THE LIMIT DISTRIBUTION OF THE MIDDLE PRIME FACTORS OF AN INTEGER

Jean-Marie De Koninck[1]
Département de mathématiques et de statistique, Université Laval, Québec G1V 0A6, Canada
jmdk@mat.ulaval.ca

Nicolas Doyon[2]
Département de mathématiques et de statistique, Université Laval, Québec G1V 0A6, Canada
nicolas.doyon@mat.ulaval.ca

Vincent Ouellet
Département de mathématiques et de statistique, Université Laval, Québec G1V 0A6, Canada
vincent.ouellet.7@ulaval.ca

Received: , Revised: , Accepted: , Published:

Abstract

Writing an integer $n \geq 2$ as $n = p_1^{\alpha_1} p_2^{\alpha} \cdots p_k^{\alpha_k}$ where $p_1 < p_2 < \cdots < p_k$ are its prime factors, for any real $\beta \in (0,1)$, we define the *β-positioned prime factor* of $n > 1$ as $p^{(\beta)}(n) := p_{\max(1,\lfloor \beta(k+1) \rfloor)}$. We obtain the limit distribution of $p^{(\beta)}(n)$.

1. Introduction

Writing an integer $n \geq 2$ as $n = p_1^{\alpha_1} p_2^{\alpha} \cdots p_k^{\alpha_k}$ where $p_1 < p_2 < \cdots < p_k$ are its prime factors, for any real $\beta \in (0,1)$, we define the *β-positioned prime factor* of n as $p^{(\beta)}(n) := p_{\max(1,\lfloor \beta(k+1) \rfloor)}$. For convenience, we set $p^{(\beta)}(1) = 1$. Recently, Ouellet [9] improved a result of De Koninck and Luca [2] by showing that $p^{(1/2)}(n)$, the *middle prime factor* of n, satisfies the relation

$$\sum_{n \leq x} \frac{1}{p^{(1/2)}(n)} = \frac{x}{\log x} \exp\left(\sqrt{2 \log_2 x \log_3 x}\left(H(x) + O\left(\frac{1}{(\log_3 x)^2}\right)\right)\right),$$

---

[1]Research supported in part by a grant from NSREC.
[2]Research supported in part by a grant from NSREC.

where $\log_k x$ stands for the k-iterated logarithm of x assuming that x is large enough for $\log_k x$ to be well defined and positive, and where

$$H(x) = 1 - \frac{3\log_4 x}{2\log_3 x} + \left(\frac{3}{2}\log 2 - 1\right)\frac{1}{\log_3 x} - \frac{9}{8}\left(\frac{\log_4 x}{\log_3 x}\right)^2 + \left(\frac{9\log 2}{4} + 1\right)\frac{\log_4 x}{(\log_3 x)^2}$$

and more generally that

$$\sum_{n\le x}\frac{1}{p^{(\beta)}(n)} = \frac{x}{\log x}\exp\left(C(\log_2 x)^{1-\beta}(\log_3 x)^{\beta}\left(G(x,\beta) + O\left(\frac{1}{(\log_3 x)^2}\right)\right)\right),$$

where $C = \frac{(1-\beta)^{2\beta-1}}{\beta^\beta}$ and

$$G(x,\beta) = 1 + c_1\frac{\log_4 x}{\log_3 x} + c_2\frac{1}{\log_3 x} + c_3\left(\frac{\log_4 x}{\log_3 x}\right)^2 + c_4\frac{\log_4 x}{(\log_3 x)^2}$$

with

$$c_1 = \frac{-\beta(2-\beta)}{1-\beta}, \qquad c_2 = \beta\left(\log\beta - \frac{3-2\beta}{1-\beta}\log(1-\beta) - \frac{1}{1-\beta}\right),$$
$$c_3 = \frac{2-\beta}{2}c_1, \qquad c_4 = (2-\beta)c_2 - c_1 + \frac{\beta}{1-\beta}.$$

Since the main contribution to the sum of the reciprocals of $p^{(\beta)}(n)$ comes from a set of integers of zero density, these results and the investigation of their proofs do not reveal anything concerning the normal value of $p^{(\beta)}(n)$, nor for the distribution of the values of $p^{(\beta)}(n)$. On the other hand, De Koninck and Kátai [1], using a Turán-Kubilius type inequality, showed that the normal order of $\log_2 p^{(1/2)}(n)$ is $\frac{1}{2}\log_2 n$. Explicitly, they proved that for any function $g(x)$ tending to infinity with x,

$$\lim_{x\to\infty}\frac{1}{x}\#\left\{n\le x : \left|\log_2 p^{(1/2)}(n) - \frac{1}{2}\log_2 x\right| > g(x)\sqrt{\log_2 x}\right\} = 0.$$

Tenenbaum, in his book [11], provides an estimate for $p_j(n)$, the j-th prime factor of n, namely

$$p_j(n) = e^{e^{j+O(\sqrt{j})}} \qquad \text{almost everywhere.}$$

Therefore, since the normal order of $\omega(n) := \sum_{p|n} 1$ is $\log_2 n$ and since $\log_2 n = \log_2 x + O(1)$ for almost all $n \le x$, one can expect that for any given $\varepsilon > 0$,

$$\lim_{x\to\infty}\frac{1}{x}\#\left\{n\le x : \left|\log_2 p^{(\beta)}(n) - \beta\log_2 x\right| > \varepsilon\log_2 x\right\} = 0.$$

Here, we show the following stronger result.

Theorem 1. *For any given real number t such that $|t| \ll (\log_2 x)^\varepsilon$ for some fixed $0 < \varepsilon < \frac{1}{8}$, we have*

$$\frac{1}{x}\#\left\{n \le x : \frac{\log_2 p^{(1/2)}(n) - \frac{1}{2}\log_2 x}{\sqrt{\log_2 x}} < t\right\} = \Phi(2t) + O\left(\frac{1}{\sqrt{\log_3 x}}\right),$$

where $\Phi(\tau) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\tau} e^{-v^2/2}\,\mathrm{d}v$ *stands for the normal distribution function.*

We also provide a generalization valid for any $\beta \in (0,1)$.

Theorem 2. *For $\beta \in (0,1)$ and any real number t such that $|t| \ll (\log_2 x)^\varepsilon$ for some fixed $0 < \varepsilon < \frac{1}{8}$, we have*

$$\frac{1}{x}\#\left\{n \le x : \frac{\log_2 p^{(\beta)}(n) - \beta\log_2 x}{\sqrt{\log_2 x}} < t\right\} = \Phi\left(\frac{t}{\sqrt{\beta(1-\beta)}}\right) + O_\beta\left(\frac{1}{\sqrt{\log_3 x}}\right).$$

Remark. Doyon and Ouellet [4] have shown that the sum of the reciprocals of the middle prime factors behaves very differently depending on whether the prime multiplicity is considered or not. However, Theorems 1 and 2 both hold whether prime multiplicity is taken into account or not. This follows from the fact that if we set $\Omega(n) := \sum_{p^a \| n} a$, then, for any function $\xi(x)$ tending to infinity as x tends to infinity, we have

$$\frac{1}{x}\#\{n \le x : |\Omega(n) - \omega(n)| > \xi(x)\} = o(1) \qquad (x \to \infty).$$

A more explicit bound is used below to prove it.

2. Preliminary results

Our proofs of Theorems 1 and 2 will make much use of the following completely additive function $\Omega_y^{(\beta)}(n)$ (here y is any given positive real number) defined on primes p by

$$\Omega_y^{(\beta)}(p) := \begin{cases} -1 & \text{if } p \le y, \\ \frac{\beta}{1-\beta} & \text{if } p > y. \end{cases}$$

When the superscript β is omitted, we set $\Omega_y(n) := \sum_{\substack{p \le y \\ p^a \| n}} a$. We also define the strongly additive function $\omega_y^{(\beta)}(n)$ on primes p by $\omega_y^{(\beta)}(p) := \Omega_y^{(\beta)}(p)$ and, when the superscript β is omitted, we set $\omega_y(n) = \sum_{\substack{p \le y \\ p | n}} 1$. Finally, we let $p(n)$ (resp. $P(n)$) stand for the smallest (resp. largest) prime factor of n.

