# SIGN CHANGES OF THE ERROR TERM IN THE PILTZ DIVISOR PROBLEM

SIEGFRED BALUYOT AND CRUZ CASTILLO

Abstract. We study the function $\Delta_k(x) := \sum_{n\le x} d_k(n) - \mathrm{Res}_{s=1}(\zeta^k(s)x^s/s)$, where $k\geq 3$ is an integer, $d_k(n)$ is the $k$-fold divisor function, and $\zeta(s)$ is the Riemann zeta-function. For a large parameter $X$, we show that there exist at least $X^{\frac{349}{864}-\varepsilon}$ disjoint subintervals of $[X,2X]$, each of length $X^{\frac{13}{27}-\varepsilon}$, such that $|\Delta_3(x)|\gg x^{1/3}$ for all $x$ in the subinterval. For $k\geq 4$, we assume the Lindelöf Hypothesis and show that there exist at least $X^{\frac14-\frac{k^2-5k+8}{4k(k-1)(k-2)}-\varepsilon}$ disjoint subintervals of $[X,2X]$, each of length $X^{\frac34-\frac{3k-8}{4k(k-2)}-\varepsilon}$, such that $|\Delta_k(x)|\gg x^{\frac12-\frac{1}{2k}}$ for all $x$ in the subinterval. In particular, $\Delta_k(x)$ does not change sign in any of these subintervals. To prove these, we bound the mean square of $\sup_{0\le h\le H}|\Delta_k(x+h)-\Delta_k(x)|$, and we bound the fourth moment of $\Delta_k(x)$. Our results may be viewed as higher-degree analogues of a theorem of Heath-Brown and Tsang, who studied the case $k=2$. A crucial ingredient of our proofs is a technique recently developed by Lester, who used bounds for moments of $\zeta(s)$ to estimate the contribution of large frequencies in the trigonometric polynomial approximation to $\Delta_k(x)$.

## 1. Introduction and results

For each integer $k\geq 2$, let $d_k(n)$ be the number of ways to write $n$ as a product $n_1n_2\cdots n_k$ with each $n_i$ a positive integer. Define

$$\Delta_k(x) := \sum_{n\le x} d_k(n) - \operatorname*{Res}_{s=1}\left(\frac{\zeta^k(s)x^s}{s}\right), \tag{1.1}$$

where $\zeta(s)$ is the Riemann zeta-function. A great deal of work has been done towards finding upper bounds for the order of magnitude of $\Delta_k(x)$. The well-known Dirichlet divisor problem concerns finding the smallest $\alpha$ such that $\Delta_2(x)\ll x^{\alpha+\varepsilon}$ for all fixed $\varepsilon>0$. More generally, the *Piltz divisor problem* asks for the smallest $\alpha_k$ for which $\Delta_k(x)\ll_\varepsilon x^{\alpha_k+\varepsilon}$. It is known that $\alpha_k\geq (k-1)/(2k)$ [24, Theorem 12.6(B)], and it is widely believed that $\alpha_k=(k-1)/(2k)$. The current record for the smallest upper bound for $\alpha_2$ is $\alpha_2\le 131/416$, due to Huxley [8, 9]. Kolesnik [15] has shown that $\alpha_3\le 43/96$, and upper bounds for $\alpha_k$ for $k\geq 4$ have been obtained by Ivić [10, Theorems 13.2 and 13.3]. The Lindelöf Hypothesis is equivalent to the statement that $\alpha_k\le 1/2$ for all $k\geq 2$ [24, Theorem 13.4].

Aside from upper bounds, many other interesting properties of $\Delta_k$ have been studied. Omega results more precise than $\alpha_k\geq (k-1)/(2k)$ are known. The current best omega result is due to Soundararajan [22], who has shown, by refining ideas of Hafner [5], that

$$\Delta_k(x)=\Omega\Big((x\log x)^{\frac{k-1}{2k}}(\log\log x)^{\frac{k+1}{2k}(k^{2k/(k+1)}-1)}(\log\log\log x)^{-\frac12-\frac{k-1}{4k}}\Big).$$

2010 *Mathematics Subject Classification.* 11N37, 11M06.

*Keywords and phrases*: divisor function, Dirichlet divisor problem, moments, Riemann zeta-function, short intervals, fourth moment.

Cramér [3] has proved an asymptotic formula for the second moment of $\Delta_2$, while Tong [26] has shown that, unconditionally for $k = 3$ and assuming the Lindelöf Hypothesis for $k \geq 4$,

$$\int_X^{2X} \Big(\Delta_k(x)\Big)^2\,dx \sim \int_X^{2X} \big(C_k x^{\frac{1}{2}-\frac{1}{2k}}\big)^2\,dx \tag{1.2}$$

as $X \to \infty$, where

$$C_k = \frac{1}{\pi}\left(\frac{1}{2k}\sum_{n=1}^{\infty}\frac{d_k^2(n)}{n^{1+\frac{1}{k}}}\right)^{1/2}. \tag{1.3}$$

The error term in Cramér's asymptotic formula has been examined more closely by Lau and Tsang [17, 18, 28] and Ge and Gonek [4]. Tsang [27] has proved asymptotic formulas for the third and fourth moments of $\Delta_2$, while Ivić [10, Chapter 13] has obtained bounds for higher moments of $\Delta_2$ and $\Delta_3$. For further interesting research on moments and various other properties of $\Delta_2$, see the informative survey [29].

While moments of $\Delta_k$ have been extensively studied, much work has also been done towards understanding the mean square of $\Delta_k(x;h) := \Delta_k(x+h) - \Delta_k(x)$ with $h$ a parameter. Moments of $\Delta_k$ present data about the size of $\Delta_k(x)$, while moments of $\Delta_k(x;h)$ give information about the fluctuations of $\Delta_k$. Bounds for the mean square of $\Delta_k(x;h)$ over $[X, 2X]$ with $X^\varepsilon \ll h \ll X^{1-\frac{1}{k}-\varepsilon}$ have been obtained by Jutila [13] for $k = 2$ and Ivić [11] for $k \geq 3$. Ivić [12] has proved an asymptotic formula in the case $k = 2$ for the same range of $h$. If $h$ is instead equal to $x/T$ with $2 \leq T \leq X$, then a method of Selberg leads to a bound for the mean square of $\Delta_k(x;h)$ (see [20] and [19, (1.2)]). In recent work, Lester [19] has shown that, for certain constants $b_k$,

$$\frac{1}{X}\int_X^{2X}\left(\Delta_k\left(x + \frac{x^{1-\frac{1}{k}}}{L}\right) - \Delta_k(x)\right)^2 dx = \frac{b_k X^{1-\frac{1}{k}}}{L}(\log L)^{k^2-1} + O\left(\frac{X^{1-\frac{1}{k}}}{L}(\log L)^{k^2-2}\right) \tag{1.4}$$

unconditionally for $k = 3$ and $2 \leq L \ll X^{\frac{1}{12}-\varepsilon}$, and assuming the Lindelöf Hypothesis for $k \geq 3$ and $2 \leq L \ll X^{\frac{1}{k(k-1)}-\varepsilon}$. This agrees with a conjecture of Keating, Rodgers, Roditty-Gershon, and Rudnick [14], who studied the analogous problem in function fields and used their results to predict an asymptotic formula for the mean square of $\Delta_k(x;h)$ over $[X, 2X]$ with $h = X^\vartheta$ and $0 < \vartheta < 1 - 1/k$. Through their conjecture, Keating et al. have found an interesting connection between the mean square of $\Delta_k(x;h)$ and averages of coefficients of characteristic polynomials of random matrices. Bettin and Conrey [1] have shown that the conjecture of Keating et al. would follow from a (yet unproved) conjecture for moments of $\zeta(s)$.

Information about moments of $\Delta_k$ and $\Delta_k(x;h)$ can be used to closely examine the behavior of $\Delta_k$. Indeed, using a bound for the mean square of $\Delta_2(x;h)$ and formulas for the second and fourth moments of $\Delta_2$, Heath-Brown and Tsang [7] find long intervals on which $\Delta_2(x)$ does not change sign. More precisely, they prove the existence of at least $\gg \sqrt{X}\log^5 X$ disjoint subintervals of $[X, 2X]$, each of length a constant times $\sqrt{X}(\log X)^{-5}$, such that $|\Delta_2(x)| \gg x^{1/4}$ for all $x$ in any of the subintervals. Earlier, Tong [25] has shown for each $k \geq 2$ that there exists a constant $\beta_k > 0$ such that, for all large enough $X$, $\Delta_k(x)$ changes sign at least once in the interval $[X, X + \beta_k X^{1-\frac{1}{k}}]$. Thus, Heath-Brown and Tsang prove that the case $k = 2$ of Tong's theorem is best possible up to factors of $\log X$, since its statement becomes false if we replace $\beta_2\sqrt{X}$ by some constant times $\sqrt{X}(\log X)^{-5}$. This observation leads to the question: Is Tong's result for $k \ge 3$ also best possible?

We prove analogues of the theorem of Heath-Brown and Tsang for $k \ge 3$. Although we are not able to answer the question in the previous paragraph, we make progress towards its resolution. We prove an unconditional theorem for $k = 3$, while we assume the Lindelöf Hypothesis to handle the case $k \ge 4$. We also show that our result for $k = 3$ improves if we assume the Lindelöf Hypothesis. Recall that the constant $C_k$ is defined by (1.3).

