_N = \eta_{b,f(N)}$.

Assume first that $f \in \mathcal{F} \setminus \mathcal{F}_0$. Since $-f \in \mathcal{F} \setminus \mathcal{F}_0$, $b_{-f} = -b$, and

$$D_f(N) = d_{\mathbb{T}}(\mu_N[f(n)], \eta_{b,f(N)}) = d_{\mathbb{T}}(\mu_N[-f(n)], \eta_{-b,-f(N)}) = D_{-f}(N)$$

for every $N \in \mathbb{N}$, no generality is lost by assuming henceforth that $b > 0$. Fix any $0 < \varepsilon < \frac{1}{2}b$ and pick $N_\varepsilon \in \mathbb{N}$ so large that $|\Delta(r)| < \varepsilon$ for all $r \geq N_\varepsilon$. Observing that $f'(r) > 0$ whenever $r \geq N_\varepsilon$, denote by $g = g(r)$ the inverse of f, defined for all $r \geq f(N_\varepsilon)$. For every $r \in \mathbb{I}$, note that

$$f(r) = f(1) + b\log r + \int_1^r \frac{\Delta(s)}{s}\,\mathrm{d}s = f(N_\varepsilon) - b\log N_\varepsilon + b\log r + \int_{N_\varepsilon}^r \frac{\Delta(s)}{s}\,\mathrm{d}s\,, \tag{3.0.2}$$

and hence in particular $0 \leq f(r) - f(N_\varepsilon) \leq (b+\varepsilon)\log r$ for all $r \geq N_\varepsilon$. Now, fix any $h \in \mathrm{Lip}_1(\mathbb{T})$. Euler summation yields for every $N \geq N_\varepsilon$,

$$\begin{aligned} N\int_{\mathbb{T}} h\,\mathrm{d}(\mu_N - \omega_N) &= \sum\nolimits_{n=1}^{N} h\circ f(n) - N\int_{\mathbb{T}} h(s+f(N))\,\mathrm{d}\eta_b(s) \\ &= c_N + Ne^{-f(N)/b}\int_{-\infty}^{f(N_\varepsilon)} h'(s)e^{s/b}\,\mathrm{d}s + \int_{f(N_\varepsilon)}^{f(N)} h'(s)g_N(s)\,\mathrm{d}s + \\ &\qquad + \int_{f(N_\varepsilon)}^{f(N)} h'(s)\left(\langle g(s)\rangle - \frac{1}{2}\right)\,\mathrm{d}s\,, \end{aligned} \tag{3.0.3}$$

where c_N depends on f, h, N_ε but $|c_N| \leq \frac{1}{2}N_\varepsilon$ for all $N \geq N_\varepsilon$, and the smooth function $g_N : [f(N_\varepsilon), f(N)] \to \mathbb{R}$ is given by

$$g_N(s) = Ne^{-f(N)/b}e^{s/b} - g(s) \qquad \forall N \geq N_\varepsilon, s \in [f(N_\varepsilon), f(N)]\,.$$

For convenience, define a function $\hat{h} : \mathbb{R} \to \mathbb{R}$ as

$$\hat{h}(t) = \int_{-\infty}^{t} h'(s)e^{(s-t)/b}\,\mathrm{d}s = \frac{1}{e^{1/b}-1}\int_0^1 h'(s+t)e^{s/b}\,\mathrm{d}s \qquad \forall t \in \mathbb{R}\,.$$

It is readily seen that $\hat{h}$ is 1-periodic, with $\left|\hat{h}(t_1) - \hat{h}(t_2)\right| \leq 2d_{\mathbb{T}}(t_1,t_2)$ for all $t_1, t_2 \in \mathbb{R}$ and $\int_0^1 \hat{h}(t)\,\mathrm{d}t = 0$. Moreover,

$$\hat{h}(t) = h(t) - \frac{1}{b(e^{1/b}-1)}\int_0^1 h(s+t)e^{s/b}\,\mathrm{d}s = \int_{\mathbb{T}} h\,\mathrm{d}(\delta_{t+\mathbb{Z}} - \eta_{b,t}) \qquad \forall t \in \mathbb{R}\,,$$

and consequently

$$\left|\hat{h}(t)\right| \leq d_{\mathbb{T}}(\delta_{t+\mathbb{Z}}, \eta_{b,t}) = d_{\mathbb{T}}(\delta_{0+\mathbb{Z}}, \eta_b) = b\tanh\frac{1}{4b} \qquad \forall t \in \mathbb{R}\,.$$

Utilize $\widehat{h}$ to re-write two terms in (3.0.3) as follows: On the one hand, by (3.0.2)

$$-\frac{f(N)}{b} = -\frac{f(N_\varepsilon)}{b} + \log\frac{N_\varepsilon}{N} - \frac{1}{b}\int_{N_\varepsilon}^{N}\frac{\Delta(s)}{s}\,\mathrm{d}s \qquad \forall N \geq N_\varepsilon\,,$$

and hence

$$Ne^{-f(N)/b}\int_{-\infty}^{f(N_\varepsilon)} h'(s)e^{s/b}\,\mathrm{d}s = N_\varepsilon\, e^{-1/b\int_{N_\varepsilon}^{N}\Delta(s)/s\,\mathrm{d}s}\,\widehat{h}\circ f(N_\varepsilon) \qquad \forall N \geq N_\varepsilon\,.$$