Lemma 1. *Let x be a large real number. Given positive real numbers $y < x$ and $0 < \beta < 1$, we have*

$$\#\left\{n \le x : p^{(\beta)}(n) \le y\right\} = \#\left\{n \le x : \omega(n) > \frac{1-\beta}{\beta}, \omega_y^{(\beta)}(n) < 1\right\} + R(x,y),$$

where $R(x,y) := \#\left\{n \le x : \omega(n) \le \frac{1-\beta}{\beta}, p(n) \le y\right\}$. *In particular,*

$$\#\left\{n \le x : p^{(\beta)}(n) \le y\right\} = \#\left\{n \le x : \omega_y^{(\beta)}(n) < 1\right\} + O_\beta\left(\frac{x\left(\log_2 x\right)^{\frac{1-2\beta}{\beta}}}{\log x}\right).$$

Proof. It follows from the definitions of $\omega_y^{(\beta)}(n)$ and $\omega_y(n)$ that

$$\omega_y^{(\beta)}(n) = \frac{\beta}{1-\beta}\left(\omega(n) - \omega_y(n)\right) - \omega_y(n) = \frac{1}{1-\beta}\left(\beta\omega(n) - \omega_y(n)\right). \tag{1}$$

Moreover, we have for any integer $n > 1$ that $p^{(\beta)}(n) > y$ if and only if $\omega_y(n) \le k_0 - 1$, where $k_0 = \max\left(1, \lfloor \beta\left(\omega(n)+1\right)\rfloor\right)$. Thus, when $\omega(n) > \frac{1-\beta}{\beta}$, we obtain from (1) that

$$p^{(\beta)}(n) \le y \quad \text{if and only if} \quad \omega_y^{(\beta)}(n) < 1.$$

When $\omega(n) \le \frac{1-\beta}{\beta}$, we have $k_0 = 1$, so that $p^{(\beta)}(n) \le y$ if and only if $p(n) \le y$. Using the well known Hardy-Ramanujan inequality, the error term follows. □

From here on, we focus our attention on the distribution of the function $\omega_y^{(\beta)}(n)$. In their paper, De Koninck and Kátai [1] used the function $\Delta_y(n)$ which is the same as our function $\Omega_y^{(1/2)}(n)$. Observe that when $\beta = c/d$ is rational, we have

$$\Omega_y^{(\beta)}(p) = \begin{cases} -1 & \text{if } p \le y, \\ \frac{c}{d-c} & \text{if } p > y, \end{cases}$$

which would allow one to work with the integer-valued function $(c-d)\Omega_y^{(\beta)}(n)$ and thus to use tools such as the Selberg-Delange method (see for instance Delange [3]). For $k \ge 1$ and $z \ge 2$, we define the function $\mathcal{D}_k(z)$ by

$$\mathcal{D}_k(z) = \#\left\{n \ge 1 : \omega(n) \le k, P(n) \le z, \mu^2(n) = 1\right\} \ll \frac{1}{k!}\left(\frac{z}{\log z}\right)^k \tag{2}$$

uniformly for $k \le (\log_2 x)^{1/3}$, where $z = x^{1/\ell}$ for any real number $1 \le \ell \le (\log_2 x)^{1/3}$. We begin by citing a particular case of Proposition 4 of Granville and Soundararajan [7] which provides information on the moments of the $\omega_y^{(\beta)}(n)$ function.

Lemma 2. *For any real numbers $\ell \geq 1$ and $\beta \in (0,1)$, set $z := x^{1/\ell}$ and consider the functions*

$$\mu_y^{(\beta)}(z) := \sum_{p \leq z} \frac{\Omega_y^{(\beta)}(p)}{p} \qquad \text{and} \qquad \left(\sigma_y^{(\beta)}(z)\right)^2 := \sum_{p \leq z} \frac{\left(\Omega_y^{(\beta)}(p)\right)^2}{p}\left(1 - \frac{1}{p}\right).$$

Then, uniformly for all even integers $k \leq \left(\sigma_y^{(\beta)}(z)/M\right)^{2/3}$ and for all real numbers $2 \leq y < x$ and $1 \leq \ell \leq (\log_2 x)^{1/3}$, we have

$$\sum_{n \leq x} \left(\sum_{\substack{p \mid n \\ p \leq z}} \omega_y^{(\beta)}(p) - \mu_y^{(\beta)}(z) \right)^k = C_k x \left(\sigma_y^{(\beta)}(z)\right)^k + O\left(x C_k k^3 M^2 \left(\sigma_y^{(\beta)}(z)\right)^{k-2} \right) + O\left(M^k \left(\sum_{p \leq z} \frac{1}{p} \right)^k \mathcal{D}_k(z) \right),$$

where $M = \max \left|\Omega_y^{(\beta)}(p)\right| = \max\left(1, \frac{\beta}{1-\beta}\right)$. The constants C_k are given by

$$C_k := \frac{\Gamma(k+1)}{\Gamma(k/2+1)\, 2^{k/2}} \tag{3}$$

which for k even corresponds to the Gaussian moments. Moreover, uniformly for all odd integers $k \leq \left(\sigma_y^{(\beta)}(z)/M\right)^{2/3}$, for all real numbers $2 \leq y < x$ and $1 \leq \ell \leq (\log_2 x)^{1/3}$, we have

$$\sum_{n \leq x} \left(\sum_{\substack{p \mid n \\ p \leq z}} \omega_y^{(\beta)}(p) - \mu_y^{(\beta)}(z) \right)^k \ll C_k x \left(\sigma_y^{(\beta)}(z)\right)^{k-1} k^{3/2} M + M^k \left(\sum_{p \leq z} \frac{1}{p} \right)^k \mathcal{D}_k(z).$$

The following two easy results (which are consequences of Mertens' formula) will allow for an application of Lemma 2 to our problem.

Lemma 3. *For any real number $\ell \geq 1$, set $z = x^{1/\ell}$. Then, given a real number $\beta \in (0,1)$, we have uniformly for any real numbers $1 \leq \ell \leq (\log_2 x)^{1/3}$ and $e^e < y = o(z)$ that*

$$\mu_y^{(\beta)}(z) = \sum_{p \leq z} \frac{\Omega_y^{(\beta)}(p)}{p} = \frac{\beta \log_2 z - \log_2 y}{1-\beta} + O_\beta(1).$$

In particular, for any fixed $\xi \in (0,1)$*, we have uniformly for any* y *such that* $\log_2 y = (1+o(1))\,\xi \log_2 x$ *as* $x \to \infty$ *that*

$$\mu_y^{(\beta)}(z) = (1+o(1))\frac{\beta-\xi}{1-\beta}\log_2 x.$$

Proof. By Mertens' formula, we have

$$\begin{aligned}\sum_{p\le z}\frac{\Omega_y^{(\beta)}(p)}{p} &= \sum_{p\le y}\frac{-1}{p} + \sum_{y<p\le z}\frac{\beta}{p(1-\beta)}\\ &= -\log_2 y + \frac{\beta(\log_2 z - \log_2 y)}{1-\beta} + O_\beta(1) = \frac{\beta\log_2 z - \log_2 y}{1-\beta} + O_\beta(1).\end{aligned}$$

Since $\log_2 z = \log_2\left(x^{1/\ell}\right) = \log_2 x - \log \ell$, we have, when $\log_2 y = (1+o(1))\,\xi\log_2 x$, that, as $x\to\infty$,

$$\begin{aligned}\sum_{p\le z}\frac{\Omega_y^{(\beta)}(p)}{p} &= \frac{1}{1-\beta}\left(\beta\log_2 x - \beta\log\ell - (1+o(1))\,\xi\log_2 x\right) + O_\beta(1)\\ &= (1+o(1))\,\frac{\beta-\xi}{1-\beta}\log_2 x - \frac{\beta}{1-\beta}\log\ell = (1+o(1))\,\frac{\beta-\xi}{1-\beta}\log_2 x.\end{aligned}$$

□

Lemma 4. *For any real number* $\ell \ge 1$*, set* $z = x^{1/\ell}$*. Then, given a real number* $\beta \in (0,1)$*, we have uniformly for any real numbers* $1 \le \ell \le (\log_2 x)^{1/3}$ *and* $e^e < y = o(z)$ *that*

$$\left(\sigma_y^{(\beta)}(z)\right)^2 = \sum_{p\le z}\frac{\left(\Omega_y^{(\beta)}(p)\right)^2}{p}\left(1-\frac{1}{p}\right) = \frac{\beta^2\log_2 z + (1-2\beta)\log_2 y}{(1-\beta)^2} + O_\beta(1).$$