**Theorem 1.1.** *Let $\xi$ and $\varepsilon$ be arbitrarily small positive constants. There exists a constant $X_0$ depending only on $\xi$ and $\varepsilon$ such that if $X \ge X_0$, then there are at least $X^{\frac{349}{864}+\xi-\varepsilon}$ disjoint subintervals of $[X, 2X]$, each of length $X^{\frac{13}{27}-\xi}$, such that $|\Delta_3(x)| > (\frac{1}{2}C_3 - \varepsilon)x^{1/3}$ for all $x$ in the subinterval. In particular, $\Delta_3(x)$ does not change sign in any of these subintervals.*

**Theorem 1.2.** *Assume the Lindelöf Hypothesis and let $k \ge 3$ be an integer. Let $\xi$ and $\varepsilon$ be arbitrarily small positive constants. There exists a constant $X_0$ depending only on $k$, $\xi$, and $\varepsilon$ such that if $X \ge X_0$, then there are at least*

$$\begin{cases} X^{\frac{137}{480}+\xi-\varepsilon} & \text{if } k = 3 \\ X^{\frac{1}{4}+\frac{1}{k-1}-\frac{1}{4k-8}-\frac{1}{k}+\xi-\varepsilon} & \text{if } k \ge 4 \end{cases}$$

*disjoint subintervals of $[X, 2X]$, each of length equal to*

$$\begin{cases} X^{\frac{3}{5}-\xi} & \text{if } k = 3 \\ X^{\frac{3}{4}+\frac{1}{4k-8}-\frac{1}{k}-\xi} & \text{if } k \ge 4, \end{cases}$$

*such that $|\Delta_k(x)| > (\frac{1}{2}C_k - \varepsilon)x^{\frac{1}{2}-\frac{1}{2k}}$ for all $x$ in the subinterval. In particular, $\Delta_k(x)$ does not change sign in any of these subintervals.*

Note that, similarly to the result of Heath-Brown and Tsang, Theorems 1.1 and 1.2 do not rule out the possibility that some of the disjoint subintervals may have a union that is contained in a longer subinterval on which $\Delta_k(x)$ does not change sign. On the other hand, Tong's theorem [25] implies that this longer subinterval cannot have length larger than $X^{1-\frac{1}{k}}$ times a certain constant.

To prove Theorems 1.1 and 1.2, we shall use the method of Heath-Brown and Tsang [7] for detecting intervals that do not contain sign changes of $\Delta_k(x)$. Their method requires a bound for the mean square of $\sup_{0\le h\le H} |\Delta_k(x; h)|$, with $H$ a parameter. For $k = 2$, Heath-Brown and Tsang find one by applying Jutila's bound [13] for the mean square of $\Delta_2(x; h)$. If we assume the Lindelöf Hypothesis, then we may apply Ivić's bound [11] to handle the case $k \ge 3$. This would lead to results that are weaker than those in Theorem 1.3 above and Corollary 1.5 below. One may wish to be able to apply Lester's formula (1.4) to obtain strong results. However, the method of Heath-Brown and Tsang needs information about $\Delta_k(x; h)$ when $\log h$ is small compared to $\log x$, while Lester's formula gives information only when $\log h \ge (1 - \frac{1}{k-1}) \log x$. We instead adapt Lester's proof of (1.4) and combine it with the method of Heath-Brown and Tsang to deduce a bound for the mean square of $\sup_{0\le h\le H} |\Delta_k(x; h)|$. The key idea in Lester's proof is the use of moments of $\zeta(s)$ to estimate the contribution of large frequencies in the trigonometric polynomial approximation to $\Delta_k(x)$ (see Section 4 below for more details).

Our bound for the mean square of $\sup_{0\le h\le H}|\Delta_k(x;h)|$ is unconditional, but it depends on bounds for the growth of $\zeta(s)$ on the critical line and moments of $\zeta(s)$. Define $\lambda\ge 0$ by

$$\lambda=\inf\{\ \vartheta\ge 0\ :\ \zeta(\tfrac12+it)\ll t^{\vartheta+\varepsilon}\ \text{as } t\to\infty,\ \text{for all fixed } \varepsilon>0\}. \tag{1.5}$$

The classical Weyl-Hardy-Littlewood bound [16],[24, Theorem 5.5] implies that $\lambda\le 1/6$. The current record for the smallest upper bound for $\lambda$ is $\lambda\le 13/84$, due to Bourgain [2]. Our main results do not depend on Bourgain's deep result, and the upper bound $\lambda\le 1/6$ is sufficient for our proof of Theorem 1.1. The yet unproved Lindelöf Hypothesis states that $\lambda=0$.

**Theorem 1.3.** *Let $k\ge 3$ be a fixed integer. Let $\lambda$ be defined by* (1.5)*, and suppose that $\delta$ is a fixed constant that satisfies $0\le\delta\le 1/(2k)$ and the property that*

$$\int_0^\tau |\zeta(\tfrac12+\delta+it)|^{2k}\,dt\ll\tau^{1+\varepsilon}\ \ \textit{as } \tau\to\infty,\ \textit{for all fixed } \varepsilon>0. \tag{1.6}$$

*Suppose further that $\varepsilon$ is an arbitrarily small positive constant. If $1\le H\le X$ and $1\le Y\le T\le X$, then*

$$\begin{aligned}\frac{1}{X}\int_X^{2X}\sup_{0\le h\le H}\Big(\Delta_k(x+h)-\Delta_k(x)\Big)^2dx &\ll_{k,\delta,\varepsilon}\Big(1+\frac{Y^{k-1}}{X}\Big)HX^{\varepsilon}+HX^{-2\delta+\varepsilon}T^{\frac12+\delta k}Y^{\delta k-\frac12}\\ &+\frac{X^{2+\varepsilon}}{Y^{k+1}}+X^{\varepsilon}Y^{k-2}+\frac{X^{2+\varepsilon}}{T^2}+X^{1-2\delta}Y^{2k(\delta+\lambda-2\lambda\delta)-2+\varepsilon}+X^{1-2\delta}T^{2k(\delta+\lambda-2\lambda\delta)-2+\varepsilon}.\end{aligned}\tag{1.7}$$

A theorem of Heath-Brown [6] allows us to take $\delta=1/12$ if $k=3$ and $\delta=1/8$ if $k=4$. Various $\delta$ satisfying (1.6) for other $k$ may be deduced from Theorem 8.4 of Ivić [10], but our proof of Theorem 1.3 requires $\delta\le 1/(2k)$. A well-known fact is that the Lindelöf Hypothesis is equivalent to the statement that $\delta=0$ satisfies (1.6) for all $k$ [24, Theorem 13.2].

In order for the detection method of Heath-Brown and Tsang [7] to succeed, our bound for the mean square of $\sup_{0\le h\le H}|\Delta_k(x+h)-\Delta_k(x)|$ must be of lower order of magnitude than the right-hand side of (1.2). Thus, we will require the bound

$$\frac{1}{X}\int_X^{2X}\sup_{0\le h\le H}\Big(\Delta_k(x+h)-\Delta_k(x)\Big)^2dx\ll_{k,\varepsilon}X^{1-\frac1k-\varepsilon}, \tag{1.8}$$

where $\varepsilon>0$ is an arbitrarily small positive constant. We shall prove this for certain values of $H$ by choosing specific values for the parameters in Theorem 1.3. The role of $H$ is important because it is the length of the subintervals we find in Theorems 1.1 and 1.2.

**Corollary 1.4.** *If $k=3$, then* (1.8) *holds for $H=X^{\frac{13}{27}-\varepsilon}$.*

**Corollary 1.5.** *Assume the Lindelöf Hypothesis. If $k=3$, then* (1.8) *holds for $H=X^{\frac35-\varepsilon}$. If $k\ge 4$, then* (1.8) *holds for $H=X^{\frac34+\frac{1}{4k-8}-\frac1k-\varepsilon}$.*

More than giving intervals on which $\Delta_2(x)$ does not change sign, the theorem of Heath-Brown and Tsang [7] implies that $|\Delta_2(x)|\gg x^{1/4}$ on a subset of $[X,2X]$ whose measure is $\gg X$. To deduce this lower bound for the measure, Heath-Brown and Tsang use an estimate for the fourth moment of $\Delta_2$, due to Tsang [27], who applied the Erdös-Turán inequality and van der Corput's bound for exponential sums to prove an asymptotic formula.

We combine Tsang's technique with the method of Lester [19] to find a conditional bound for the fourth moment of $\Delta_k$. We shall apply this bound to deduce a lower bound for the measure of the set of $x \in [X, 2X]$ for which $|\Delta_k(x)| \gg x^{\frac{1}{2}-\frac{1}{2k}}$. This leads to our lower bound for the number of disjoint subintervals in Theorem 1.2.

**Theorem 1.6.** *Assume the Lindelöf Hypothesis, and let $\varepsilon > 0$ be an arbitrarily small positive constant. If $k \geq 3$ and $X \geq 1$, then*

$$\frac{1}{X}\int_X^{2X} \big(\Delta_k(x)\big)^4\, dx \ll_{k,\varepsilon} X^{2-\frac{1}{k-1}+\varepsilon}.$$

Our proof of the case $k = 3$ of Theorem 1.6 can in fact be made unconditional. However, Ivić [10, Theorem 13.10] has proved through a different method that

$$\frac{1}{X}\int_X^{2X} \big(\Delta_3(x)\big)^4\, dx \ll_{\varepsilon} X^{\frac{139}{96}+\varepsilon} \tag{1.9}$$

by applying Kolesnik's [15] pointwise bound $\Delta_3(x) \ll x^{\frac{43}{96}+\varepsilon}$. This bound for the fourth moment of $\Delta_3$ is stronger than the case $k = 3$ of Theorem 1.6.