On the other hand, observe that

$$g_N(s) = -\int_{s}^{f(N)} e^{(s-r)/b}\frac{g(r)\Delta\circ g(r)}{b^2 + b\Delta\circ g(r)}\,\mathrm{d}r \qquad \forall N \geq N_\varepsilon, s \in [f(N_\varepsilon), f(N)]\,,$$

and consequently, for every $N \geq N_\varepsilon$,

$$\begin{aligned}
\int_{f(N_\varepsilon)}^{f(N)} h'(s)g_N(s)\,\mathrm{d}s &= \int_{f(N_\varepsilon)}^{f(N)} h'(s)\left(-\int_{s}^{f(N)} e^{(s-r)/b}\frac{g(r)\Delta\circ g(r)}{b^2+b\Delta\circ g(r)}\,\mathrm{d}r\right)\mathrm{d}s\\
&= -\frac{1}{b}\int_{N_\varepsilon}^{N}\Delta(r)\left(\int_{f(N_\varepsilon)}^{f(r)} h'(s)e^{(s-f(r))/b}\,\mathrm{d}s\right)\mathrm{d}r\\
&= -\frac{1}{b}\int_{N_\varepsilon}^{N}\Delta(r)\,\widehat{h}\circ f(r)\,\mathrm{d}r + \frac{1}{b}\,\widehat{h}\circ f(N_\varepsilon)\int_{N_\varepsilon}^{N}\Delta(r)e^{(f(N_\varepsilon)-f(r))/b}\,\mathrm{d}r\\
&= -\frac{1}{b}\int_{N_\varepsilon}^{N}\Delta(r)\,\widehat{h}\circ f(r)\,\mathrm{d}r + \frac{N_\varepsilon}{b}\,\widehat{h}\circ f(N_\varepsilon)\int_{N_\varepsilon}^{N}\frac{\Delta(r)}{r}e^{-1/b\int_{N_\varepsilon}^{r}\Delta(s)/s\,\mathrm{d}s}\,\mathrm{d}r\\
&= -\frac{1}{b}\int_{N_\varepsilon}^{N}\Delta(r)\,\widehat{h}\circ f(r)\,\mathrm{d}r + N_\varepsilon\,\widehat{h}\circ f(N_\varepsilon)\left(1 - e^{-1/b\int_{N_\varepsilon}^{N}\Delta(s)/s\,\mathrm{d}s}\right),
\end{aligned}$$

where the fourth equality is due to (3.0.2). For every $N \geq N_\varepsilon$, therefore, (3.0.3) simplifies to

$$\begin{aligned}
N\int_{\mathbb{T}} h\,\mathrm{d}(\mu_N - \omega_N) = c_N + N_\varepsilon\,\widehat{h}\circ f(N_\varepsilon) - \frac{1}{b}\int_{N_\varepsilon}^{N}\Delta(r)\,\widehat{h}\circ f(r)\,\mathrm{d}r +\\
+\int_{f(N_\varepsilon)}^{f(N)} h'(s)\left(\langle g(s)\rangle - \frac{1}{2}\right)\mathrm{d}s\,.
\end{aligned} \tag{3.0.4}$$

Fixing any $N_f \in \mathbb{N}$ with $N_f \geq \max\{N_\varepsilon, e^{2N_\varepsilon/\varepsilon}\}$, deduce from (3.0.4) that for all $N \geq N_f$,

$$\begin{aligned}
N\int_{\mathbb{T}} h\,\mathrm{d}(\mu_N - \omega_N) &\leq N_\varepsilon + \left(\tanh\frac{1}{4b}\right)\int_{N_\varepsilon}^{N}|\Delta(r)|\,\mathrm{d}r + \frac{f(N)-f(N_\varepsilon)}{2}\\
&\leq \frac{b+2\varepsilon}{2}\,\log N + \left(\tanh\frac{1}{4b}\right)\int_{N_\varepsilon}^{N}|\Delta(r)|\,\mathrm{d}r \leq b\log N + \int_{1}^{N}|\Delta(r)|\,\mathrm{d}r\,.
\end{aligned}$$

Since $h \in \mathrm{Lip}_1(\mathbb{T})$ has been arbitrary, and N_f is independent of h,

$$ND_f(N) = Nd_{\mathbb{T}}(\mu_N, \omega_N) \le b \log N + \int_1^N |\Delta(r)|\,\mathrm{d}r \qquad \forall N \ge N_f\,.$$

Clearly, this proves (3.0.1) whenever $f \in \mathcal{F} \setminus \mathcal{F}_0$.

Assume in turn that $f \in \mathcal{F}_0$. Euler summation now yields, for every $N \in \mathbb{N}$ and $h \in \mathrm{Lip}_1(\mathbb{T})$,

$$\begin{aligned} N \int_{\mathbb{T}} h\,\mathrm{d}(\mu_N - \omega_N) &= \sum\nolimits_{n=1}^{N} h \circ f(n) - Nh \circ f(N) \\ &= \frac{h \circ f(1) - h \circ f(N)}{2} - \int_1^N \Delta(r) \left(1 - \frac{1}{r}\right) h' \circ f(r)\,\mathrm{d}r + \\ &\quad + \int_1^N \Delta(r) \left(\frac{\langle r \rangle}{r} - \frac{1}{2r}\right) h' \circ f(r)\,\mathrm{d}r \\ &\le \frac{1}{4} + \int_1^N |\Delta(r)| \left|1 - \frac{1 + 2\langle r \rangle}{2r}\right| \mathrm{d}r \le \frac{1}{4} + \int_1^N |\Delta(r)|\,\mathrm{d}r\,. \end{aligned}$$

Since $h \in \mathrm{Lip}_1(\mathbb{T})$ has been arbitrary, this proves (3.0.1) whenever $f \in \mathcal{F}_0$.

Corollary 3.2. *Assume that $f \in \mathcal{F}$. Then $\lim_{N\to\infty} \mu_N[f(n) - f(N)] = \eta_{b_f}$. Moreover, for every increasing sequence (N_j) in $\mathbb{N}$ and every $t \in \mathbb{R}$,*

$$\lim\nolimits_{j\to\infty} \mu_{N_j}[f(n)] = \eta_{b_f,t} \quad \iff \quad \lim\nolimits_{j\to\infty} d_{\mathbb{T}}(f(N_j), t) = 0\,.$$