Moreover, for any fixed $\xi \in (0,1)$*, we have uniformly for any* y *such that* $\log_2 y = (1+o(1))\,\xi\log_2 x$ *as* $x\to\infty$ *that*

$$\left(\sigma_y^{(\beta)}(z)\right)^2 = (1+o(1))\,\frac{\beta^2+\xi-2\beta\xi}{(1-\beta)^2}\log_2 x.$$

Proof. On the one hand, using Mertens' formula, we have

$$\begin{aligned}\sum_{p\le z}\frac{(\Omega_y^{(\beta)}(p))^2}{p} &= \sum_{p\le y}\frac{1}{p} + \sum_{y<p\le z}\frac{\beta^2}{(1-\beta)^2 p}\\ &= \log_2 y + \frac{\beta^2(\log_2 z - \log_2 y)}{(1-\beta)^2} + O_\beta(1)\\ &= \frac{\beta^2\log_2 z + (1-2\beta)\log_2 y}{(1-\beta)^2} + O_\beta(1).\end{aligned}$$

On the other hand,

$$\sum_{p\le z}\frac{\left(\Omega_y^{(\beta)}(p)\right)^2}{p^2}\le \max\left(1,\frac{\beta^2}{(1-\beta)^2}\right)\sum_p\frac{1}{p^2}=O_\beta(1).$$

Hence,

$$\sum_{p\le z}\frac{(\Omega_y^{(\beta)}(p))^2}{p}=\frac{\beta^2\log_2 z+(1-2\beta)\log_2 y}{(1-\beta)^2}+O_\beta(1),$$

where the error term is uniform in ℓ and y. If $\log_2 y=(1+o(1))\,\xi\log_2 x$ as $x\to\infty$, then

$$\begin{aligned}\sum_{p\le z}\frac{(\Omega_y^{(\beta)}(p))^2}{p}&=(1+o(1))\,\frac{\beta^2+\xi-2\beta\xi}{(1-\beta)^2}\log_2 x-\frac{\beta^2}{(1-\beta)^2}\log\ell\\&=(1+o(1))\,\frac{\beta^2+\xi-2\beta\xi}{(1-\beta)^2}\log_2 x.\end{aligned}$$

□

From Lemmas 2, 3 and 4, we will deduce the following corollary.

Corollary 1. *Let $\beta\in(0,1)$ be a real number and set $z=x^{1/k}$ for any integer $k\ge 1$. Fix $\xi\in(0,1)$. Then, uniformly for even integers $k\le\frac{N^{1/3}}{2}(\log_2 x)^{1/3}$, where $N:=\min\left(\left(\frac{1-\beta}{\beta}\right)^2,\left(\frac{\beta}{1-\beta}\right)^2\right)$, and for y such that $\log_2 y=(1+o(1))\xi\log_2 x$ as $x\to\infty$, we have*

$$\frac{1}{x}\sum_{n\le x}\left(\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^k=C_k+O_{\beta,\xi}\left(\frac{k^{3/2}C_k}{\sqrt{\log_2 x}}\right),$$

while uniformly for odd integers $k\le\frac{N^{1/3}}{2}(\log_2 x)^{1/3}$ and y satisfying the above bounds, we have

$$\frac{1}{x}\sum_{n\le x}\left(\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^k\ll_{\beta,\xi}\frac{k^{3/2}C_k}{\sqrt{\log_2 x}}.$$

Proof. It follows from Lemma 4 with $\ell=k$ that, as $x\to\infty$,

$$\left(\frac{\sigma_y^{(\beta)}(z)}{M}\right)^{2/3}=(1+o(1))\,\frac{(\beta^2+\xi-2\beta\xi)^{1/3}}{(1-\beta)^{2/3}}\frac{(\log_2 x)^{1/3}}{M^{2/3}}\ge(1+o(1))\,N^{1/3}\,(\log_2 x)^{1/3}$$

uniformly for $1 \leq k \leq \frac{N^{1/3}}{2} (\log_2 x)^{1/3}$. Thus, we can apply Lemma 2 for any $k \leq \frac{N^{1/3}}{2} (\log_2 x)^{1/3}$. Now, we write

$$\begin{aligned} \omega_y^{(\beta)}(n) - \mu_y^{(\beta)}(x) &= \left(\sum_{\substack{p\mid n \\ p\leq z}} \omega_y^{(\beta)}(p) - \mu_y^{(\beta)}(z) \right) + \left(\sum_{\substack{p\mid n \\ p>z}} \omega_y^{(\beta)}(p) - \sum_{z<p\leq x} \frac{\Omega_y^{(\beta)}(p)}{p} \right) \\ &= \sum_{\substack{p\mid n \\ p\leq z}} \omega_y^{(\beta)}(p) - \mu_y^{(\beta)}(z) + O(k), \end{aligned} \tag{4}$$

where this last bound comes from Mertens' formula and the fact that the integer $n \leq x$ may have at most k distinct prime factors $p > z$. Note that the error term is uniform in k. By the binomial theorem and the uniformity of the error term, letting

$$\mathcal{F}_y(z) := \sum_{\substack{p\mid n \\ p\leq z}} \omega_y^{(\beta)}(p) - \mu_y^{(\beta)}(z),$$

we get

$$\left(\omega_y^{(\beta)}(n) - \mu_y^{(\beta)}(x)\right)^k = \left(\mathcal{F}_y\left(z\right)\right)^k + O\left(\sum_{\ell=0}^{k-1} \binom{k}{\ell} k^{k-\ell} \left|\mathcal{F}_y\left(z\right)\right|^\ell \right). \tag{5}$$

Set $R\left(\ell\right) := \binom{k}{\ell} k^{k-\ell} \left|\mathcal{F}_y\left(z\right)\right|^\ell$. We now proceed in the same way as in the proof of Proposition 2 of Granville and Soundararajan [7]. We obtain

$$\frac{1}{x} \sum_{\ell=0}^{k-1} \sum_{n\leq x} R\left(\ell\right) \ll_{\beta,\xi} k^{3/2} C_k \left(\sigma_y^{(\beta)}\right)^{k-1},$$

so that, from (5),

$$\sum_{n\leq x} \left(\omega_y^{(\beta)}(n) - \mu_y^{(\beta)}(x)\right)^k = \sum_{n\leq x} \left(\mathcal{F}_y(z)\right)^k + O_{\beta,\xi}\left(x k^{3/2} C_k \left(\sigma_y^{(\beta)}(z)\right)^{k-1} \right). \tag{6}$$

When k is even, we have from Lemma 2 that

$$\sum_{n\leq x} \left(\mathcal{F}_y(z)\right)^k = \left(\sigma_y^{(\beta)}\right)^k \left(x C_k + O_{\beta,\xi}\left(x C_k \frac{k^{3/2}}{\sqrt{\log_2 x}} \right) \right)$$

since $k^{3/2} \leq \sigma_y^{(\beta)}(z)$. When k is odd, we have

$$\sum_{n\leq x} \left(\mathcal{F}_y(z)\right)^k \ll_{\beta,\xi} x C_k k^{3/2} \left(\sigma_y^{(\beta)}(z)\right)^{k-1} = \left(\sigma_y^{(\beta)}(z)\right)^k x C_k \frac{k^{3/2}}{\sqrt{\log_2 x}}.$$

Finally, we have

$$\mu_y^{(\beta)}(x) = \frac{\beta \log_2 x - \log_2 y}{1-\beta} + O(1),$$

so that

$$\omega_y^{(\beta)}(n) - \mu_y^{(\beta)}(x) = \omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta} + O(1)$$

uniformly for k in our range. Comparing this estimate with (4) and (5), we can replace $\mu_y^{(\beta)}(x)$ in (6) while keeping the same error term, which almost completes the proof of the corollary. Indeed, what is left is to estimate $\left(\sigma_y^{(\beta)}(z)\right)^k$. From Lemma 4, we obtain

$$\left(\sigma_y^{(\beta)}(z)\right)^k = \left(\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}\right)^k \left(1 + O_{\beta,\xi}\left(\frac{k \log k}{\log_2 x}\right)\right),$$

which completes the proof of the corollary. □

We need two additional lemmas that will provide upper bounds on the frequencies of large deviations of $\omega_y^{(\beta)}(n)$ and $\Omega_y^{(\beta)}(n)$ from its mean value. We first recall Theorem 3.8 of Tenenbaum [11], chapter 3, which we state as a lemma.