To obtain results as strong as Heath-Brown and Tsang [7] and show that Tong's theorem [25] is best possible, we would need (1.8) to hold for $H$ as large as $X^{1-\frac{1}{k}-\varepsilon}$. This would follow from (the yet unproved) Conjecture 1.1 of Keating et al. [14]. The difficulty in increasing the size of $H$ in Corollaries 1.4 and 1.5 is that the presence of the term $X^{\varepsilon}Y^{k-2}$ in (1.7) necessitates choosing $Y \ll X^{\frac{k-1}{k(k-2)}-\varepsilon}$ in order to deduce (1.8) from (1.7). This restriction for $Y$ limits the size of $H$ for which we are able to prove (1.8), because there occurs the term $HX^{-2\delta+\varepsilon}T^{\frac{1}{2}+\delta k}Y^{\delta k-\frac{1}{2}}$ in (1.7). As discussed in Lester [19], this difficulty is essentially the same as the one encountered in finding an asymptotic formula for the mean square of $\Delta_k(x; h)$ for $h$ outside the range covered by Lester's theorem. The conjecture of Keating et al. predicts that, for such $h$, this asymptotic formula will be of a different form than (1.4). This suggests a significant change in the behavior of $\Delta_k(x; h)$ for $h$ outside the range covered by Lester's theorem.

The rest of the paper is organized as follows. In Section 2, we set some notation and conventions that hold throughout this work. In Section 3, we prove some technical lemmas that are used in the proofs of our main results. We use Lester's method in Section 4 to bound moments involving the contribution of large frequencies in the trigonometric polynomial approximation to $\Delta_k(x)$. We prove Theorem 1.3 and its corollaries in Section 5. We prove Theorem 1.6 in Section 6, and prove Theorems 1.1 and 1.2 in Section 7.

## 2. Notation and conventions

For the rest of this paper, $k$ denotes an integer $\geq 3$. Most of our arguments will work for $k = 2$, but this special case is already well-understood in the context of our main results through the works of Heath-Brown and Tsang [7], Ivić [12], and Tsang [27].

We follow standard convention in analytic number theory and use $\varepsilon$ to denote an arbitrarily small positive constant whose value may vary from one line to the next. We allow implied constants to depend on $\varepsilon$ and $k$ without necessarily indicating so. We will sometimes display the dependence of implied constants on $\varepsilon$, $k$, or other quantities by using subscripts such as those in $A \ll_B C$ or $r = O_s(t)$. Implied constants will never depend on the parameters $T, X, Y, H$.

We use $e(x)$ to denote $e^{2\pi ix}$. For $\delta \geq 0$ and $x, V, Y, T > 0$, we define $Q_k(x; V)$ and $I_k(x; \delta, Y, T)$ by

$$Q_k(x;V) := \frac{x^{\frac{1}{2}-\frac{1}{2k}}}{\pi\sqrt{k}} \sum_{n\leq V/x} \frac{d_k(n)}{n^{\frac{1}{2}+\frac{1}{2k}}} \cos\left(2\pi k(nx)^{1/k} + \frac{(k-3)\pi}{4}\right) \tag{2.1}$$

and

$$I_k(x;\delta,Y,T) := \mathrm{Re}\left\{ \frac{1}{\pi i} \int_{\frac{1}{2}-\delta+iY}^{\frac{1}{2}-\delta+iT} \zeta^k(s) \frac{x^s}{s}\, ds \right\}. \tag{2.2}$$

## 3. Lemmata

The following lemma is a slight generalization of Lemma 2.5 of [19], and is the starting point of all the proofs of our main results.

**Lemma 3.1.** *Let $0 \leq \delta < 1/2$ be fixed, and let $\lambda$ be defined by* (1.5). *If $x, T \geq 1$ and $1 \leq Y \leq \min\{x, T\}$, then*

$$\Delta_k(x) = Q_k(x; Y^k/(2\pi)^k) + I_k(x; \delta, Y, T) + E_k(x; \delta, Y, T),$$

*where $Q_k$ is defined by* (2.1), *$I_k$ is defined by* (2.2), *and*

$$\begin{aligned} E_k(x;\delta,Y,T) \ll x^{1+\varepsilon}Y^{-\frac{k}{2}-\frac{1}{2}} + x^{\varepsilon}Y^{\frac{k}{2}-1} + x^{\frac{1}{2}-\delta}Y^{k(\delta+\lambda-2\lambda\delta)-1+\varepsilon} \\ + x^{1+\varepsilon}T^{-1+\varepsilon} + x^{\frac{1}{2}-\delta}T^{k(\delta+\lambda-2\lambda\delta)-1+\varepsilon}. \end{aligned} \tag{3.1}$$

*Proof.* The proof is similar to that of [19, Lemma 2.5], but we provide it since our situation is slightly more general. A standard argument using Perron's formula leads to

$$\sum_{n\leq x} d_k(n) = \frac{1}{2\pi i}\int_{1+\varepsilon-iT}^{1+\varepsilon+iT} \zeta^k(s)\frac{x^s}{s}\, ds + O\left(x^{\varepsilon} + \frac{x^{1+\varepsilon}}{T}\right).$$

We deform the contour of integration to the path consisting of line segments connecting the points $1+\varepsilon-iT$, $\frac{1}{2}-\delta-iT$, $\frac{1}{2}-\delta+iT$, and $1+\varepsilon+iT$, leaving a residue from the pole of $\zeta(s)$ at $s=1$. We estimate the contribution of the horizontal line segments using the Phragmén-Lindelöf Theorem, (1.5), and the functional equation of $\zeta(s)$, and then insert the definitions (1.1) and (2.2) to deduce that

$$\begin{aligned} \Delta_k(x) =& \frac{1}{2\pi i}\int_{\frac{1}{2}-\delta-iY}^{\frac{1}{2}-\delta+iY} \zeta^k(s)\frac{x^s}{s}\, ds + I_k(x;\delta,Y,T) \\ &+ O\left(x^{\varepsilon} + x^{1+\varepsilon}T^{-1+\varepsilon} + x^{\frac{1}{2}-\delta}T^{k(\delta+\lambda-2\lambda\delta)-1+\varepsilon}\right). \end{aligned}$$

We evaluate the integral on the right-hand side by deforming its contour of integration to the path consisting of line segments connecting the points $\frac{1}{2}-\delta-iY$, $-\varepsilon-iY$, $-\varepsilon+iY$, and $\frac{1}{2}-\delta+iY$, leaving a residue of size $O(1)$ from the pole of $1/s$ at $s=0$. We use the Phragmén-Lindelöf Theorem in the same way as before to bound the contribution of the horizontal line segments, and arrive at

$$\begin{aligned}\Delta_k(x) =&\frac{1}{2\pi i}\int_{-\varepsilon-iY}^{-\varepsilon+iY} \zeta^k(s)\frac{x^s}{s}\,ds + I_k(x;\delta,Y,T)\\ &+O\Big(x^\varepsilon Y^{\frac{k}{2}-1} + x^{\frac12-\delta}Y^{k(\delta+\lambda-2\lambda\delta)-1+\varepsilon} + x^{1+\varepsilon}T^{-1+\varepsilon} + x^{\frac12-\delta}T^{k(\delta+\lambda-2\lambda\delta)-1+\varepsilon}\Big).\end{aligned}$$

Lemma 3.1 now follows from this and Lemma 2.4 of Lester [19], which states that

$$\frac{1}{2\pi i}\int_{-\varepsilon-iY}^{-\varepsilon+iY} \zeta^k(s)\frac{x^s}{s}\,ds = Q_k(x;Y^k/(2\pi)^k) + O\big(x^\varepsilon Y^{\frac{k}{2}-1} + x^{1+\varepsilon}Y^{-\frac{k}{2}-\frac12}\big)$$

for $Y \le x$, where $Q_k$ is defined by (2.1). □

In their study of sign changes of $\Delta_2(x)$, Heath-Brown and Tsang [7] used binary expansions and the Cauchy-Schwarz inequality to bound $\sup_{0\le h\le H}|\Delta_2(x+h)-\Delta_2(x)|^2$. We portray their device as a general lemma here, because we will use it more than once in our proofs. The main point of this lemma is that it gives a bound for $|f(x+jb)-f(x)|^2$ that does not depend on $j$, which may be a function of $x$ in our applications.

**Lemma 3.2.** *Let $\ell$ be a nonnegative integer, $b\in\mathbb{R}$, and $j$ an integer such that $0\le j\le 2^\ell$. If $f:\mathbb{R}\to\mathbb{C}$ is a function and $x\in\mathbb{R}$, then*

$$|f(x+jb)-f(x)|^2 \le (\ell+1)\sum_{0\le\mu\le\ell}\sum_{0\le\nu<2^\mu}|f(x+(\nu+1)2^{\ell-\mu}b)-f(x+\nu 2^{\ell-\mu}b)|^2, \tag{3.2}$$

*where the indices of summation $\mu$ and $\nu$ run through integers.*

*Proof.* This is proved implicitly in [7], but we give a proof for completeness. Since $j$ is an integer with $0\le j\le 2^\ell$, it has a unique binary expansion

$$j=\sum_{\mu\in U}2^{\ell-\mu} \tag{3.3}$$

for some subset $U$ of $\{0,1,2,\dots,\ell\}$. We let

$$\nu_\mu=\sum_{\substack{m\in U\\ m<\mu}}2^{\mu-m} \tag{3.4}$$

for each $\mu\in U$, and write $f(x+jb)-f(x)$ as a telescoping sum

$$f(x+jb)-f(x)=\sum_{\mu\in U}\Big(f\big(x+(\nu_\mu+1)2^{\ell-\mu}b\big)-f\big(x+\nu_\mu 2^{\ell-\mu}b\big)\Big).$$

It follows from this and the Cauchy-Schwarz inequality that

$$|f(x+jb)-f(x)|^2 \le (\ell+1)\sum_{\mu\in U}\Big|f\big(x+(\nu_\mu+1)2^{\ell-\mu}b\big)-f\big(x+\nu_\mu 2^{\ell-\mu}b\big)\Big|^2. \tag{3.5}$$

Note that if $0\in U$, then $U=\{0\}$ by (3.3) and the assumption that $j\le 2^\ell$. In this case, $\nu_0=0$ by (3.4). On the other hand, if $0\notin U$, then $\nu_\mu<2^\mu$ by (3.4). In either case, it holds that $0\le\nu_\mu<2^\mu$ for all $\mu\in U$. Thus (3.2) follows from (3.5), because the right-hand side of (3.2) includes all possible values for $\nu_\mu$ and $\mu$. □

We will bound the fourth moment of $\Delta_k(x)$ by applying the Erdös-Turán inequality together with van der Corput's method for estimating exponential sums in a way similar to the proof of Lemma 4 of Tsang [27]. This technique is embodied in the following lemma.