A few comments regarding the simple upper bound on $D_f(N)$ provided by Theorem 3.1 are in order. On the one hand, if $b_f \neq 0$ and $\int_1^\infty |\Delta_f(r)|\,\mathrm{d}r < \infty$ then (3.0.1) is too pessimistic by a factor $\sqrt{\log N}$, since in this case it can be shown that

$$\lim\nolimits_{N\to\infty} \frac{N}{\sqrt{\log N}} D_f(N) \in \left\{ \sqrt{\frac{|b_f|}{6\pi}\left(1 + \frac{2}{j}\right)} : j \in \mathbb{N} \cup \{\infty\} \right\}; \tag{3.0.5}$$

see [2, Thm. 3.1]. The appearance of the factor $\sqrt{\log N}$ in (3.0.5) is *not* a coincidence: By (a slightly extended version of) a classical theorem due to Roth,

$$\limsup\nolimits_{N\to\infty} \frac{N}{\sqrt{\log N}}\, d_{\mathbb{T}}(\mu_N[x_n], \eta_{b,x_N}) > 0 \tag{3.0.6}$$

for *every* sequence (x_n) in $\mathbb{R}$ and *every* $b \in \mathbb{R} \setminus \{0\}$; see [5, 10, 11, 15, 20]. As a strengthening of (3.0.6) in the context of Theorem 3.1, i.e., for $(x_n) = (f(n))$ with $f \in \mathcal{F}$, it seems plausible to conjecture that in fact

$$\liminf_{N\to\infty} \frac{N}{\sqrt{\log N}} D_f(N) > 0 \qquad \forall f \in \mathcal{F} \setminus \mathcal{F}_0 \,. \tag{3.0.7}$$

At the time of this writing, the authors do not know of any proof of, or counter-example to (3.0.7). If indeed correct, (3.0.7) would illustrate how $f = b \log$ realizes the fastest convergence in $\mu_N[f(x_n) - f(N)] \to \eta_{b_f}$ possible for any $f \in \mathcal{F} \setminus \mathcal{F}_0$.

On the other hand, despite its simplicity the right-hand side in (3.0.1) often yields the correct order of magnitude of $D_f(N)$ as $N \to \infty$, for instance for all examples discussed in [3, 15]. To briefly outline one mechanism causing this (perhaps unexpected) sharpness of (3.0.1), assume for simplicity that $f \in \mathcal{F}$ with $b_f > 0$, and denote by $f^{-1} : \mathbb{R} \to \mathbb{I}$ any function so that $f^{-1} \circ f(r) = r$ for all large r. With this, assume that

$$r \mapsto \frac{f^{-1}(r)\Delta_f \circ f^{-1}(r)}{b_f + \Delta_f \circ f^{-1}(r)} \text{ is eventually convex or eventually concave}\,, \tag{3.0.8}$$

i.e., convex or concave on $[c, \infty[$ with some $c \in \mathbb{R}_+$; in addition, if Δ_f is not eventually zero, assume that

$$\limsup_{r\to\infty} \frac{f^{-1}(r+1)|\Delta_f \circ f^{-1}(r+1)|}{f^{-1}(r)|\Delta_f \circ f^{-1}(r)|} < \infty \,. \tag{3.0.9}$$

Utilizing the above proof of Theorem 3.1, notably (3.0.4), as well as [3, Thm. 3.2], it is straightforward to establish the following *lower* bound on $D_f(N)$; all details are left to the interested reader.

Proposition 3.3. *Assume that $f \in \mathcal{F}$ with $b_f > 0$ satisfies* (3.0.8) *and* (3.0.9)*. Then there exist $c_f \in \mathbb{R}_+$ and $N_f \in \mathbb{N}$ so that*

$$D_f(N) \geq \max\left\{\frac{c_f}{N}, c_f|\Delta_f(N)| - b_f\frac{\log N}{N}\right\} \qquad \forall N \geq N_f \,.$$

Example 3.1. To illustrate Theorem 3.1 and Proposition 3.3, assume that $f \in \mathcal{F}$ satisfies

$$f(r) = b\log r + c\log\log r \qquad \forall r \geq e\,,$$

with some $b \in \mathbb{R}_+$, $c \in \mathbb{R} \setminus \{0\}$. Then $b_f = b$, $|\Delta_f(r)| = |c|/\log r$, and both results apply, implying that the sequence $(D_f(N)\log N)_{N\geq 2}$ is bounded below and above by positive constants. A careful elementary analysis [3, Thm. 5.1] shows that in fact

$$\lim_{N\to\infty} \log N(F_{\mu_N[f(n)-f(N)]} - F_{\eta_b}) = \frac{c}{b}\,(F_{\zeta_b} - F_{\eta_b}) \qquad \text{uniformly on } \mathbb{R}\,,$$

and hence in particular

$$0 < \lim_{N\to\infty} D_f(N)\log N = \frac{|c|}{b}\, d_{\mathbb{T}}(\zeta_b,\eta_b) = |c|\left(\frac{e^{\frac{1}{4b}}\tanh\frac{1}{4b}}{\cosh\frac{1}{4b}} - 1\right) < |c|\,.$$

In this example, therefore, (3.0.1) is sharp up to a positive multiplicative constant as $N\to\infty$; see also Figure 2 below.

Remark 3.4. (i) A situation such as in Example 3.1 is not uncommon: If $f\in\mathcal{F}$ has the property that Δ_f is *regularly varying* at infinity [6] with index >-1, and if f meets the requirements of Proposition 3.3, then there exists $c\in\mathbb{R}_+$ so that for all sufficiently large N,

$$\frac{|\Delta_f(N)|}{c} \le D_f(N) \le c|\Delta_f(N)|\,.$$

In this situation, therefore, $D_f(N)\to 0$ precisely at the rate of $|\Delta_f(N)|\to 0$.

(ii) Recall that $\lim_{|b|\to\infty}\eta_b = \lambda_{\mathbb{T}}$. In view of Theorem 3.1, one might expect that if $f:\mathbb{I}\to\mathbb{R}$ is smooth with $\lim_{r\to\infty} r|f'(r)| = \infty$, then $\mu_N[f(n)]\to\lambda_{\mathbb{T}}$ as $N\to\infty$, i.e., $\big(f(n)\big)$ is u.d. mod 1. A classical theorem due to Fejér [14, Cor. 1.2.1] guarantees that under mild assumptions on f this indeed is the case; for a simple explicit upper bound in the spirit of (3.0.1) see, e.g., [3, Prop. 5.8].