Lemma 5. *Uniformly for $3 \le y \le x$ and $0 \le \epsilon < \sqrt{\log_2 y}$, we have*

$$\frac{1}{x}\#\left\{n \le x : |\omega_y(n) - \log_2 y| > \epsilon\sqrt{\log_2 y}\right\} \ll e^{-\epsilon^2/3}$$

and

$$\frac{1}{x}\#\left\{n \le x : |\Omega_y(n) - \log_2 y| > \epsilon\sqrt{\log_2 y}\right\} \ll e^{-\epsilon^2/3}.$$

From Lemma 5, we deduce the following corollary.

Corollary 2. *Given $\beta \in (0,1)$, set $c_\beta := (1-\beta)^2/12$. Uniformly for $3 \le y \le x$ and $0 \le \epsilon < \sqrt{\log_2 y}$, we have*

$$\frac{1}{x}\#\left\{n \le x : \left|\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}\right| > \epsilon\sqrt{\log_2 x}\right\} \ll_\beta e^{-c_\beta \epsilon^2}.$$

Moreover, the same estimate is valid for $\Omega_y^{(\beta)}(n)$.

Proof. From the definition of $\Omega_y^{(\beta)}(n)$, we have

$$\omega_y^{(\beta)}(n) = \frac{\beta}{1-\beta}\omega(n) - \frac{1}{1-\beta}\omega_y(n).$$

From this, it follows that

$$\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta} = \left(\frac{\beta\omega(n)}{1-\beta} - \frac{\beta \log_2 x}{1-\beta}\right) + \left(\frac{\log_2 y}{1-\beta} - \frac{\omega_y(n)}{1-\beta}\right).$$

By the triangle inequality, we have

$$\left|\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}\right| \le \left|\frac{\beta\omega(n)}{1-\beta} - \frac{\beta\log_2 x}{1-\beta}\right| + \left|\frac{\log_2 y}{1-\beta} - \frac{\omega_y(n)}{1-\beta}\right|. \quad (7)$$

Assuming that

$$\left|\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}\right| > \epsilon\sqrt{\log_2 x},$$

it follows that either

$$|\omega(n) - \log_2 x| > \frac{(1-\beta)\epsilon\sqrt{\log_2 x}}{2\beta}$$

or

$$|\omega_y(n) - \log_2 y| > \frac{(1-\beta)\epsilon\sqrt{\log_2 x}}{2} > \frac{(1-\beta)\,\epsilon\sqrt{\log_2 y}}{2}.$$

Using Lemma 5, the first part of Corollary 2 then follows. The proof for $\Omega_y^{(\beta)}(n)$ is similar. □

Lemma 6. *There exists a positive constant C such that, for all $k \le \frac{\log x}{\log 2}$, we have*

$$\#\{n \le x : \Omega(n) > k\} \le C\frac{x(\log x)(\log_2 x)^4}{2^k}.$$

Proof. Erdős and Sárközy [5, eq (17)] have shown that for all $k \ge 1$, there exists a constant $C > 0$ such that,

$$\#\{n \le x : \Omega(n) > k\} \le C\frac{x(\log x)k^4}{2^k}. \quad (8)$$

On the other hand, it follows from the results of Nicolas in [8] that for $B\log\log x \le k \le \log x/\log 2$ (with $B > 2$), we have

$$\#\{n \le x : \Omega(n) > k\} \le \frac{x\log x}{2^k}. \quad (9)$$

Gathering inequalities (8) and (9), the lemma follows. □

Finally, we will be using a technical result of Esseen [6] which has been used in the proof of Theorem 2 in Rényi and Turán [10].

Lemma 7. (Esseen) *Let $\varepsilon > 0$. Let $F(x)$ and $G(x)$ be two distribution functions such that $G'(x)$ exists for all x and $|G'(x)| \le A$ for some positive constant A. Further let $f(u) = \int_{-\infty}^{\infty} e^{iux}\mathrm{d}F(x)$ and $g(u) = \int_{-\infty}^{\infty} e^{iux}\mathrm{d}G(x)$ denote their respective characteristic functions. Then, if the condition*

$$\int_{-T}^{T}\left|\frac{f(u)-g(u)}{u}\right|\mathrm{d}u < \varepsilon$$

is satisfied, we have for all real t,

$$|F(t) - G(t)| < K\left(\varepsilon + \frac{A}{T}\right)$$

for some absolute constant $K > 0$.

3. The distribution of the $\omega_y^{(\beta)}(n)$ function

In this section, we investigate the distribution of the $\omega_y^{(\beta)}(n)$ function. Our goal is to prove to following result.

Theorem 3. *Given any fixed real numbers $\beta \in (0,1)$, $\xi \in (0,1)$ and integer y such that $\log_2 y = (\xi + o(1)) \log_2 x$ as $x \to \infty$, then, for any $t \in \mathbb{R}$ such that $t = o(\log_2 x)$, we have that*

$$\frac{1}{x}\#\left\{n \le x : \frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} < t\right\} = \Phi(t) + O_{\beta,\xi}\left(\frac{1}{\sqrt{\log_3 x}}\right),$$

which when $\beta = 1/2$, simplifies to

$$\frac{1}{x}\#\left\{n \le x : \frac{\omega_y^{(1/2)}(n) - (\log_2 x - 2\log_2 y)}{\sqrt{\log_2 x}} < t\right\} = \Phi(t) + O_{\xi}\left(\frac{1}{\sqrt{\log_3 x}}\right).$$

Corollary 3. *Given any fixed real numbers $\beta \in (0,1)$, $\xi \in (0,1)$ and integer y such that $\log_2 y = (\xi + o(1)) \log_2 x$ as $x \to \infty$ as $x \to \infty$, then, for any $t \in \mathbb{R}$ such that $t^2 \ll \sqrt{\log_2 x}$, we have that*

$$\frac{1}{x}\#\left\{n \le x : \frac{\Omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} < t\right\} = \Phi(t) + O_{\beta,\xi}\left(\frac{1}{\sqrt{\log_3 x}}\right).$$

In order to prove Theorem 3, we will apply Lemma 7 with $G(x) = \Phi(x)$, allowing us to choose $A = 2$. The characteristic function is

$$\frac{1}{\lfloor x \rfloor}\sum_{n \le x} \exp\left(iu\frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right) = f(u)\left(1 + O\left(\frac{1}{x}\right)\right),$$

where

$$f(u) := \frac{1}{x}\sum_{n \le x} \exp\left(iu\frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right). \quad (10)$$

Set $g(u)=e^{-u^2/2}$. When $\frac{-1}{\sqrt{\log_2 x}} \leq u \leq \frac{1}{\sqrt{\log_2 x}}$, we have

$$g(u) = 1 + O\left(u^2\right) = 1 + O\left(u\right).$$

Moreover,

$$f(u) = \frac{1}{x}\sum_{s\geq 0}\frac{(iu)^s}{s!}\sum_{n\leq x}\left(\frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s.$$

Hence, it follows that

$$f(u) = 1 + O\left(\frac{1}{x}\sum_{s\geq 1}\frac{|u|^s}{s!}\left|\sum_{n\leq x}\left(\frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s\right|\right). \tag{11}$$

Using Corollary 1, we get for $s \leq \frac{N^{1/3}}{2}\left(\log_2 x\right)^{1/3}$ that

$$\frac{1}{x}\sum_{n\leq x}\left(\frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s \ll_{\beta,\xi} C_s = \frac{\Gamma\left(s+1\right)}{\Gamma\left(\frac{s}{2}+1\right)2^{s/2}},$$

so that, if $N_0 := \frac{N^{1/3}}{2}\left(\log_2 x\right)^{1/3}$,

$$\frac{1}{x}\sum_{s\leq N_0}\frac{|u|^s}{s!}\left|\sum_{n\leq x}\left(\frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s\right| \ll_{\beta,\xi} u. \tag{12}$$