**Lemma 3.3.** *Let $\|x\|$ denote the distance from $x$ to the nearest integer. If $\rho > 0$, $W \geq 1$, and $0 < \alpha \ll W^{1/k}$, then*

$$\#\big\{\mu \in \mathbb{Z} \ : \ W < \mu \leq 2W \text{ and } \big\|\big(\mu^{1/k} + \alpha\big)^k\big\| \leq \rho\big\} \ll_k W\rho + W^{\frac{2}{3} - \frac{1}{3k}}\alpha^{1/3} + W^{\frac{1}{2} + \frac{1}{2k}}\alpha^{-1/2},$$

*with the implied constant depending only on $k$.*

*Proof.* The Erdös-Turán inequality (see, for example, [21, Corollary 1.1]) implies that

$$\begin{aligned} &\#\big\{\mu \in \mathbb{Z} \ : \ W < \mu \leq 2W \text{ and } \big\|\big(\mu^{1/k} + \alpha\big)^k\big\| \leq \rho\big\} \\ &\leq 2W\rho + \frac{W}{L+1} + 3\sum_{\nu=1}^{L} \frac{1}{\nu}\bigg|\sum_{W<\mu\leq 2W} e\Big(\nu\big(\mu^{1/k} + \alpha\big)^k\Big)\bigg| \end{aligned} \tag{3.6}$$

for every positive integer $L$. To estimate the exponential sum, let

$$f(x) = \nu\big(x^{1/k} + \alpha\big)^k. \tag{3.7}$$

Then

$$f''(x) = -\Big(1 - \frac{1}{k}\Big)\nu\alpha\big(x^{1/k} + \alpha\big)^{k-2}x^{\frac{1}{k}-2}.$$

Thus, since $0 < \alpha \ll W^{1/k}$, there are positive constants $A_k$ and $B_k$ that depend only on $k$ such that

$$A_k\nu\alpha W^{-1-\frac{1}{k}} \leq -f''(x) \leq B_k\nu\alpha W^{-1-\frac{1}{k}}$$

whenever $W \leq x \leq 2W$. Hence, van der Corput's method [24, Theorem 5.9] implies

$$\sum_{W<\mu\leq 2W} e(f(\mu)) \ll_k \nu^{1/2}W^{\frac{1}{2}-\frac{1}{2k}}\alpha^{1/2} + \nu^{-1/2}W^{\frac{1}{2}+\frac{1}{2k}}\alpha^{-1/2}.$$

From this, the definition (3.7) of $f$, and (3.6), we arrive at

$$\begin{aligned} &\#\big\{\mu \in \mathbb{Z} \ : \ W < \mu \leq 2W \text{ and } \big\|\big(\mu^{1/k} + \alpha\big)^k\big\| \leq \rho\big\} \\ &\ll_k \ W\rho + \frac{W}{L} + L^{1/2}W^{\frac{1}{2}-\frac{1}{2k}}\alpha^{1/2} + W^{\frac{1}{2}+\frac{1}{2k}}\alpha^{-1/2}. \end{aligned}$$

To complete the proof of the lemma, we optimize this bound and choose $L$ to be the least integer that is greater than $W^{\frac{1}{3}+\frac{1}{3k}}\alpha^{-1/3}$. $\square$

## 4. Lester's method

The following is Lemma 2.6 of Lester [19], stated in a slightly more general form. Its proof is the same as that in [19].

**Lemma 4.1.** *Let $w : (0,\infty) \to \mathbb{R}$ be a fixed nonnegative smooth function of compact support. Suppose that $\delta$ is a fixed constant that satisfies $0 \leq \delta \leq 1/(2k)$ and (1.6). If $1 \leq Y \leq T \leq X$ and $I_k$ is defined by (2.2), then*

$$\frac{1}{X}\int_0^\infty \big|I_k(x;\delta,Y,T)\big|^2 w\Big(\frac{x}{X}\Big)\,dx \ll X^{1-2\delta+\varepsilon}Y^{2\delta k-1}.$$

We next use Lester's method to bound the mean square of the change in $I_k$ from $x$ to $x+h$.

**Lemma 4.2.** *Let $\delta$ be a fixed constant that satisfies $0 \le \delta \le 1/(2k)$ and (1.6). If $1 \le Y \le T \le X$, $0 \le h \le X$, and $I_k$ is defined by (2.2), then*

$$\frac{1}{X}\int_X^{2X} \big|I_k(x+h;\delta,Y,T) - I_k(x;\delta,Y,T)\big|^2\,dx \ll \min\big\{X^{1-2\delta+\varepsilon}Y^{2\delta k-1}, h^2X^{-1-2\delta+\varepsilon}T^{2\delta k+1}\big\}.$$

*Proof.* For brevity, we denote $I_k(u;\delta,Y,T)$ by $I(u)$ in this proof. The fundamental theorem of calculus and the Cauchy-Schwarz inequality imply

(4.1)
$$\frac{1}{X}\int_X^{2X} |I(x+h)-I(x)|^2\,dx = \frac{1}{X}\int_X^{2X}\left|\int_x^{x+h} I'(u)\,du\right|^2 dx \le \frac{h}{X}\int_X^{2X}\int_x^{x+h}|I'(u)|^2\,du\,dx.$$

Now, let $w:(0,\infty)\to\mathbb{R}$ be a nonnegative smooth function of compact support such that $w(u)=1$ whenever $1\le u\le 3$. Since $h\le X$, we see from interchanging the order of integration that

$$\frac{h}{X}\int_X^{2X}\int_x^{x+h}|I'(u)|^2\,du\,dx \le \frac{h^2}{X}\int_X^{2X+h}|I'(u)|^2\,du \le \frac{h^2}{X}\int_0^\infty |I'(u)|^2 w\Big(\frac{u}{X}\Big)\,du. \tag{4.2}$$

To bound the latter integral, observe that the definition (2.2) of $I$ and differentiation under the integral sign give

$$I'(u) := \mathrm{Re}\left\{\frac{1}{\pi i}\int_{\frac12-\delta+iY}^{\frac12-\delta+iT}\zeta^k(s)u^{s-1}\,ds\right\}.$$

We use this and the inequality $|\mathrm{Re}(z)|\le|z|$ to bound $|I'(u)|^2$, and then expand the square, interchange the order of integration, and make a change of variables to arrive at

$$\int_0^\infty |I'(u)|^2 w\Big(\frac{u}{X}\Big)\,du \le \frac{X^{-2\delta}}{\pi^2}\int_Y^T\int_Y^T \zeta^k(\tfrac12-\delta+it)\zeta^k(\tfrac12-\delta-iv)X^{i(t-v)}\mathcal{J}(t-v)\,dt\,dv,$$

where $\mathcal{J}(y) := \int_0^\infty u^{-1-2\delta+iy}w(u)\,du$. Repeated integration by parts shows that $\mathcal{J}(y)\ll_A \min\{1,|y|^{-A}\}$ for arbitrarily large $A>0$. Thus, given any $\eta>0$, we have $\mathcal{J}(t-v)\ll_{A,\eta} X^{-A}$ for $|t-v|>X^\eta$, and so the part of the above integral with $|t-v|>X^\eta$ is negligible. We bound the remaining part using the inequality $|ab|\ll|a|^2+|b|^2$, the functional equation of $\zeta(s)$, and Stirling's formula to deduce that

$$\begin{aligned}\int_0^\infty |I'(u)|^2 w\Big(\frac{u}{X}\Big)\,du &\ll X^{-2\delta}\iint\limits_{\substack{Y\le t,v\le T\\ |t-v|\le X^\eta}} |\zeta(\tfrac12-\delta+it)|^k|\zeta(\tfrac12-\delta-iv)|^k\,dt\,dv + O_A\big(X^{-A}\big)\\ &\ll X^{-2\delta+\eta}\int_Y^T t^{2\delta k}|\zeta(\tfrac12+\delta+it)|^{2k}\,dt + O_A\big(X^{-A}\big).\end{aligned}$$

This bound is $\ll X^{-2\delta+\eta}T^{2\delta k+1+\varepsilon}$ by (1.6) and integration by parts. From this, (4.2), and (4.1), we arrive at

$$\frac{1}{X}\int_X^{2X}|I(x+h)-I(x)|^2\,dx \ll h^2X^{-1-2\delta+\varepsilon}T^{2\delta k+1} \tag{4.3}$$

upon taking $\eta$ to be an arbitrarily small $\varepsilon>0$.