4. Proof of the main result

Given $f\in\mathcal{F}$, this section utilizes Theorem 3.1 and its proof to provide a simple estimate for the deviation of the empirical distribution $\mu_N[f(p_n)]$ from $\eta_{b_f,f(p_N)}$ for large N. As the reader will suspect already, this deviation tends to zero as $N\to\infty$. For a precise quantitative statement, denote by $\varphi:\mathbb{R}_+\to\mathbb{R}$ the smooth function given by

$$\varphi(r) = 1 + \frac{\log r - 1}{r} + \frac{\log r - 2}{r^2} \qquad \forall r\in\mathbb{R}_+\,,$$

and let $\phi:\mathbb{I}\to\mathbb{I}$ be any increasing smooth function with

$$\phi(r) = r\log r\,\varphi(\log r) \qquad \forall r\ge e^{e^2}\,.$$

Plainly, $\varphi(r)\to 1$, $r\varphi'(r)/\varphi(r)\to 0$, and $\phi(r)/r\to\infty$ as $r\to\infty$. The specific definitions of φ, ϕ are somewhat malleable; see Example 4.1 below. However, notice that as a quantitative refinement of the prime number theorem,

$$\frac{p_n}{\phi(n)} = 1 - \frac{(\log\log n)^2}{2(\log n)^3} + \mathcal{O}\left(\frac{\log\log n}{(\log n)^3}\right) \qquad \text{as } n\to\infty\,; \tag{4.0.1}$$

see [1, Sec. 6]. The following, a concise rewording of Theorem 1.1 previewed already in the Introduction, is the main result of this note. It accomplishes for the sequence

$(f(p_n))$ what Theorem 3.1 accomplished for $(f(n))$. In analogy to $D_f(N)$, for every $f \in \mathcal{F}$ and $N \in \mathbb{N}$ let

$$E_f(N) = d_{\mathbb{T}}(\mu_N[f(p_n)], \eta_{b_f, f(p_N)}) = d_{\mathbb{T}}(\mu_N[f(p_n) - f(p_N)], \eta_{b_f})\,.$$

Theorem 4.1. *Assume that $f \in \mathcal{F}$. Then there exists $N_f \in \mathbb{N}$ so that*

$$E_f(N) \leq \frac{1}{4\log N} + \frac{1}{N}\int_1^N |\Delta_f \circ \phi(r)|\,\mathrm{d}r \qquad \forall N \geq N_f\,. \tag{4.0.2}$$

Proof. Fix $f \in \mathcal{F}$. As in the proof of Theorem 3.1, write b_f, Δ_f, $\mu_N[f(p_n)]$, and $\eta_{b_f, f(p_N)}$ simply as b, Δ, μ_N, and ω_N, respectively.

Assume first that $f \in \mathcal{F} \setminus \mathcal{F}_0$; again, no generality is lost by assuming that $b > 0$. Let $\tilde{f} : \mathbb{I} \to \mathbb{R}$ be given by

$$\tilde{f}(r) = f \circ \phi(r) = f(1) + b\log\phi(r) + \int_1^{\phi(r)} \frac{\Delta(s)}{s}\,\mathrm{d}s \qquad \forall r \in \mathbb{I}\,. \tag{4.0.3}$$

Notice that $\tilde{f}$ is smooth. From

$$r\tilde{f}'(r) = b + \frac{b}{\log r} + \frac{b\varphi'(\log r)}{\varphi(\log r)} + \Delta\circ\phi(r)\left(1 + \frac{1}{\log r} + \frac{\varphi'(\log r)}{\varphi(\log r)}\right) \quad \forall r \geq e^{e^2},$$

it is clear that in fact $\tilde{f} \in \mathcal{F}$, with $b_{\tilde{f}} = b$ and

$$\Delta_{\tilde{f}}(r) = \frac{b}{\log r}\left(1 + \frac{\varphi'(\log r)\log r}{\varphi(\log r)}\right) + \Delta\circ\phi(r)\left(1 + \frac{1}{\log r} + \frac{\varphi'(\log r)}{\varphi(\log r)}\right) \quad \forall r \geq e^{e^2}. \tag{4.0.4}$$

Now, fix any ε with

$$0 < \varepsilon < \min\left\{1, \frac{1}{4b}\right\}\coth\frac{1}{4b} - 1\,, \tag{4.0.5}$$

and pick $N_\varepsilon \in \mathbb{N}$ with $N_\varepsilon \geq \max\{e^{e^2}, e^{1+1/\varepsilon}\}$ so large that

$$\max\left\{|\Delta(r)|, |\Delta_{\tilde{f}}(r)|, \left|\frac{\varphi'(\log r)\log r}{\varphi(\log r)}\right|\right\} < \varepsilon \qquad \forall r \geq N_\varepsilon\,,$$

but also

$$\phi(n) \geq n \quad \text{and} \quad \left|\log\frac{p_n}{\phi(n)}\right| \leq \frac{(\log\log n)^2}{(\log n)^3} \qquad \forall n \geq N_\varepsilon\,.$$

With this,

$$|\tilde{f}(n) - f(p_n)| = \left| b \log \frac{\phi(n)}{p_n} + \int_{p_n}^{\phi(n)} \frac{\Delta(s)}{s}\,\mathrm{d}s \right| \leq (b+\varepsilon)\frac{(\log\log n)^2}{(\log n)^3} \quad \forall N \geq N_\varepsilon\,. \tag{4.0.6}$$

For convenience, let $\tilde{\mu}_N = \mu_N[\tilde{f}(n)]$ and $\widetilde{\omega}_N = \eta_{b,\tilde{f}(N)}$ for every $N \in \mathbb{N}$. By the triangle inequality,

$$E_f(N) = d_{\mathbb{T}}(\mu_N, \omega_N) \leq d_{\mathbb{T}}(\mu_N, \tilde{\mu}_N) + d_{\mathbb{T}}(\tilde{\mu}_N, \widetilde{\omega}_N) + d_{\mathbb{T}}(\widetilde{\omega}_N, \omega_N) \qquad \forall N \in \mathbb{N}\,. \tag{4.0.7}$$