From Stirling's formula, we also have

$$\begin{aligned}&\frac{1}{x}\sum_{s> N_0}\frac{|u|^s}{s!}\left|\sum_{n\leq x}\left(\frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s\right| \\ &\qquad \leq \frac{1}{x}\sum_{s>N_0}\left(eL_1\frac{|u|}{s\sqrt{\log_2 x}}\right)^s\left|\sum_{n\leq x}\left(\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}\right)^s\right| \end{aligned} \tag{13}$$

for some positive constant L_1 depending only on β and ξ. We bound (13) by treating separately the cases $\omega(n) \leq \left(\log_2 x\right)^{7/6}$ and $\omega(n) > \left(\log_2 x\right)^{7/6}$. First observe that, for any constant L depending only on β and ξ,

$$\begin{aligned}\sum_{s>N_0}\left(\frac{eL\,|u|}{s\sqrt{\log_2 x}}\right)^s\frac{1}{x}\sum_{\substack{n\leq x\\ \omega(n)\leq(\log_2 x)^{7/6}}}\left(\log_2 x\right)^{7s/6} &\leq \sum_{s>N_0}\left(\frac{eL\left(\log_2 x\right)^{2/3}|u|}{s}\right)^s \\ &\leq \sum_{s>N_0}\left(\frac{2eL\left(\log_2 x\right)^{1/3}}{N^{1/3}}\,|u|\right)^s \ll u.\end{aligned}$$

Since there exists a constant L_2 depending only on β and ξ such that

$$\left|\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right|<\frac{L_2\omega(n)}{\sqrt{\log_2 x}}$$

when $\omega(n)>(\log_2 x)^{7/6}$, we obtain from (13) that

$$\frac{1}{x}\sum_{s>N_0}\frac{|u|^s}{s!}\left|\sum_{\substack{n\le x\\ \omega(n)>(\log_2 x)^{7/6}}}\left(\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s\right|$$
$$\ll\sum_{s>N_0}\left(\frac{eL_2|u|}{s\sqrt{\log_2 x}}\right)^s\frac{1}{x}\sum_{\substack{n\le x\\ \omega(n)>(\log_2 x)^{7/6}}}\omega(n)^s. \qquad (14)$$

Using Lemma 6, we have

$$\frac{1}{x}\sum_{\substack{n\le x\\ \omega(n)>(\log_2 x)^{7/6}}}\omega(n)^s\ll(\log x)(\log_2 x)^4\sum_{(\log_2 x)^{7/6}<j\le\frac{\log x}{\log 2}}\left(\frac{e^{\frac{s\log j}{j}}}{2}\right)^j.$$

For $t>1$, the function

$$h(t)=\left(\frac{e^{\frac{s\log t}{t}}}{2}\right)^t$$

reaches its maximum when $t=\frac{s}{\log 2}$, so that

$$\frac{1}{x}\sum_{\substack{n\le x\\ \omega(n)>(\log_2 x)^{7/6}}}\omega(n)^s\ll(\log x)^2(\log_2 x)^4\left(\frac{s}{e\log 2}\right)^s.$$

It follows from (14) that

$$\frac{1}{x}\sum_{s>\log_2 x}\frac{|u|^s}{s!}\left|\sum_{\substack{n\le x\\ \omega(n)>(\log_2 x)^{7/6}}}\left(\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s\right|\ll u. \qquad (15)$$

When $N_0<s\le\log_2 x$, we obtain from Lemma 6 that

$$\frac{1}{x}\sum_{\substack{n\le x\\ \omega(n)>(\log_2 x)^{7/6}}}\omega(n)^s\ll(\log x)^2(\log_2 x)^4\frac{(\log_2 x)^{7s/6}}{2^{(\log_2 x)^{7/6}}}.$$

Thus,

$$\frac{1}{x}\sum_{N_0<s\leq\log_2 x}\frac{|u|^s}{s!}\left|\sum_{\substack{n\leq x\\ \omega(n)>(\log_2 x)^{7/6}}}\left(\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s\right|\ll u. \tag{16}$$

Combining (12), (13), (15) and (16), it follows from (11) that

$$f(u)=1+O_{\beta,\xi}(u).$$

Thus, we obtain

$$f(u)=g(u)+O_{\beta,\xi}\left(u\right) \tag{17}$$

for $\frac{-1}{\sqrt{\log_2 x}}\leq u\leq\frac{1}{\sqrt{\log_2 x}}$. Now, set $T:=\sqrt{\frac{\log_3 x}{2}}$. We want to show that $g(u)$ is a good estimation of $f(u)$ when $\frac{1}{\sqrt{\log_2 x}}<|u|\leq T$. By expanding the exponential, we obtain

$$f(u)=\frac{1}{x}\sum_{s\geq0}\frac{1}{s!}\sum_{n\leq x}\left(iu\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s=f_1(u)+f_2(u)+f_3(u)+f_4(u),$$

where $0\leq s<S_1$ in f_1, $S_1\leq s<S_2$ in f_2, $S_2\leq s<S_3$ in f_3 and $s\geq S_3$ in f_4. For reasons that will become clear during the course of the proof, we set

$$S_1:=\exp\left(\frac{\log_3 x}{4\log_4 x}\right),\quad S_2:=(\log_2 x)^{1/3-\varepsilon},\quad S_3:=(\log_2 x)^{1+\varepsilon},$$

where ε is a real number satisfying $0<\varepsilon<1/10$.

3.1. The estimation of $f_1(u)$

Lemma 8. *Uniformly for* $|u|\leq\sqrt{\frac{\log_3 x}{2}}$, *we have*

$$f_1(u)=1+\sum_{s=2,\, s\text{ even}}^{\infty}\frac{C_s(iu)^s}{s!}+O_{\beta,\xi}\left(\frac{(\log_3 x)^{5/2}}{(\log_2 x)^{1/4}}\right),$$

where the C_s*'s are the constants defined in (3).*

Proof. Recall that we chose $S_1<\frac{N^{1/3}}{2}(\log_2 x)^{1/3}$ so that $f_1(u)$ can be estimated using the result of Soundararajan and Granville [7], here given by Lemma 2 and Corollary 1. Indeed, using Corollary 1, we have

$$f_1(u)=\frac{1}{x}\sum_{0\leq s<S_1}\frac{1}{s!}\sum_{n\leq x}\left(iu\frac{\omega_y^{(\beta)}(n)-\frac{\beta\log_2 x-\log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}}\right)^s=\Sigma_1+O\left(\Sigma_2\right),$$

where

$$\Sigma_1 = \sum_{0 \le s < S_1,\, s \text{ even}} \frac{C_s (iu)^s}{s!}$$

and

$$\Sigma_2 = \sum_{0 \le s < S_1} \frac{|u|^s}{s!} \frac{s^{3/2} C_s}{\sqrt{\log_2 x}}.$$

We have

$$\begin{aligned} \Sigma_2 &\ll_{\beta,\xi} \frac{1}{\sqrt{\log_2 x}} \sum_{0 \le s < S_1} s^2 \frac{C_s}{s!} \left(\frac{\log_3 x}{2}\right)^{s/2} \\ &\ll \frac{1}{\sqrt{\log_2 x}} \sum_{3 \le s < S_1} \frac{s^2}{\sqrt{s/2}} \left(\frac{e}{s/2}\right)^{s/2} \left(\frac{\log_3 x}{4}\right)^{s/2} \ll_{\beta,\xi} \frac{(\log_3 x)^{5/2}}{(\log_2 x)^{1/4}}. \end{aligned} \tag{18}$$

From Equation (18), we find that

$$f_1(u) = \sum_{0 \le s < S_1,\, s \text{ even}} \frac{C_s (iu)^s}{s!} + O_{\beta,\xi}\left(\frac{(\log_3 x)^{5/2}}{(\log_2 x)^{1/4}}\right). \tag{19}$$

Furthermore, we have

$$\begin{aligned} \left| \sum_{0 \le s < S_1,\, s \text{ even}} \frac{C_s (iu)^s}{s!} - \sum_{s \text{ even}} \frac{C_s (iu)^s}{s!} \right| &\le \sum_{s > S_1} \frac{C_s |u|^s}{s!} \ll \frac{1}{\sqrt{S_1}} \sum_{s > S_1} \left(\frac{e \log_3 x}{2S_1}\right)^{s/2} \\ &\ll \frac{1}{\sqrt{S_1}} \left(\frac{e \log_3 x}{2S_1}\right)^{S_1/2} = o\left(\frac{1}{\sqrt{\log_2 x}}\right). \end{aligned} \tag{20}$$