We have bounded the mean square of $I(x+h)-I(x)$ through the derivative $I'(x)$. Alternatively, we may use the inequality $|a+b|^2 \le 2|a|^2 + 2|b|^2$ and make a change of variables to write

$$\frac{1}{X}\int_X^{2X} |I(x+h)-I(x)|^2\,dx \le \frac{2}{X}\int_{X+h}^{2X+h} |I(x)|^2\,dx + \frac{2}{X}\int_X^{2X} |I(x)|^2\,dx.$$

Since $0 \le h \le X$, it follows that

$$\frac{1}{X}\int_X^{2X} |I(x+h)-I(x)|^2\,dx \ll \frac{1}{X}\int_0^{\infty} |I(x)|^2 w\Big(\frac{x}{X}\Big)\,dx,$$

where the function $w$ is as in (4.2). The right-hand side is $\ll X^{1-2\delta+\varepsilon}Y^{2\delta k-1}$ by Lemma 4.1. Combining this and (4.3), we arrive at the conclusion of Lemma 4.2. □

In proving Theorem 1.6, we will bound the fourth moment of $I_k(x;0,Y,T)$ by applying Lester's method together with the Riesz-Thorin Interpolation Theorem. For the rest of this section, let $1 \le Y \le T \le X$. Let $\Xi$ be the line segment from $\frac12 + iY$ to $\frac12 + iT$, and view $\Xi$ as a measure space in such a way that

$$\int_\Xi f = -i\int_{\frac12+iY}^{\frac12+iT} f(s)\,ds = \int_Y^T f(\tfrac12+it)\,dt$$

for all continuous functions $f : \Xi \to \mathbb{C}$. Define the operator $\mathcal{T}$ by

$$\mathcal{T}f(x) = \frac{1}{\pi i}\int_{\frac12+iY}^{\frac12+iT} f(s)x^s\,ds. \tag{4.4}$$

Note that if $f \in L^p(\Xi)$ for some $p \ge 1$, then Hölder's inequality implies that $\mathcal{T}f(x)$ exists for all $x > 0$, and that $\mathcal{T}f$ is continuous on $(0,\infty)$. Thus, if $f \in L^p(\Xi)$ for some $p \ge 1$, then $\mathcal{T}f \in L^q([X,2X])$ for all $q \ge 1$. In the next two lemmas, we use $\|f\|_p$ to denote the norm of $f$ in $L^p(\Xi)$, and we use $\|\mathcal{T}f\|_q$ to denote the norm of $\mathcal{T}f$ in $L^q([X,2X])$.

**Lemma 4.3.** *If $f \in L^2(\Xi)$, then $\|\mathcal{T}f\|_2 \ll X^{1+\varepsilon}\|f\|_2$. The implied constant here depends only on $\varepsilon$.*

*Proof.* Let $w : (0,\infty) \to \mathbb{R}$ be a nonnegative smooth function of compact support such that $w(u) = 1$ whenever $1 \le u \le 2$. Then

$$\int_X^{2X} |\mathcal{T}f(x)|^2\,dx \le \int_0^\infty |\mathcal{T}f(x)|^2 w\Big(\frac{x}{X}\Big)\,dx.$$

We replace $\mathcal{T}f(x)$ on the right-hand side by its definition (4.4), and then expand the square, apply Fubini's theorem, and make a change of variables to arrive at

$$\int_X^{2X} |\mathcal{T}f(x)|^2\,dx \le \frac{X^2}{\pi^2}\int_Y^T\int_Y^T f(\tfrac12+it)\overline{f(\tfrac12+iv)}X^{i(t-v)}\mathcal{J}(t-v)\,dv\,dt, \tag{4.5}$$

where $\mathcal{J}(y) := \int_0^\infty u^{1+iy}w(u)\,du$. Repeated integration by parts shows that $\mathcal{J}(y) \ll_A \min\{1, |y|^{-A}\}$ for arbitrarily large $A > 0$. From this and the inequality $|ab| \ll |a|^2 + |b|^2$, we deduce for any given $\eta > 0$ that

$$\begin{aligned}(4.6)\quad \int_{Y}^{T}\int_{\substack{Y\\|t-v|>X^\eta}}^{T} f(\tfrac12+it)\overline{f(\tfrac12+iv)}X^{i(t-v)}\mathcal{J}(t-v)\,dv\,dt &\ll_{A,\eta} \frac{1}{X^A}\int_Y^T\int_Y^T |f(\tfrac12+it)|^2\,dv\,dt\\ &= \frac{T-Y}{X^A}\|f\|_2^2.\end{aligned}$$

On the other hand, the bound $\mathcal{J}(y) \ll 1$ and the inequality $|ab| \ll |a|^2+|b|^2$ imply that

$$\int_{Y}^{T}\int_{\substack{Y\\|t-v|\le X^\eta}}^{T} f(\tfrac12+it)\overline{f(\tfrac12+iv)}X^{i(t-v)}\mathcal{J}(t-v)\,dv\,dt \ll \int_{Y}^{T}\int_{\substack{Y\\|t-v|\le X^\eta}}^{T} |f(\tfrac12+it)|^2\,dv\,dt \le X^\eta\|f\|_2^2.$$

From this, (4.6), (4.5), and the fact that $T-Y\le X$, we arrive at

$$\int_X^{2X} |\mathcal{T}f(x)|^2\,dx \ll X^{2+\eta}\|f\|_2^2.$$

Taking the square root of both sides, we finish the proof upon choosing $\eta$ to be an arbitrarily small $\varepsilon > 0$. □

**Lemma 4.4.** *If $f \in L^{4/3}(\Xi)$, then $\|\mathcal{T}f\|_4 \ll X^{\frac34+\varepsilon}\|f\|_{4/3}$. The implied constant here depends only on $\varepsilon$.*

*Proof.* By taking the absolute value of the integrand on the right-hand side of (4.4), we see that $\|\mathcal{T}f\|_\infty \ll X^{1/2}\|f\|_1$ for all $f\in L^1(\Xi)$. Lemma 4.3 states that $\|\mathcal{T}f\|_2 \ll X^{1+\varepsilon}\|f\|_2$ for all $f \in L^2(\Xi)$. It follows from these and the Riesz-Thorin Interpolation Theorem (see, for example, [23, p. 52]) that

$$\|\mathcal{T}f\|_4 \ll \big(X^{1/2}\big)^{1/2}\big(X^{1+\varepsilon}\big)^{1/2}\|f\|_{4/3} = X^{\frac34+\varepsilon}\|f\|_{4/3}$$

for all $f \in L^{4/3}(\Xi)$. □

**Lemma 4.5.** *Assume the Lindelöf Hypothesis. If $1 \le Y \le T \le X$ and $I_k$ is defined by* (2.2)*, then*

$$\int_X^{2X} \big|I_k(x;0,Y,T)\big|^4\,dx \ll \frac{X^{3+\varepsilon}}{Y}.$$

*Proof.* Let $f(s) = s^{-1}\zeta(s)^k$. Then the definitions (2.2) and (4.4) of $I_k$ and $\mathcal{T}$ imply that

$$I_k(x;0,Y,T) = \operatorname{Re}\big(\mathcal{T}f(x)\big).$$

From this, the inequality $|\operatorname{Re}(z)| \le |z|$, and Lemma 4.4, we arrive at

$$\int_X^{2X} \big|I_k(x;0,Y,T)\big|^4\,dx \ll X^{3+\varepsilon}\left(\int_Y^T \frac{|\zeta(\frac12+it)|^{4k/3}}{t^{4/3}}\,dt\right)^3.$$

The right-hand side is $\ll X^{3+\varepsilon}Y^{-1}$ if the Lindelöf Hypothesis is true. □

We remark that the Lindelöf Hypothesis is unnecessary for the case $k=3$ of Lemma 4.5 because the size of the fourth moment of $\zeta(s)$ is known [24, (7.6.2)]. This leads to an unconditional proof of Theorem 1.6 for $k=3$. However, as mentioned earlier, the better bound (1.9) has been found by Ivić [10].

The proofs of Lemmas 4.3, 4.4, and 4.5 also lead to a bound for the fourth moment of $I_k(x;\delta,Y,T)$. However, this more general bound does not improve Theorem 1.6, and the case $\delta=0$ is adequate for proving Theorem 1.6.

## 5. The mean square of $\sup_{0\le h\le H}|\Delta_k(x+h)-\Delta_k(x)|$

Our goal in this section is to prove Theorem 1.3 and Corollaries 1.4 and 1.5. Our first task is to estimate the second moment of $Q_k(x+h;V)-Q_k(x;V)$, where $Q_k$ is defined by (2.1).