An upper bound is now going to be established for each of the three terms on the right in (4.0.7) individually. To this end, observe first that for every $h \in \mathrm{Lip}_1(\mathbb{T})$ and every $N \geq N_\varepsilon$,

$$\begin{aligned} N\int_{\mathbb{T}} h\,\mathrm{d}(\mu_N - \tilde{\mu}_N) &= \sum\nolimits_{n=1}^{N} (h\circ f(p_n) - h\circ\tilde{f}(n)) \\ &\leq \frac{N_\varepsilon - 1}{2} + \sum\nolimits_{n=N_\varepsilon}^{N} |f(p_n) - \tilde{f}(n)| \\ &\leq \frac{N_\varepsilon - 1}{2} + (b+\varepsilon)\sum\nolimits_{n=N_\varepsilon}^{N} \frac{(\log\log n)^2}{(\log n)^3}\,. \end{aligned}$$

Since $\lim_{N\to\infty} \frac{(\log N)^3}{N(\log\log N)^2} \sum_{n=N_\varepsilon}^{N} \frac{(\log\log n)^2}{(\log n)^3} = 1$, there exists $N_{f,0} \geq N_\varepsilon$ so that

$$d_{\mathbb{T}}(\mu_N, \tilde{\mu}_N) \leq (b+2\varepsilon)\frac{(\log\log N)^2}{(\log N)^3} \qquad \forall N \geq N_{f,0}\,.$$

Next, recall from the proof of Theorem 3.1 that with the appropriate $\widetilde{N} \geq N_\varepsilon$,

$$d_{\mathbb{T}}(\tilde{\mu}_N, \widetilde{\omega}_N) \leq \frac{b+2\varepsilon}{2}\cdot\frac{\log N}{N} + \left(\tanh\frac{1}{4b}\right)\frac{1}{N}\int_{N_\varepsilon}^{N} |\Delta_{\tilde{f}}(r)|\,\mathrm{d}r \qquad \forall N \geq \widetilde{N}\,,$$

and consequently, utilizing (4.0.4), for every $N \geq \widetilde{N}$,

$$\begin{aligned} &d_{\mathbb{T}}(\tilde{\mu}_N, \widetilde{\omega}_N) \leq \\ &\leq \frac{b+2\varepsilon}{2}\cdot\frac{\log N}{N} + \left(\tanh\frac{1}{4b}\right)\frac{1}{N}\int_{N_\varepsilon}^{N}\left(\frac{b}{\log r}(1+\varepsilon) + |\Delta\circ\phi(r)|\left(1+\frac{1+\varepsilon}{\log r}\right)\right)\mathrm{d}r \\ &\leq \frac{b+2\varepsilon}{2}\cdot\frac{\log N}{N} + \left(\tanh\frac{1}{4b}\right)\frac{1+\varepsilon}{N}\left(b\int_{N_\varepsilon}^{N}\frac{\mathrm{d}r}{\log r} + \int_{N_\varepsilon}^{N}|\Delta\circ\phi(r)|\mathrm{d}r\right)\,. \end{aligned}$$

Finally, observe that by (2.0.2),

$$d_{\mathbb{T}}(\tilde{\omega}_N, \omega_N) = d_{\mathbb{T}}(\eta_{b,\tilde{f}(N)}, \eta_{b,f(p_N)}) \leq \left(\tanh\frac{1}{4b}\right) d_{\mathbb{T}}(\tilde{f}(N), f(p_N))$$
$$\leq (b+\varepsilon)\left(\tanh\frac{1}{4b}\right)\frac{(\log\log N)^2}{(\log N)^3} \leq \left(\frac{1}{4}+\varepsilon\right)\frac{(\log\log N)^2}{(\log N)^3} \qquad \forall N \geq N_\varepsilon\,.$$

Plugging these upper bounds into (4.0.7) yields, for every $N \geq \max\{N_{f,0}, \widetilde{N}\}$,

$$E_f(N) \leq \left(b+\frac{1}{4}+3\varepsilon\right)\frac{(\log\log N)^2}{(\log N)^3} + \left(\frac{b}{2}+\varepsilon\right)\frac{\log N}{N} +$$
$$+ b\left(\tanh\frac{1}{4b}\right)\frac{1+\varepsilon}{N}\int_{N_\varepsilon}^{N}\frac{\mathrm{d}r}{\log r} + \left(\tanh\frac{1}{4b}\right)\frac{1+\varepsilon}{N}\int_{N_\varepsilon}^{N}|\Delta\circ\phi(r)|\,\mathrm{d}r\,.$$

From (4.0.5) and $\lim_{N\to\infty}\dfrac{\log N}{N}\displaystyle\int_{N_\varepsilon}^{N}\frac{\mathrm{d}r}{\log r} = 1$, it is readily deduced that taking $N_f \geq \max\{N_{f,0}, \widetilde{N}\}$ sufficiently large yields (4.0.2).

Assume in turn that $f \in \mathcal{F}_0$. Then $\omega_N = \delta_{f(p_N)+\mathbb{Z}}$ as well as $\tilde{f} = f\circ\phi \in \mathcal{F}_0$, and (4.0.4) remains valid (with $b=0$). Fix any $0<\varepsilon<\frac{1}{6}$, and pick $N_\varepsilon \in \mathbb{N}$ satisfying exactly the same requirements as above. By identical reasoning, (4.0.6) remains valid also, and hence with $\tilde{\mu}_N$ as above and the appropriate $N_{f,1} \geq N_\varepsilon$,

$$d_{\mathbb{T}}(\mu_N, \tilde{\mu}_N) \leq 2\varepsilon\frac{(\log\log N)^2}{(\log N)^3} \qquad \forall N \geq N_{f,1}\,.$$

Note that now $\tilde{\mu}_N = \delta_{\tilde{f}(N)+\mathbb{Z}}$, and the proof of Theorem 3.1 shows that for all $N \geq N_\varepsilon$,

$$d_{\mathbb{T}}(\tilde{\mu}_N, \tilde{\omega}_N) \leq \frac{1}{4N} + \frac{1}{N}\int_1^N |\Delta_{\tilde{f}}(r)|\,\mathrm{d}r$$
$$\leq \frac{1}{N}\left(\frac{1}{4}+\int_1^{N_\varepsilon}|\Delta_{\tilde{f}}(r)|\,\mathrm{d}r\right) + \frac{1}{N}\int_{N_\varepsilon}^{N}|\Delta\circ\phi(r)|\left(1+\frac{1+\varepsilon}{\log r}\right)\mathrm{d}r$$
$$\leq \frac{1}{N}\left(\frac{1}{4}+\int_1^{N_\varepsilon}|\Delta_{\tilde{f}}(r)|\,\mathrm{d}r\right) + \frac{\varepsilon(1+\varepsilon)}{N}\int_{N_\varepsilon}^{N}\frac{\mathrm{d}r}{\log r} + \frac{1}{N}\int_{N_\varepsilon}^{N}|\Delta\circ\phi(r)|\,\mathrm{d}r\,.$$