Using (19) and (20), the proof of Lemma 8 is thus complete. □

3.2. The estimation of $f_2(u)$

Lemma 9. *Uniformly for* $|u| \le \sqrt{\frac{\log_3 x}{2}}$, *we have*

$$|f_2(u)| = o_{\beta,\xi}\left(\frac{1}{\log_2 x}\right) \qquad (x \to \infty).$$

Proof. Since $S_2 = o\left((\log_2 x)^{1/3}\right)$ as $x \to \infty$, we can once again use Corollary 1 and obtain

$$\begin{aligned} |f_2(u)| &= \left| \frac{1}{x} \sum_{S_1 \le s < S_2} \frac{1}{s!} \sum_{n \le x} \left(iu \frac{\omega_y^\beta(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta) \log_2 y}{(1-\beta)^2}}} \right)^s \right| \ll_{\beta,\xi} \sum_{S_1 \le s < S_2} \frac{C_s |u|^s}{s!} \\ &\ll \frac{1}{\sqrt{S_1}} \sum_{s \ge S_1} \left(\frac{e \log_3 x}{2S_1}\right)^{s/2} = o\left(\frac{1}{\sqrt{\log_2 x}}\right). \end{aligned}$$

□

3.3. The estimation of $f_3(u)$

Lemma 10. *Uniformly for* $|u| \le \sqrt{\frac{\log_3 x}{2}}$, *we have*

$$|f_3(u)| \ll \frac{1}{\sqrt{\log_2 x}}.$$

Proof. We begin by writing

$$f_3(u) = \frac{1}{x} \sum_{S_2 \le s < S_3} f_{3,s}(u), \tag{21}$$

where

$$f_{3,s}(u) = \frac{1}{s!} \sum_{n \le x} \left(iu \frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} \right)^s.$$

To simplify the notation, we set

$$A(n) := \frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}. \tag{22}$$

With this, we get

$$|f_{3,s}(u)| \le \frac{|u|^s}{s!} \sum_{n \le x} |A(n)|^s \ll \left(\frac{e \cdot \log_3 x}{s} \right)^s \sum_{n \le x} |A(n)|^s. \tag{23}$$

We will now provide upper bounds for $\sum_{S_2 \le s < S_3} |f_{3,s}(u)|$ depending on the size of $A(n)$. For this, we let $0 < \delta < 1$ be an arbitrarily small number and examine three cases separately.

Case 1. $|A(n)| < \left(\frac{s}{e \cdot \log_3 x} \right)^{1-\delta}$,

Case 2. $\left(\frac{s}{e \cdot \log_3 x} \right)^{1-\delta} \le |A(n)| < (\log_2 x)^{1/2+2\delta}$,

Case 3. $|A(n)| \ge (\log_2 x)^{1/2+2\delta}$.

Case 1. In this case, we have, using Equation (23),

$$|f_{3,s}(u)| \le x \left(\frac{e \cdot \log_3 x}{s} \right)^{\delta s}.$$

Summing the above with respect to s with $S_2 < s \le S_3$, we have

$$\sum_{S_2 \le s < S_3} |f_{3,s}(u)| \ll x \sum_{s \ge S_2} \left(\frac{e \log_3 x}{s} \right)^{\delta s} \le x \sum_{s \ge S_2} \left(\frac{e \log_3 x}{(\log_2 x)^{1/3-\delta}} \right)^{\delta s} = O\left(\frac{x}{\log_2 x} \right).$$

Case 2. In this case, we have

$$\begin{aligned} |f_{3,s}(u)| &\le \left(\frac{e \log_3 x}{s} \right)^s \sum_{\substack{n \le x \\ \left(\frac{s}{e \log_3 x}\right)^{1-\delta} \le |A(n)| \\ |A(n)| < (\log_2 x)^{1/2+2\delta}}} |A(n)|^s \\ &\le \left(\frac{e (\log_3 x)(\log_2 x)^{1/2+2\delta}}{s} \right)^s \sum_{\substack{n \le x \\ \left(\frac{s}{e \cdot \log_3 x}\right)^{1-\delta} \le |A(n)|}} 1. \end{aligned}$$

Using Corollary 2, since

$$|A(n)| \ge \left(\frac{s}{e \log_3 x} \right)^{1-\delta} \quad \text{if and only if} \quad \left| \frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta) \log_2 y}{(1-\beta)^2}}} \right| \ge \left(\frac{s}{e \log_3 x} \right)^{1-\delta}$$

so that $\left| \omega_y^{(\beta)}(n) - \dfrac{\beta \log_2 x - \log_2 y}{1-\beta} \right| > \dfrac{1}{2} \sqrt{\dfrac{\beta^2 + \xi - 2\beta\xi}{(1-\beta)^2}} \left(\dfrac{s}{e \log_3 x} \right)^{1-\delta} \sqrt{\log_2 x},$

we obtain

$$|f_{3,s}(u)| \ll_\beta x \left(\frac{e (\log_3 x)(\log_2 x)^{1/2+2\delta}}{s} \right)^s \exp\left(-\frac{\beta^2 + \xi - 2\beta\xi}{48} \left(\frac{s}{e \log_3 x} \right)^{2-2\delta} \right).$$

This yields

$$|f_{3,s}(u)| \ll x \exp\left(-\frac{\beta^2 + \xi - 2\beta\xi}{96} s^{2-2\delta} \right).$$

Summing the above over s, we obtain that

$$\sum_{S_2 \le s < S_3} |f_{3,s}(u)| \ll x \sum_{S_2 \le s < S_3} \exp\left(-\frac{\beta^2 + \xi - 2\beta\xi}{96} s^{2-2\delta} \right) \ll \frac{x}{\log_2 x}.$$

Case 3. In this case, since $|A(n)| \ge (\log_2 x)^{1/2+2\delta}$, it follows that there exist positive constants $c_{\beta,\xi}$ and $d_{\beta,\xi}$ depending only on β and ξ such that $\omega(n) \ge c_{\beta,\xi} (\log_2 x)^{1+2\delta}$ and that

$$|A(n)| \le \frac{2\omega(n)}{d_{\beta,\xi} \sqrt{\log_2 x}}.$$

Thus, we have

$$|f_{3,s}(u)| \ll \left(\frac{2e\log_3 x}{sd_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \sum_{\substack{n\le x\\ \omega(n)\ge c_{\beta,\xi}(\log_2 x)^{1+2\delta}}} \omega(n)^s$$
$$= \left(\frac{2e\log_3 x}{sd_\beta\sqrt{\log_2 x}}\right)^s \sum_{j\ge c_{\beta,\xi}(\log_2 x)^{1+2\delta}} \sum_{\substack{n\le x\\ \omega(n)=j}} j^s.$$

Using Lemma 6, we obtain

$$|f_{3,s}(u)| \ll x(\log x)(\log_2 x)^4 \left(\frac{2e\log_3 x}{sd_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \sum_{c_{\beta,\xi}(\log_2 x)^{1+2\delta}\le j\le \frac{\log x}{\log 2}} \frac{j^s}{2^j}. \tag{24}$$

Now observe that , choosing $\delta=\varepsilon$, we obtain

$$j^{1-\varepsilon/2} \ge (c_{\beta,\xi})^{1-\varepsilon/2}(\log_2 x)^{(1+2\varepsilon)(1-\varepsilon/2)} \gg_{\beta,\xi} (\log_2 x)^{1+\frac{3}{2}\varepsilon-\varepsilon^2},$$

so that $(\log_2 x)^{1+\varepsilon} = o\left(j^{1-\varepsilon/2}\right)$. From this, it follows that

$$\sum_{j\ge c_{\beta,\xi}(\log_2 x)^{1+2\delta}} \frac{j^s}{2^j} \le \sum_{j\ge c_{\beta,\xi}(\log_2 x)^{1+2\delta}} \left(\frac{3}{4}\right)^j \ll \exp\left(c_{\beta,\xi}(\log_2 x)^{1+2\delta}(\log 3-\log 4)\right),$$

which allows us to conclude from (24) that

$$\sum_{S_2\le s<S_3} |f_{3,s}(u)| \ll xS_3 \exp\left(\frac{c_{\beta,\xi}}{2}(\log_2 x)^{1+2\delta}(\log 3-\log 4)\right) = o\left(\frac{x}{\sqrt{\log_2 x}}\right).$$