**Lemma 5.1.** *If $X\ge 1$, $V\ge 0$, and $0\le h\le X$, then*

$$\frac{1}{X}\int_X^{2X}\big(Q_k(x+h;V)-Q_k(x;V)\big)^2\,dx \ll hV^{\varepsilon}+\frac{hV^{1-\frac{1}{k}+\varepsilon}}{X}.$$

*Proof.* The conclusion holds trivially for $0\le V<1$ since $Q_k(x;V)=0$ for $x\ge 1$ and $0\le V<1$. We may thus assume for the rest of the proof that $V\ge 1$. For brevity, define $\Sigma(u,v)$ by

$$\Sigma(u,v)=\Sigma(u,v;V):=\frac{1}{\pi\sqrt{k}}\sum_{n\le V/v}\frac{d_k(n)}{n^{\frac12+\frac{1}{2k}}}\cos\left(2\pi k(nu)^{1/k}+\frac{(k-3)\pi}{4}\right), \tag{5.1}$$

so that $Q_k(x;V)=x^{\frac12-\frac{1}{2k}}\Sigma(x,x)$ by (2.1). We use this and the inequality $|a+b|^2\ll|a|^2+|b|^2$ to write

$$\frac{1}{X}\int_X^{2X}\big(Q_k(x+h;V)-Q_k(x;V)\big)^2\,dx\ll \mathcal{I}_1+\mathcal{I}_2+\mathcal{I}_3, \tag{5.2}$$

where

$$\mathcal{I}_1:=\frac{1}{X}\int_X^{2X}(x+h)^{1-\frac1k}\big|\Sigma(x+h,x+h)-\Sigma(x+h,x)\big|^2\,dx, \tag{5.3}$$

$$\mathcal{I}_2:=\frac{1}{X}\int_X^{2X}\left|(x+h)^{\frac12-\frac{1}{2k}}-x^{\frac12-\frac{1}{2k}}\right|^2|\Sigma(x+h,x)|^2\,dx, \tag{5.4}$$

and

$$\mathcal{I}_3:=\frac{1}{X}\int_X^{2X}x^{1-\frac1k}\big|\Sigma(x+h,x)-\Sigma(x,x)\big|^2\,dx. \tag{5.5}$$

We first bound $\mathcal{I}_1$. The definition (5.1) of $\Sigma(u,v)$ implies that

$$\Sigma(x+h,x+h)-\Sigma(x+h,x)=-\frac{1}{\pi\sqrt{k}}\sum_{V/(x+h)<n\le V/x}\frac{d_k(n)}{n^{\frac12+\frac{1}{2k}}}\cos\left(2\pi kn^{1/k}(x+h)^{1/k}+\frac{(k-3)\pi}{4}\right)$$

We insert this in (5.3), expand the square, and interchange the order of summation to deduce that

$$\begin{aligned}\mathcal{I}_1=&\frac{1}{\pi^2kX}\sum_{V/(2X+h)<m,n\le V/X}\frac{d_k(m)d_k(n)}{(mn)^{\frac12+\frac{1}{2k}}}\int_{A(m,n)}^{B(m,n)}(x+h)^{1-\frac1k}\\&\times\cos\left(2\pi km^{1/k}(x+h)^{1/k}+\frac{(k-3)\pi}{4}\right)\cos\left(2\pi kn^{1/k}(x+h)^{1/k}+\frac{(k-3)\pi}{4}\right)dx,\end{aligned}$$

where $A(m,n) = \max\{X, (V/m) - h, (V/n) - h\}$ and $B(m,n) = \min\{2X, V/m, V/n\}$. Since $B(m,n) - A(m,n) \le h$, $(x+h)^{1-\frac{1}{k}} \asymp X^{1-\frac{1}{k}}$ for $x \asymp X$, and $|\cos y| \le 1$ for all real $y$, it follows that

$$\mathcal{I}_1 \ll hX^{-1/k} \sum_{m,n \le V/X} \frac{d_k(m)d_k(n)}{(mn)^{\frac{1}{2}+\frac{1}{2k}}} \ll hX^{-1/k} \left(\frac{V}{X}\right)^{1-\frac{1}{k}+\varepsilon} \le \frac{hV^{1-\frac{1}{k}+\varepsilon}}{X}. \tag{5.6}$$

We next bound $\mathcal{I}_2$. The mean value theorem of differential calculus implies that

$$(x+h)^{\frac{1}{2}-\frac{1}{2k}} - x^{\frac{1}{2}-\frac{1}{2k}} \ll hx^{-\frac{1}{2}-\frac{1}{2k}}. \tag{5.7}$$

We insert the definition (5.1) of $\Sigma(x+h,x)$ into (5.4), and then use the identity $\cos(2\pi y) = (e(y)+e(-y))/2$, the inequality $|z+\overline{z}|^2 \ll |z|^2$, and (5.7) to write

$$\mathcal{I}_2 \ll \frac{h^2}{X^{2+\frac{1}{k}}} \int_X^{2X} \left| \sum_{n \le V/x} \frac{d_k(n)}{n^{\frac{1}{2}+\frac{1}{2k}}} e\bigg(kn^{1/k}(x+h)^{1/k} + \frac{k-3}{8}\bigg) \right|^2 dx.$$

We expand the square and interchange the order of summation to arrive at

$$\mathcal{I}_2 \ll \frac{h^2}{X^{2+\frac{1}{k}}} \sum_{m,n \le V/X} \frac{d_k(m)d_k(n)}{(mn)^{\frac{1}{2}+\frac{1}{2k}}} \int_X^{B(m,n)} e\Big(k(x+h)^{1/k}\big(m^{1/k} - n^{1/k}\big)\Big)\, dx,$$

where $B(m,n) = \min\{2X, V/m, V/n\}$. Now split the above sum to write

$$\mathcal{I}_2 \ll \mathcal{I}_{2,D} + \mathcal{I}_{2,O}, \tag{5.8}$$

where $\mathcal{I}_{2,D}$ is the sum of the "diagonal" terms with $m = n$ and $\mathcal{I}_{2,O}$ is the sum of the "off-diagonal" terms with $m \ne n$. Since $B(m,n) - X \le X$, we have

$$\mathcal{I}_{2,D} \le \frac{h^2}{X^{1+\frac{1}{k}}} \sum_{n \le V/X} \frac{d_k^2(n)}{n^{1+\frac{1}{k}}} \ll \frac{h^2}{X^{1+\frac{1}{k}}}. \tag{5.9}$$

To bound $\mathcal{I}_{2,O}$, it suffices to bound the sum of the terms with $m > n$, as the terms with $m < n$ are their complex conjugates. If $m > n$, then integration by parts gives

$$\int_X^{B(m,n)} e\Big(k(x+h)^{1/k}\big(m^{1/k} - n^{1/k}\big)\Big)\, dx \ll \frac{X^{1-\frac{1}{k}}}{m^{1/k} - n^{1/k}} \ll \frac{X^{1-\frac{1}{k}} m^{1-\frac{1}{k}}}{m-n}.$$

Thus, using $m - n \gg m$ for $n \le m/2$ and $n^{\frac{1}{2}+\frac{1}{2k}} \asymp m^{\frac{1}{2}+\frac{1}{2k}}$ for $m/2 < n < m$, we deduce that

$$\mathcal{I}_{2,O} \ll \frac{h^2}{X^{1+\frac{2}{k}}} \sum_{n < m \le V/X} \frac{m^{\frac{1}{2}-\frac{3}{2k}+\varepsilon}}{n^{\frac{1}{2}+\frac{1}{2k}}(m-n)} \ll \frac{h^2}{X^{1+\frac{2}{k}}} \sum_{m \le V/X} m^{\frac{1}{2}-\frac{3}{2k}+\varepsilon} \left(\frac{\log m}{m^{\frac{1}{2}+\frac{1}{2k}}}\right) \ll \frac{h^2 V^{1-\frac{2}{k}+\varepsilon}}{X^2}.$$

From this, (5.9), and (5.8), we arrive at

$$\mathcal{I}_2 \ll \frac{h^2}{X^{1+\frac{1}{k}}} + \frac{h^2 V^{1-\frac{2}{k}+\varepsilon}}{X^2} \tag{5.10}$$

Having estimated $\mathcal{I}_1$ and $\mathcal{I}_2$, we now turn to $\mathcal{I}_3$. The definitions (5.1) and (5.5) together with the identity $\cos(A+B)-\cos(A-B)=-2\sin A\sin B$ imply

$$\mathcal{I}_3=\frac{4}{\pi^2kX}\int_X^{2X}x^{1-\frac{1}{k}}\Bigg|\sum_{n\le V/x}\frac{d_k(n)}{n^{\frac12+\frac{1}{2k}}}\sin\left(\pi kn^{1/k}g(x;h)\right)\times\sin\left(2\pi k(nx)^{1/k}+\pi kn^{1/k}g(x;h)+\frac{(k-3)\pi}{4}\right)\Bigg|^2\,dx,$$

where $g(x;h)$ is defined by

$$g(x;h):=(x+h)^{1/k}-x^{1/k}. \tag{5.11}$$

We write the second sine in terms of exponential functions via $\sin(2\pi y)=i(e(-y)-e(y))/2$ and then apply the inequalities $|z-\overline{z}|^2\ll|z|^2$ and $x^{1-\frac1k}\asymp X^{1-\frac1k}$ to deduce that

$$\mathcal{I}_3\ll X^{-1/k}\int_X^{2X}\Bigg|\sum_{n\le V/x}\frac{d_k(n)}{n^{\frac12+\frac{1}{2k}}}\sin\left(\pi kn^{1/k}g(x;h)\right)\times e\left(k(nx)^{1/k}+\frac12kn^{1/k}g(x;h)+\frac{k-3}{8}\right)\Bigg|^2\,dx.$$

We expand the square and interchange the order of summation to arrive at

$$\mathcal{I}_3\ll X^{-1/k}\sum_{m,n\le V/X}\frac{d_k(m)d_k(n)}{(mn)^{\frac12+\frac{1}{2k}}}\int_X^{B(m,n)}\sin\left(\pi km^{1/k}g(x;h)\right)\sin\left(\pi kn^{1/k}g(x;h)\right)\times e\left(k\left(m^{1/k}-n^{1/k}\right)\left(x^{1/k}+\frac12g(x;h)\right)\right)dx,$$

where $B(m,n)=\min\{2X,V/m,V/n\}$. Split the above sum to write

$$\mathcal{I}_3\ll\mathcal{I}_{3,D}+\mathcal{I}_{3,O}, \tag{5.12}$$

where $\mathcal{I}_{3,D}$ is the sum of the diagonal terms with $m=n$ and $\mathcal{I}_{3,O}$ is the sum of the off-diagonal terms with $m\neq n$.