Since $0<\varepsilon(1+\varepsilon)<\frac{1}{5}$ it is clear that there exists $N_{f,2} \geq N_\varepsilon$ so that

$$d_{\mathbb{T}}(\tilde{\mu}_N, \tilde{\omega}_N) \leq \frac{1}{5\log N} + \frac{1}{N}\int_{N_\varepsilon}^{N}|\Delta\circ\phi(r)|\,\mathrm{d}r \qquad \forall N \geq N_{f,2}\,.$$

Finally, observe that

$$d_{\mathbb{T}}(\tilde{\omega}_N, \omega_N) = d_{\mathbb{T}}(\tilde{f}(N), f(p_N)) \le \varepsilon \frac{(\log\log N)^2}{(\log N)^3} \qquad \forall N \ge N_\varepsilon\,.$$

In total, for all $N \ge \max\{N_{f,1}, N_{f,2}\}$,

$$E_f(N) = d_{\mathbb{T}}(\mu_N, \omega_N) \le 3\varepsilon \frac{(\log\log N)^2}{(\log N)^3} + \frac{1}{5\log N} + \frac{1}{N}\int_{N_\varepsilon}^{N} |\Delta \circ \phi(r)|\,\mathrm{d}r\,,$$

from which it is clear that taking $N_f \ge \max\{N_{f,1}, N_{f,2}\}$ sufficiently large again yields (4.0.2).

Corollary 4.2. *Assume that $f \in \mathcal{F}$. Then $\lim_{N\to\infty} \mu_N[f(p_n) - f(p_N)] = \eta_{b_f}$. Moreover, for every increasing sequence (N_j) in $\mathbb{N}$ and every $t \in \mathbb{R}$,*

$$\lim_{j\to\infty} \mu_{N_j}[f(p_n)] = \eta_{b_f,t} \quad \Longleftrightarrow \quad \lim_{j\to\infty} d_{\mathbb{T}}(f(p_{N_j}), t) = 0\,.$$

Notice that by Corollary 4.2,

$$\mathcal{M}[f(p_n)] \subset \{\eta_{b_f,s} : s \in [0,1[\} \qquad \forall f \in \mathcal{F}\,. \tag{4.0.8}$$

If $f \in \mathcal{F} \setminus \mathcal{F}_0$ then the inclusion in (4.0.8) actually is an equality. By contrast, if $f \in \mathcal{F}_0$ then the inclusion may be strict, e.g., for $f(r) = (1+r)^{-1}$ where $\mathcal{M}[f(p_n)] = \{\delta_{0+\mathbb{Z}}\}$, but may also be an equality, e.g., for $f(r) = \log\log(1+r)$.

Just as in Section 3, a few comments regarding the sharpness of (4.0.2) are in order. On the one hand, the appearance of the term $1/\log N$ in (4.0.2) generally suggests a much slower convergence in $E_f(N) \to 0$ than in $D_f(N) \to 0$. For instance, if $f \in \mathcal{F}$ with $b_f > 0$ has the property that either Δ_f is regularly varying at infinity with index < 0 or else $\int_1^\infty |\Delta_f(r)|\,\mathrm{d}r < \infty$, and if $\tilde{f}$ given by (4.0.3) satisfies the assumptions of Proposition 3.3 then $\inf_{N\ge 2} E_f(N)\log N > 0$ whereas $\lim_{N\to\infty} D_f(N)\log N = 0$. The simplest concrete example is $f = \log$ for which (4.0.2) yields $\limsup_{N\to\infty} E_f(N)\log N \le \frac{1}{4}$. A careful analysis reveals that in fact

$$\lim_{N\to\infty} \log N (F_{\mu_N[\log p_n - \log p_N]} - F_{\eta_1}) = F_{\zeta_1} - F_{\eta_1} \qquad \text{uniformly on } \mathbb{R}\,;$$

see [4, Lem. 3.3] and also Figure 1. In this example, therefore, (4.0.2) is asymptotically sharp, as $\lim_{N\to\infty} E_f(N)\log N = d_{\mathbb{T}}(\zeta_1, \eta_1) \approx 3.076 \cdot 10^{-2}$.

On the other hand, the convergence in $E_f(N) \to 0$ may well be faster than in $D_f(N) \to 0$, and (4.0.2) may be too pessimistic. For instance, if

$$f(r) = \log r - \log\log r \qquad \forall r > e\,,$$

then (4.0.2) yields $\sup_{N\in\mathbb{N}} E_f(N)\log N < \infty$. However, notice that $b_f = 1$, $\Delta_f(r) = -1/\log r$, and deduce from (4.0.4) that

$$\begin{aligned}\Delta_{\tilde f}(r) &= \frac{1}{\log r}\left(1+\frac{\varphi'(\log r)\log r}{\varphi(\log r)}\right) + \Delta_f\circ\phi(r)\left(1+\frac{1}{\log r}+\frac{\varphi'(\log r)}{\varphi(\log r)}\right)\\ &= \frac{1}{(\log r)^2} + \mathcal{O}\left(\frac{(\log\log r)^2}{(\log r)^3}\right) \qquad \text{as } r\to\infty\,.\end{aligned}$$

The proof of Theorem 4.1 now shows that $\sup_{N\in\mathbb{N}} E_f(N)(\log N)^2 < \infty$ and also, via Proposition 3.3 applied with $\tilde f$ instead of f, that $\inf_{N\ge 2} E_f(N)(\log N)^2 > 0$. Clearly, therefore, $\lim_{N\to\infty} E_f(N)\log N = 0$, yet $\lim_{N\to\infty} D_f(N)\log N \in \mathbb{R}_+$ by Example 3.1. A careful analysis [15] reveals that in fact

$$\lim_{N\to\infty}(\log N)^2(F_{\mu_N[f(p_n)-f(p_N)]} - F_{\eta_1}) = F_{\zeta_1} - F_{\eta_1} \qquad \text{uniformly on } \mathbb{R}\,.$$

(The intrepid reader may want to try reaching this conclusion via [19, Thm. 1].) In this example, (4.0.2) is too pessimistic by a factor $\log N$; see also Figure 2.