Gathering the estimates from cases 1, 2 and 3 in relation (21), the proof of Lemma 10 is complete. □

3.4. The estimation of $f_4(u)$

Lemma 11. *Uniformly for $|u|\le\sqrt{\frac{\log_3 x}{2}}$, we have*

$$|f_4(u)| = o\left(\frac{1}{\sqrt{\log_2 x}}\right).$$

Proof. Recalling the definition of $A(n)$ given in (22), we have

$$|f_4(u)| \ll \frac{1}{x\sqrt{S_3}} \sum_{s\ge S_3} \frac{(e\log_3 x)^s}{s^s} \sum_{j\ge 1} \sum_{\substack{n\le x\\ \omega(n)=j}} |A(n)|^s. \tag{25}$$

Given that

$$|A(n)| \le \frac{c_1\omega(n) + c_2\log_2 x}{d_{\beta,\xi}\sqrt{\log_2 x}}$$

for some positive constants c_1 and c_2 depending only on β, and $d_{\beta,\xi}$ depending only on β and ξ, we have from Equation (25) that

$$|f_4(u)| \ll \frac{1}{x\sqrt{\log_2 x}} \sum_{s\ge S_3} \left(\frac{e\log_3 x}{sd_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \sum_{j\ge 1} \sum_{\substack{n\le x\\ \omega(n)=j}} (c_1 j + c_2\log_2 x)^s. \tag{26}$$

From Lemma 6 and Inequality (26), we obtain

$$|f_4(u)| \ll \frac{(\log x)(\log_2 x)^4}{\sqrt{\log_2 x}} \sum_{s\ge S_3} \left(\frac{e\log_3 x}{sd_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \sum_{1\le j\le \frac{\log x}{\log 2}} \frac{(c_1 j + c_2\log_2 x)^s}{2^j}. \tag{27}$$

We will evaluate the above sum treating the cases $j \le \frac{c_2}{c_1}\log_2 x$ and $j > \frac{c_2}{c_1}\log_2 x$ separately. Let

$$T_1(x) := \sum_{s\ge S_3} \left(\frac{e\log_3 x}{sd_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \sum_{1\le j\le \frac{c_2}{c_1}\log_2 x} \frac{(c_1 j + c_2\log_2 x)^s}{2^j}$$

and

$$T_2(x) := \sum_{s\ge S_3} \left(\frac{e\log_3 x}{sd_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \sum_{\frac{c_2}{c_1}\log_2 x< j\le \frac{\log x}{\log 2}} \frac{(c_1 j + c_2\log_2 x)^s}{2^j}.$$

First assuming $j \le \frac{c_2}{c_1}\log_2 x$, we get $c_1 j + c_2\log_2 x \le 2c_2\log_2 x$, so that

$$T_1(x) \le \sum_{s\ge S_3} \left(\frac{2ec_2\log_2 x\log_3 x}{S_3\cdot d_{\beta,\xi}\sqrt{\log_2 x}}\right)^s \ll \left(\frac{2ec_2\log_2 x\log_3 x}{S_3\cdot d_{\beta,\xi}\sqrt{\log_2 x}}\right)^{S_3}. \tag{28}$$

On the other hand, assuming that $j > \frac{c_2}{c_1}\log_2 x$, we have $c_1 j + c_2\log_2 x \le 2c_1 j$, in which case

$$T_2(x) \le \sum_{s\ge S_3} \left(\frac{2c_1 e\log_3 x}{sc_\beta\sqrt{\log_2 x}}\right)^s \sum_{\frac{c_2}{c_1}\log_2 x< j\le \frac{\log x}{\log 2}} \frac{j^s}{2^j}. \tag{29}$$

One can easily check that the maximum value of $j^s/2^j$ is reached when $j = s/\log 2$ and is therefore equal to $\left(\frac{s}{e\log 2}\right)^s$. Thus,

$$\sum_{\frac{c_2}{c_1}\log_2 x< j\le \frac{\log x}{\log 2}} \frac{j^s}{2^j} \ll \log x \left(\frac{s}{e\log 2}\right)^s.$$

Substituting this bound in (29), we find that, as $x \to \infty$

$$T_2(x) \ll \log x \sum_{s \geq S_3} \left(\frac{2c_1 e \log_3 x}{s d_{\beta,\xi} \sqrt{\log_2 x}} \right)^s \left(\frac{s}{e \log 2} \right)^s \ll \log x \left(\frac{2c_1 \log_3 x}{d_{\beta,\xi} \log 2 \sqrt{\log_2 x}} \right)^{S_3}$$
$$= o\left(\frac{1}{\sqrt{\log_2 x}} \right). \tag{30}$$

Combining (28) and (30) in (27), we conclude that $|f_4(u)| = o\left(\frac{1}{\sqrt{\log_2 x}} \right)$. □

Gathering the estimates from Lemmas 8, 9, 10 and 11, we conclude that

$$|f(u) - g(u)| \ll_{\beta,\xi} \frac{(\log_3 x)^{5/2}}{(\log_2 x)^{1/4}} \tag{31}$$

uniformly for $|u| \leq \sqrt{\frac{\log_3 x}{2}}$.

3.5. Completion of the proof of Theorem 3

We have

$$\int_{-T}^{T} \left| \frac{f(u) - g(u)}{u} \right| \mathrm{d}u = \Delta_1(x) + \Delta_2(x), \tag{32}$$

where

$$\Delta_1(x) := \int_{\frac{-1}{\sqrt{\log_2 x}}}^{\frac{1}{\sqrt{\log_2 x}}} \left| \frac{f(u) - g(u)}{u} \right| \mathrm{d}u$$

and

$$\Delta_2(x) := \int_{\frac{1}{\sqrt{\log_2 x}} < |u| \leq T} \left| \frac{f(u) - g(u)}{u} \right| \mathrm{d}u.$$

From (31), we have

$$\Delta_2(x) \ll_\beta \frac{(\log_3 x)^{5/2}}{(\log_2 x)^{1/4}} \int_{\frac{1}{\sqrt{\log_2 x}}}^{T} \frac{1}{u} \mathrm{d}u \ll \frac{(\log_3 x)^{7/2}}{(\log_2 x)^{1/4}}.$$

On the other hand, from (17), we easily get that

$$\Delta_1(x) \ll_{\beta,\xi} \int_{\frac{-1}{\sqrt{\log_2 x}}}^{\frac{1}{\sqrt{\log_2 x}}} \mathrm{d}u \leq \frac{2}{\sqrt{\log_2 x}}.$$

Hence,

$$\int_{-T}^{T} \left| \frac{f(u) - g(u)}{u} \right| \mathrm{d}u \ll_\beta \frac{(\log_3 x)^{7/2}}{(\log_2 x)^{1/4}}.$$

From Esseen's result (Lemma 7), it follows that, for any $t \in \mathbb{R}$,

$$\frac{1}{x}\#\left\{n \le x : \frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} \le t\right\} = \Phi(t) + O_{\beta,\xi}\left(\frac{1}{\sqrt{\log_3 x}}\right). \tag{33}$$

In particular, for any $t \in \mathbb{R}$ such that $t = o(\log_2 x)$ as $x \to \infty$, we have

$$\Phi\left(t + O\left(\frac{1}{(\log x)^2}\right)\right) = \Phi(t) + O\left(\frac{1}{\log x}\right). \tag{34}$$

Indeed, such a formula can be obtained by the fact that

$$\frac{2}{\sqrt{\pi}}\int_0^x e^{-v^2}\mathrm{d}v = 1 + O\left(\frac{e^{-x^2}}{x}\right) \qquad (x \to \infty),$$

so that, for any $Y \ge 0$,

$$\Phi(t+h) = \frac{1}{\sqrt{2\pi}}\int_{-Y}^{t+h} e^{-v^2/2}\mathrm{d}v + O\left(\frac{e^{-Y^2}}{Y}\right).$$

In particular, if $Y = Y(x)$, $t = t(x)$ and $h = h(x)$ are such that $h = o(Y)$, $th = o(1)$, and $Yh = o(1)$, then

$$\Phi(t+h) = \Phi(t)\left(1 + O\left(Y|h| + |th| + |h|^2\right)\right) + O\left(\frac{e^{-Y^2}}{Y}\right). \tag{35}$$

Estimate (34) follows by choosing $Y = \log x$, $h = O\left((\log x)^{-2}\right)$ and $t = o(\log_2 x)$.