To bound $\mathcal{I}_{3,D}$, observe first that the definition (5.11) of $g(x;h)$ and the mean value theorem of differential calculus implies that

$$g(x;h)\ll hx^{\frac1k-1}. \tag{5.13}$$

Thus, since $B(m,n)\le 2X$ and $\sin^2 y\le|y|$ for all real $y$, we have

$$\begin{aligned}\mathcal{I}_{3,D}&\ll X^{-1/k}\sum_{n\le V/X}\frac{d_k^2(n)}{n^{1+\frac1k}}\int_X^{B(n,n)}\sin^2\left(\pi kn^{1/k}g(x;h)\right)dx\\&\ll X^{-1/k}\sum_{n\le V/X}\frac{d_k^2(n)}{n}\int_X^{B(n,n)}hx^{\frac1k-1}\,dx\ \ll hV^{\varepsilon}.\end{aligned} \tag{5.14}$$

Our next task is to bound the sum $\mathcal{I}_{3,O}$ in (5.12). Observe that the terms in $\mathcal{I}_{3,O}$ that have $m < n$ are complex conjugates of those with $m > n$. Thus

$$\mathcal{I}_{3,O} \ll X^{-1/k} \sum_{n<m\le V/X} \frac{d_k(m)d_k(n)}{(mn)^{\frac{1}{2}+\frac{1}{2k}}} \left| \int_X^{B(m,n)} \sigma_m(x)\sigma_n(x)e^{iG(x)}\,dx \right|, \tag{5.15}$$

where $\sigma_j(x) = \sigma_j(x;h)$ is defined by

$$\sigma_j(x) := \sin\left(\pi k j^{1/k} g(x;h)\right) \tag{5.16}$$

for all positive integers $j$ and $G(x) = G(x,m,n;h)$ is defined by

$$G(x) := 2\pi k\big(m^{1/k} - n^{1/k}\big)\left(x^{1/k} + \frac{1}{2}g(x;h)\right). \tag{5.17}$$

To bound the integral on the right-hand side of (5.15), we will first estimate $\sigma_j(x)$, $G(x)$, and some of their derivatives. The definition (5.11) of $g(x;h)$ and the mean value theorem of differential calculus imply

$$\left|\frac{d}{dx}g(x;h)\right| = \frac{1}{k}\left|(x+h)^{\frac{1}{k}-1} - x^{\frac{1}{k}-1}\right| \le \frac{1}{k}hx^{\frac{1}{k}-2} \tag{5.18}$$

and

$$\frac{d^2}{dx^2}g(x;h) = \frac{1}{k}\left(\frac{1}{k}-1\right)\left((x+h)^{\frac{1}{k}-2} - x^{\frac{1}{k}-2}\right) \ll hx^{\frac{1}{k}-3}. \tag{5.19}$$

By (5.16), the fact that $|\sin y| \le |y|$ for all real $y$, and (5.13), we have

$$\sigma_j(x) \ll j^{1/k}hx^{\frac{1}{k}-1}, \tag{5.20}$$

while (5.16), the fact that $|\cos y| \le 1$ for all real $y$, and (5.18) imply

$$\sigma_j'(x) = \pi k j^{1/k}\cos\left(\pi k j^{1/k}g(x;h)\right)\frac{d}{dx}g(x;h) \ll j^{1/k}hx^{\frac{1}{k}-2}. \tag{5.21}$$

It also follows from (5.16) that

$$|\sigma_j(x)| \le 1 \tag{5.22}$$

because $|\sin y| \le 1$ for all real $y$. If $X \le x \le 2X$, then (5.18) implies that $|(d/dx)g(x;h)| \le (1/k)x^{\frac{1}{k}-1}$ by our assumption that $0 \le h \le X$. It follows from this, (5.17), and the triangle inequality that

$$|G'(x)| = \left|2\pi k\big(m^{1/k} - n^{1/k}\big)\left(\frac{1}{k}x^{\frac{1}{k}-1} + \frac{1}{2}\frac{d}{dx}g(x;h)\right)\right| \gg \big(m^{1/k} - n^{1/k}\big)x^{\frac{1}{k}-1} \tag{5.23}$$

for $m > n$ and $X \le x \le 2X$. Furthermore, from (5.17) and (5.19), we deduce for $m > n$ and $X \le x \le 2X$ that

$$G''(x) = 2\pi k\big(m^{1/k} - n^{1/k}\big)\left(\left(\frac{1}{k^2} - \frac{1}{k}\right)x^{\frac{1}{k}-2} + \frac{1}{2}\frac{d^2}{dx^2}g(x;h)\right) \ll \big(m^{1/k} - n^{1/k}\big)x^{\frac{1}{k}-2} \tag{5.24}$$

since $0 \le h \le X \le x$.

Having listed estimates for $\sigma_j(x)$, $G(x)$, and their derivatives, we now bound the integral in (5.15). We integrate by parts to write

$$\begin{aligned}\int_X^{B(m,n)} &\sigma_m(x)\sigma_n(x)e^{iG(x)}\,dx \\ &= \frac{\sigma_m(x)\sigma_n(x)e^{iG(x)}}{iG'(x)}\bigg|_{x=X}^{x=B(m,n)} - \int_X^{B(m,n)} \frac{\sigma_m'(x)\sigma_n(x)e^{iG(x)}}{iG'(x)}\,dx \\ &\quad - \int_X^{B(m,n)} \frac{\sigma_m(x)\sigma_n'(x)e^{iG(x)}}{iG'(x)}\,dx + \int_X^{B(m,n)} \frac{\sigma_m(x)\sigma_n(x)G''(x)e^{iG(x)}}{i(G'(x))^2}\,dx \\ &= J_1 + J_2 + J_3 + J_4,\end{aligned}$$

say, where we recall that $B(m,n) = \min\{2X, V/m, V/n\}$. For each of $J_1, J_2, J_3, J_4$, we apply (5.23) to bound the factor $1/G'(x)$. We also use the fact that $|e^{iG(x)}| = 1$. In $J_1$, we bound the factor $\sigma_m(x)$ via (5.20), while we bound $\sigma_n(x)$ using (5.22). For $J_2$, we use (5.21) to estimate $\sigma_m'(x)$ and (5.22) to estimate $\sigma_n(x)$. We similarly apply (5.21) and (5.22) to bound $J_3$. Finally, to estimate $J_4$, we use (5.20) on $\sigma_m(x)$, (5.22) on $\sigma_n(x)$, and (5.24) on $G''(x)$. The result of these estimations is

$$\int_X^{B(m,n)} \sigma_m(x)\sigma_n(x)e^{iG(x)}\,dx \ll \frac{m^{1/k}h}{m^{1/k}-n^{1/k}}$$

for $m > n$. From this, (5.15), and the fact that $m^{1/k} - n^{1/k} \gg (m-n)m^{\frac{1}{k}-1}$ for $m > n$, we arrive at

$$\mathcal{I}_{3,O} \ll hX^{-1/k} \sum_{n<m\le V/X} \frac{d_k(m)d_k(n)m^{\frac12-\frac{1}{2k}}}{n^{\frac12+\frac{1}{2k}}(m-n)}.$$

Thus, using $m-n \gg m$ for $n \le m/2$ and $n^{\frac12+\frac{1}{2k}} \asymp m^{\frac12+\frac{1}{2k}}$ for $m/2 < n < m$, we deduce that

$$\mathcal{I}_{3,O} \ll hX^{-1/k} \sum_{n<m\le V/X} \frac{m^{\frac12-\frac{1}{2k}+\varepsilon}}{n^{\frac12+\frac{1}{2k}}(m-n)} \ll hX^{-1/k} \sum_{m\le V/X} m^{\frac12-\frac{1}{2k}+\varepsilon}\left(\frac{\log m}{m^{\frac12+\frac{1}{2k}}}\right) \ll \frac{hV^{1-\frac1k+\varepsilon}}{X}.$$

From this, (5.14), and (5.12), we arrive at

$$\mathcal{I}_3 \ll hV^{\varepsilon} + \frac{hV^{1-\frac1k+\varepsilon}}{X}. \tag{5.25}$$

It now follows from (5.2), (5.6), (5.10), and (5.25) that

$$\frac{1}{X}\int_X^{2X} \big(Q_k(x+h;V) - Q_k(x;V)\big)^2\,dx \ll hV^{\varepsilon} + \frac{hV^{1-\frac1k+\varepsilon}}{X} + \frac{h^2}{X^{1+\frac1k}} + \frac{h^2V^{1-\frac2k+\varepsilon}}{X^2}.$$

The latter two terms are bounded by the sum of the first two terms on the right-hand side because $h \le X$ and $X, V \ge 1$. This completes the proof of Lemma 5.1. $\square$

To prove Theorem 1.3, we will estimate $\Delta_k(x)$ via Lemma 3.1 and then apply Lemmas 4.2 and 5.1. Before we can do so, we need to approximate $\sup_{0\le h\le H}|\Delta_k(x+h) - \Delta_k(x)|^2$ by a quantity that allows us to apply Lemma 3.2. To this end, suppose that $1 \le H \le X$ and $X \le x \le 2X$. We write $H$ as

$$H = 2^{\ell}b \tag{5.26}$$

for some unique $\ell, b$ such that $\ell$ is a nonnegative integer and $1 \le b < 2$. The definition (1.1) of $\Delta_k(x)$ implies that

$$\Delta_k(x) = \sum_{n \le x} d_k(n) - xP_k(\log x) \tag{5.27}$$

for some polynomial $P_k$ of degree $k-1$. Thus $\Delta_k(x)$ is continuous except at points $x = n$ with $n$ an integer, where it is continuous from the right and has left-hand limit $\Delta_k(n) - d_k(n)$. It follows that there is an $h_0 \in [0, H]$ such that either

$$\sup_{0 \le h \le H} |\Delta_k(x+h) - \Delta_k(x)|^2 = |\Delta_k(x+h_0) - \Delta_k(x)|^2 \tag{5.28}$$

or

$$\sup_{0 \le h \le H} |\Delta_k(x+h) - \Delta_k(x)|^2 = |\Delta_k(x+h_0) - d_k(x+h_0) - \Delta_k(x)|^2. \tag{5.29}$$