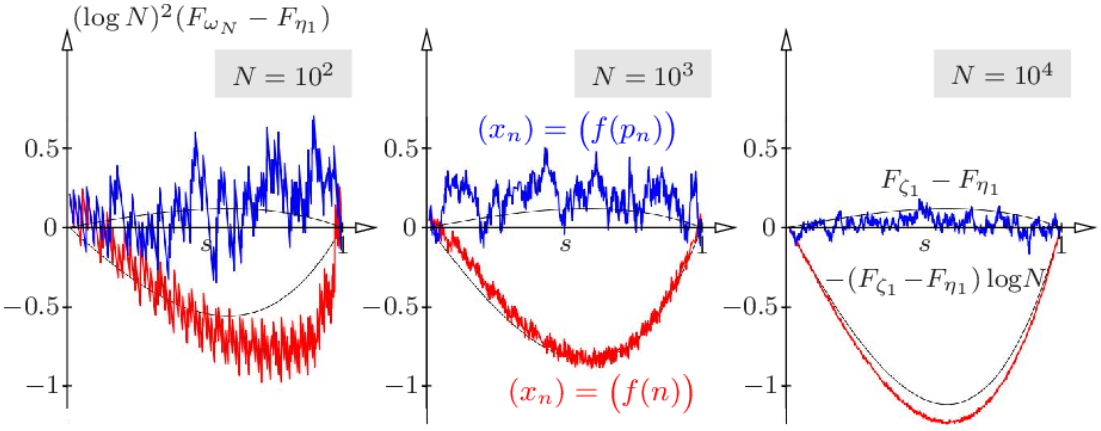


Figure 2: For $f(r) = \log r - \log\log r$ the rotated empirical distributions $\omega_N := \mu_N[f(x_n) - f(x_N)]$ converge to η_1 along $\pm(\zeta_1 - \eta_1)$ as $N\to\infty$, both for $(x_n) = (p_n)$ at rate $(\log N)^{-2}$ (blue data), and for $(x_n) = (n)$ at rate $(\log N)^{-1}$ (red data; see also Example 3.1).

Example 4.1. Given any $q \in \mathbb{N}$, there exists $f \in \mathcal{F}$ with $b_f = 1$ so that

$$E_f(N) = \mathcal{O}\big((\log N)^{-q}\big) \qquad \text{as } N\to\infty\,. \tag{4.0.9}$$

To explicitly construct such an f, recall, e.g., from [1] that as a refinement of the prime number theorem,

$$\frac{p_n}{n\log n} = 1 + \frac{\log\log n - 1}{\log n} + \sum\nolimits_{j=1}^{q-1}\frac{(-1)^{j+1}}{j}\cdot\frac{Q_j(\log\log n)}{(\log n)^{j+1}} + \mathcal{O}\left(\frac{(\log\log n)^q}{(\log n)^{q+1}}\right) \tag{4.0.10}$$

as $n \to \infty$, where each Q_j is a monic polynomial of degree j with rational coefficients; for instance, $Q_1(x) = x-2$, $Q_2(x) = x^2-6x+11$, and $Q_3(x) = x^3 - \frac{21}{2}x^2 + 42x + \frac{131}{2}$. Inspired by (4.0.10), denote by $\varphi_q : \mathbb{R}_+ \to \mathbb{R}$ the smooth function given by

$$\varphi_q(r) = 1 + \frac{\log r - 1}{r} + \sum\nolimits_{j=1}^{q-1} \frac{(-1)^{j+1}}{j} \cdot \frac{Q_j(\log r)}{r^{j+1}} \qquad \forall r \in \mathbb{R}_+ ,$$

and let $\phi_q : \mathbb{I} \to \mathbb{I}$ be such that $\phi_q(r) = r \log r \, \varphi_q(\log r)$ for all large r. (With this notation, φ, ϕ chosen earlier simply are φ_2, ϕ_2.) As a refinement of (4.0.1),

$$\frac{p_n}{\phi_q(n)} = 1 + \frac{(-1)^{q+1}}{q} \cdot \frac{(\log\log n)^q}{(\log n)^{q+1}} + \mathcal{O}\left(\frac{(\log\log n)^{q-1}}{(\log n)^{q+1}}\right) \qquad \text{as } n \to \infty .$$

As in the proof of Theorem 4.1, and with $b_f = 1$ but $f \in \mathcal{F}$ yet to be determined, let $\tilde{f} = f \circ \phi$, and notice that $\tilde{f} \in \mathcal{F}$, $b_{\tilde{f}} = 1$, and for all sufficiently large r,

$$\Delta_{\tilde{f}}(r) = \frac{1}{\log r} + \frac{\varphi_q'(\log r)}{\varphi_q(\log r)} + \Delta_f \circ \phi_q(r) \left(1 + \frac{1}{\log r} + \frac{\varphi_q'(\log r)}{\varphi_q(\log r)}\right) .$$

By the same argument as in the proof of Theorem 4.1,

$$E_f(N) \le \frac{1}{N} \int_1^N |\Delta_{\tilde{f}}(r)| \, \mathrm{d}r + \mathcal{O}\left(\frac{(\log\log N)^q}{(\log N)^{q+1}}\right) \qquad \text{as } N \to \infty .$$

Thus f has the desired property (4.0.9), provided that $\Delta_{\tilde{f}} = 0$, or equivalently

$$\Delta_f(r) = - \left. \frac{\varphi_q(\log s) + \varphi_q'(\log s) \log s}{\varphi_q(\log s) \log s + \varphi_q(\log s) + \varphi_q'(\log s) \log s} \right|_{\phi_q(s)=r} \tag{4.0.11}$$

for all large r. With Δ_f given by (4.0.11) for $r \ge N_0$, therefore, simply determine f from (3.0.2) with N_0 instead of N_ε.