We therefore obtain from (33) and (34) that

$$\frac{1}{x}\#\left\{n \le x : \frac{\omega_y^{(\beta)}(n) - \frac{\beta\log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} < t\right\} = \Phi(t) + O_{\beta,\xi}\left(\frac{1}{\sqrt{\log_3 x}}\right), \tag{36}$$

thus completing the proof of Theorem 3. To prove Corollary 3, first observe that

$$\Omega_y^{(\beta)}(n) - \omega_y^{(\beta)}(n) = \frac{\beta}{1-\beta}\left(\Omega(n) - \omega(n)\right) - \frac{1}{1-\beta}\left(\Omega_y(n) - \omega_y(n)\right).$$

Uniformly for real $k \ge 1$, we have

$$\#\left\{n \le x : \Omega(n) - \omega(n) \ge k\right\} \ll \frac{x}{2^{k/2}}.$$

By choosing $k = \log_3 x$, it follows that

$$\#\left\{n \le x : \Omega_y^{(\beta)}(n) - \omega_y^{(\beta)}(n) \ge \log_3 x\right\} \ll_\beta \frac{x}{(\log_2 x)^{\frac{\log 2}{2}}}. \tag{37}$$

Hence, we have

$$\Omega_y^{(\beta)}(n) = \omega_y^{(\beta)}(n) + O\left(\log_3 x\right)$$

for almost all integers $n \le x$. Moreover, for any real t satisfying $t^2 \ll_\beta \sqrt{\log_2 x}$, we have

$$\Phi\left(t + O_\beta\left(\frac{\log_3 x}{\sqrt{\log_2 x}}\right)\right) = \Phi(t) + O_\beta\left(\frac{\log_3 x}{(\log_2 x)^{1/4}}\right), \tag{38}$$

which can be obtained from (35) by choosing $Y = (\log_2 x)^{1/4}$, $h \ll \frac{\log_3 x}{\sqrt{\log_2 x}}$ and $t \ll (\log_2 x)^{1/4}$.

The proof of Corollary 3 then follows from (36), (37) and (38).

4. Proof of the main results

Theorems 1 and 2 will follow rather directly from Theorem 3. Indeed, we obtain from Lemma 1 that, for any $t \in \mathbb{R}$ such that $t^2 = o\left(\sqrt{\log_2 x}\right)$,

$$\begin{aligned}
&\#\left\{n \le x : \frac{\log_2 p^{(\beta)}(n) - \beta \log_2 x}{\sqrt{\log_2 x}} < t\right\} \\
&\quad = \#\left\{n \le x : p^{(\beta)}(n) < \exp\left(\exp\left(\beta \log_2 x + t\sqrt{\log_2 x}\right)\right)\right\} \\
&\quad = \#\left\{n \le x : \omega_y^{(\beta)}(n) < 1\right\} + O_\beta\left(\frac{x\left(\log_2 x\right)^{\frac{1-2\beta}{\beta}}}{\log x}\right) \qquad (39)
\end{aligned}$$

provided we choose $y := \exp\left(\exp\left(\beta \log_2 x + t\sqrt{\log_2 x}\right)\right)$. Now,

$$\omega_y^{(\beta)}(n) < 1 \quad \text{if and only if} \quad \frac{\omega_y^{(\beta)}(n) - \frac{\beta \log_2 x - \log_2 y}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} < \frac{1 + \frac{\log_2 y - \beta \log_2 x}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}}.$$

Given our choice of y, we have $\log_2 y = \beta \log_2 x + t\sqrt{\log_2 x}$, so that

$$1 + \frac{\log_2 y - \beta \log_2 x}{1-\beta} = 1 + \frac{t\sqrt{\log_2 x}}{1-\beta}.$$

For any $t \in \mathbb{R}$ such that $t = o\left(\log_2 x\right)$ as $x \to \infty$, we have

$$\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}} = \sqrt{\frac{\beta}{1-\beta}\log_2 x}\left(1 + O_\beta\left(\frac{t}{\sqrt{\log_2 x}}\right)\right).$$

Hence,

$$\frac{1 + \frac{\log_2 y - \beta \log_2 x}{1-\beta}}{\sqrt{\frac{\beta^2 \log_2 x + (1-2\beta)\log_2 y}{(1-\beta)^2}}} = \frac{t}{\sqrt{\beta\left(1-\beta\right)}} + O_\beta\left(\frac{1+t^2}{\sqrt{\log_2 x}} + \frac{t}{\log_2 x}\right).$$

Thus, if $t^2 \ll (\log_2 x)^{1/2-\varepsilon}$ for some $1/4 < \varepsilon < 1/2$, we find that

$$\frac{1+\frac{\log_2 y-\beta\log_2 x}{1-\beta}}{\sqrt{\frac{\beta^2\log_2 x+(1-2\beta)\log_2 y}{(1-\beta)^2}}} = \frac{t}{\sqrt{\beta\,(1-\beta)}} + O_\beta\left((\log_2 x)^{-\epsilon}\right). \tag{40}$$

Notice that $|t| \ll (\log_2 x)^{1/4-\varepsilon/2}$, so that, for any $h = O\left((\log_2 x)^{-\epsilon}\right)$, we have

$$\Phi\,(t+h) = \Phi(t) + O\left((\log_2 x)^{1/4-3\varepsilon/2} + (\log_2 x)^{-\epsilon/2}\right), \tag{41}$$

which can be obtained from (35) by choosing $Y = (\log_2 x)^{\epsilon/2}$, $h \ll (\log_2 x)^{-\epsilon}$, and $t \ll (\log_2 x)^{1/4-\epsilon/2}$. It follows from Theorem 3, (39), (40) and (41) that, for any real $\frac{1}{8} < \varepsilon < \frac{1}{4}$ and $t \in \mathbb{R}$ such that $|t| \ll (\log_2 x)^{1/4-\varepsilon}$, we have

$$\frac{1}{x}\#\left\{n \le x : \frac{\log_2 p^{(\beta)}(n) - \beta\log_2 x}{\sqrt{\log_2 x}} < t\right\} = \Phi\left(\frac{t}{\sqrt{\beta\,(1-\beta)}}\right) + O_\beta\left(\frac{1}{\sqrt{\log_3 x}}\right),$$

thus completing the proofs of Theorems 1 and 2. Moreover, the proof is the same in the case when the multiplicity of each prime factor is taken into account to define the β-positioned prime factor.

Acknowledgement: The authors would like to thank the referee for a careful reading.

References

[1] J.M. De Koninck and I. Kátai, Normal numbers and the middle prime factor of an integer, *Colloq. Math.* **135** (2014), No. 1, 69-77.

[2] J.M. De Koninck and F. Luca, On the middle prime factor of an integer, *J. Integer Seq.* **16** (2013), Article 13.5.5.

[3] H. Delange, Sur les formules de Atle Selberg, *Bull. Sc. Math.*, 2e série **83** (1959), 101-111.

[4] N. Doyon and V. Ouellet, On the sum of the reciprocals of the middle prime factors taking into account their multiplicity, *Ann. Univ. Sci. Budapest. Sect. Comput.* **47** (2018), 249-259.

[5] P. Erdős and A. Sárközy, On the number of prime factors of integers, *Acta Sci. Math.* **42** (1980), 237-246.

[6] C.G. Esseen, Fourier analysis of distribution functions, A mathematical study of the Laplace-Gaussian law, *Acta. Math.* **77** (1945), 1-125.

[7] A. Granville and K. Soundararajan, Sieving and the Erdős-Kac Theorem. Equidistribution in Number Theory, An Introduction. *NATO Science Series II* **237** (2007) Springer, Dordrecht, 15-27.

[8] J.-L. Nicolas, Sur la distribution des nombres entiers ayant une quantité fixée de facteurs premiers, *Acta Arith.* **44** (1984), 191-200.

[9] V. Ouellet, On the sum of the reciprocals of the middle prime factors, *J. Integer Seq.* **20** (2017), Article 17.10.1.

[10] A. Rényi and P. Turán, On a theorem of Erdős-Kác, *Acta Arith.* **4** (1958), 71-84.

[11] G. Tenenbaum, *Introduction à la Théorie Analytique et Probabiliste des Nombres*, Éditions Belin, Paris, 2008.