Given $q \in \mathbb{N}$, Example 4.1 demonstrates that for a carefully chosen $f \in \mathcal{F}$ the convergence in $E_f(N) \to 0$ as $N \to \infty$ is no slower than that in $(\log N)^{-q} \to 0$. The authors do not know whether that rate may be significantly faster at all, for any $f \in \mathcal{F} \setminus \mathcal{F}_0$. Suspecting that it may not, they specifically conjecture that

$$\limsup\nolimits_{N\to\infty} E_f(N) N^\varepsilon = \infty \qquad \forall f \in \mathcal{F} \setminus \mathcal{F}_0, \varepsilon > 0 . \tag{4.0.12}$$

Notice that (4.0.12) is consistent with the following observation: If $f : \mathbb{I} \to \mathbb{R}$ is smooth with $f(p_n) = \tilde{f}(n)$ for all large $n \in \mathbb{N}$ and some $\tilde{f} \in \mathcal{F} \setminus \mathcal{F}_0$, then known results about prime gaps [9, 25] imply that, in analogy to (1.0.7),

$$\liminf\nolimits_{r\to\infty} r f'(r) = 0 , \quad \limsup\nolimits_{r\to\infty} r f'(r) = \infty ,$$

so clearly $f \notin \mathcal{F}$. Taking, for instance, $\tilde{f}(r) = b \log r + c r^{-\varepsilon}$ with $b, c \in \mathbb{R} \setminus \{0\}$ and $0 < \varepsilon < 1$ suggests the validity of (4.0.12). If indeed correct, (4.0.12) would illustrate how the convergence $\mu_n[x_n - x_N] \to \eta_{b_f}$ as $N \to \infty$ often is much slower for

$(x_n) = (f(p_n))$ than for $(x_n) = (f(n))$, and hence would constitute a new quantitative manifestation of the irregularity inherent to the prime numbers.

Acknowledgement

The first author was partially supported by an Nserc Discovery Grant. Both authors are grateful to an anonymous referee for pertinent comments that helped improve the exposition.

REFERENCES

[1] J. Arias de Reyna and J. Toulisse, The n-th prime asymptotically, *J. Théor. Nombres Bordeaux* **25** (2013), 521–555.

[2] A. Berger, A note on the distribution of $(\log n)$ mod 1, *Unif. Distrib. Theory* **17** (2022), 77–100.

[3] A. Berger and A. Rahmatidehkordi, Circling the uniform distribution, *J. Math. Anal. Appl.* **527** (2023), No. 127495, 33pp.

[4] A. Berger and A. Rahmatidehkordi, The primes perform a Benford dance, *Anal. Appl.* **23** (2025), 553–576.

[5] D. Bilyk, On Roth's orthogonal function method in discrepancy theory, *Unif. Distrib. Theory* **6** (2011), 143–184.

[6] N.H. Bingham, C.M. Goldie and J.L. Teugels, *Regular variation*, Cambridge University Press, 1987.

[7] L. Brown and S. Steinerberger, On the Wasserstein distance between classical sequences and the Lebesgue measure, *Trans. Amer. Math. Soc.* **373** (2020), 8943–8962.

[8] C.A. Cabrelli and U.M. Molter, The Kantorovich metric for probability measures on the circle, *J. Comput. Appl. Math.* **57** (1995), 345–361.

[9] K. Ford, B. Green, S. Konyagin, J. Maynard and T. Tao, Long gaps between primes, *J. Amer. Math. Soc.* **31** (2018), 65–105.

[10] C. Graham, Irregularity of distribution in Wasserstein distance, *J. Fourier Anal. Appl.* **26** (2020), No. 75, 21pp.

[11] G. Halász, On Roth's method in the theory of irregularities of point distributions, pp. 79–94 in: *Recent progress in analytic number theory* (Durham, 1979), Academic Press, 1981.

[12] H.L. Hutson, On the distribution of $\log(p)$, *Int. J. Pure Appl. Math.* **7** (2003), 499–508.

[13] J.H.B. Kemperman, Distributions modulo 1 of slowly changing sequences, *Nieuw Arch. Wisk.* **21** (1973), 138–163.

[14] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, 1974.

[15] M.J. Liu, *Quantitative Aspects of Point Distributions modulo one*, MSc thesis, University of Alberta, 2026.

[16] H. Niederreiter, Distribution mod 1 of monotone sequences, *Nederl. Akad. Wetensch. Indag. Math.* **46** (1984), 315–327.

[17] Y. Ohkubo, On sequences involving primes, *Unif. Distrib. Theory* **6** (2011), 221–238.

[18] Y. Ohkubo and O. Strauch, Distribution of leading digits of numbers, *Unif. Distrib. Theory* **11** (2016), 23–45.

[19] Y. Ohkubo and O. Strauch, Distribution of leading digits of numbers II, *Unif. Distrib. Theory* **14** (2019), 19–42.

[20] S. Steinerberger, Wasserstein distance, Fourier series and applications, *Monatsh. Math.* **194** (2021), 305–338.

[21] O. Strauch, *Distribution of sequences: a theory*, Veda a Academia Nakladatelství, 2019.

[22] O. Strauch and O. Blažeková, Distribution of the sequence p_n/n mod 1, *Unif. Distrib. Theory* **1** (2006), 45–63.

[23] A. Wintner, On the cyclical distribution of the logarithms of the prime numbers, *Quart. J. Math. Oxford Ser.* **6** (1935), 65–68.

[24] C. Xu, The distributional asymptotics mod 1 of $(\log_b n)$, *Unif. Distrib. Theory* **14** (2019), 105–122.

[25] Y. Zhang, Bounded gaps between primes, *Ann. Math* **179** (2014), 1121–1174.

Received April 20, 2026
Accepted September 26, 2026
Published October 30, 2026

Arno Berger
Department of Mathematical and Statistical Sciences,
University of Alberta,
Edmonton T6G 2N8, Canada.
berger@ualberta.ca

Matthew Jamie Liu
Department of Mathematical and Statistical Sciences,
University of Alberta,
Edmonton T6G 2N8, Canada.
mjliu1@ualberta.ca