Suppose first that (5.28) holds. By (5.26) and the fact that $0 \le h_0 \le H$, we have

$$jb \le h_0 \le (j+1)b \tag{5.30}$$

for some integer $j$ satisfying $0 \le j \le 2^\ell - 1$. The expression (5.27) and the mean value theorem of differential calculus imply that

$$\Delta_k(u_2) - \Delta_k(u_1) = \sum_{u_1 < n \le u_2} d_k(n) + O((u_2 - u_1)\log^k(X+2))$$

for $1 \le u_1 \le u_2 \ll X$. Since $d_k(n) \ge 0$ for all $n$, it follows that

$$\Delta_k(u_2) \ge \Delta_k(u_1) - O((u_2 - u_1)\log^k(X+2)) \tag{5.31}$$

for $1 \le u_1 \le u_2 \ll X$. If $\Delta_k(x+h_0) \ge \Delta_k(x)$, then (5.30) and (5.31) give

$$0 \le \Delta_k(x+h_0) - \Delta_k(x) \le \Delta_k(x+(j+1)b) - \Delta_k(x) + O(b\log^k(X+2)),$$

while if $\Delta_k(x+h_0) \le \Delta_k(x)$, then (5.30) and (5.31) imply

$$0 \ge \Delta_k(x+h_0) - \Delta_k(x) \ge \Delta_k(x+jb) - \Delta_k(x) - O(b\log^k(X+2)).$$

In either case, we have

$$|\Delta_k(x+h_0) - \Delta_k(x)| \le \max_{0 \le j \le 2^\ell} |\Delta_k(x+jb) - \Delta_k(x)| + O(\log^k(X+2)).$$

From this and (5.28), we arrive at

$$\sup_{0 \le h \le H} |\Delta_k(x+h) - \Delta_k(x)|^2 \ll \max_{0 \le j \le 2^\ell} |\Delta_k(x+jb) - \Delta_k(x)|^2 + O(X^\varepsilon). \tag{5.32}$$

We have shown that if (5.28) holds, then (5.32) is true. Now suppose that (5.29) holds and $x + h_0$ is a positive integer. Then

$$\Delta_k(x+h_0) - d_k(x+h_0) - \Delta_k(x) < 0$$

since otherwise $|\Delta_k(x+h_0) - \Delta_k(x)| > |\Delta_k(x+h_0) - d_k(x+h_0) - \Delta_k(x)|$, which contradicts (5.29). Hence (5.30) and (5.31) imply

$$0 > \Delta_k(x+h_0) - d_k(x+h_0) - \Delta_k(x) \ge \Delta_k(x+jb) - d_k(x+h_0) - \Delta_k(x) - O(b\log^k(X+2)),$$

and (5.32) again follows because $d_k(x+h_0) \ll X^{\varepsilon}$. We have thus proved that (5.32) holds in either case. Consequently, for each $x$ with $X \le x \le 2X$, there is an integer $j_0 = j_0(x)$ such that

$$0 \le j_0 \le 2^{\ell} \tag{5.33}$$

and

$$\sup_{0\le h\le H} |\Delta_k(x+h) - \Delta_k(x)|^2 \ll |\Delta_k(x+j_0 b) - \Delta_k(x)|^2 + O(X^{\varepsilon}). \tag{5.34}$$

We now apply Lemma 3.1 to estimate the right-hand side. Let $\delta \in [0, 1/(2k)]$ be a fixed constant satisfying (1.6), and let $\lambda$ be defined by (1.5). Using the inequality $|a+b|^2 \ll |a|^2 + |b|^2$, we deduce from (5.34) and Lemma 3.1 that

$$\begin{aligned} &\sup_{0\le h\le H} |\Delta_k(x+h) - \Delta_k(x)|^2 \\ &\ll |Q_k(x+j_0 b) - Q_k(x)|^2 + |I_k(x+j_0 b) - I_k(x)|^2 + |E_k(x+j_0 b) - E_k(x)|^2 + O(X^{\varepsilon}) \end{aligned} \tag{5.35}$$

for $1 \le Y \le T \le X$, where we use the abbreviations $Q_k(x) = Q_k(x; Y^k/(2\pi)^k)$, $I_k(x) = I_k(x;\delta, Y, T)$, and $E_k(x) = E_k(x;\delta, Y, T)$. We integrate both sides of (5.35) from $x = X$ to $x = 2X$, divide both sides by $X$, and then apply Lemma 3.2 to arrive at

$$\begin{aligned} &\frac{1}{X}\int_X^{2X} \sup_{0\le h\le H} |\Delta_k(x+h) - \Delta_k(x)|^2\,dx \\ &\ll \frac{\ell+1}{X}\sum_{0\le\mu\le\ell}\sum_{0\le\nu<2^{\mu}} \int_X^{2X} |Q_k(x+(\nu+1)2^{\ell-\mu}b) - Q_k(x+\nu 2^{\ell-\mu}b)|^2\,dx \\ &\quad + \frac{\ell+1}{X}\sum_{0\le\mu\le\ell}\sum_{0\le\nu<2^{\mu}} \int_X^{2X} |I_k(x+(\nu+1)2^{\ell-\mu}b) - I_k(x+\nu 2^{\ell-\mu}b)|^2\,dx \\ &\quad + \frac{1}{X}\int_X^{2X} |E_k(x+j_0 b) - E_k(x)|^2\,dx + O(X^{\varepsilon}) \\ &= \mathcal{J}_Q + \mathcal{J}_I + \mathcal{J}_E + O(X^{\varepsilon}), \end{aligned} \tag{5.36}$$

say. Now a change of variables and Lemma 5.1 with $V = (Y/2\pi)^k$ implies

$$\begin{aligned} \mathcal{J}_Q &= \frac{\ell+1}{X}\sum_{0\le\mu\le\ell}\sum_{0\le\nu<2^{\mu}} \int_{X+\nu 2^{\ell-\mu}b}^{2X+\nu 2^{\ell-\mu}b} |Q_k(x+2^{\ell-\mu}b) - Q_k(x)|^2\,dx \\ &\ll (\ell+1)\sum_{0\le\mu\le\ell}\sum_{0\le\nu<2^{\mu}} 2^{\ell-\mu}b\Big(Y^{\varepsilon} + \frac{Y^{k-1+\varepsilon}}{X}\Big) \ll \Big(1+\frac{Y^{k-1}}{X}\Big) HY^{\varepsilon}\log^2(H+2) \end{aligned} \tag{5.37}$$

by (5.26).

We next estimate the sum $\mathcal{J}_I$ in (5.36). A change of variables gives

$$\mathcal{J}_I = \frac{\ell+1}{X}\sum_{0\le\mu\le\ell}\sum_{0\le\nu<2^{\mu}} \int_{X+\nu 2^{\ell-\mu}b}^{2X+\nu 2^{\ell-\mu}b} |I_k(x+2^{\ell-\mu}b) - I_k(x)|^2\,dx. \tag{5.38}$$

Let $B = HX^{-1}T^{\frac{1}{2}+\delta k}Y^{\frac{1}{2}-\delta k}$. We split the $\mu$-sum on the right-hand side of (5.38) according to whether $2^\mu \le B$ or $2^\mu > B$, and deduce from Lemma 4.2 that

$$\begin{aligned}\mathcal{J}_I &\ll (\ell+1)\sum_{\substack{0\le\mu\le\ell\\ 2^\mu\le B}}\sum_{0\le\nu<2^\mu} X^{1-2\delta+\varepsilon}Y^{2\delta k-1} + (\ell+1)\sum_{\substack{0\le\mu\le\ell\\ 2^\mu> B}}\sum_{0\le\nu<2^\mu}\left(2^{\ell-\mu}b\right)^2 X^{-1-2\delta+\varepsilon}T^{2\delta k+1}\\ &\ll (\ell+1)BX^{1-2\delta+\varepsilon}Y^{2\delta k-1} + (\ell+1)B^{-1}\left(2^\ell b\right)^2 X^{-1-2\delta+\varepsilon}T^{2\delta k+1}.\end{aligned}$$

From this, (5.26), and the definition $B = HX^{-1}T^{\frac{1}{2}+\delta k}Y^{\frac{1}{2}-\delta k}$, we arrive at

$$\mathcal{J}_I \ll HX^{-2\delta+\varepsilon}T^{\frac{1}{2}+\delta k}Y^{\delta k-\frac{1}{2}}. \tag{5.39}$$

Finally, to estimate the integral $\mathcal{J}_E$ in (5.36), we bound its integrand pointwise via (3.1) and the inequality $|a+b|^2 \ll |a|^2+|b|^2$. In bounding $E_k(x+j_0b)$, we make use of the fact that $x+j_0b \ll X$ by (5.33), (5.26), and our assumption that $H \le X$. We arrive at

$$\mathcal{J}_E \ll \frac{X^{2+\varepsilon}}{Y^{k+1}} + X^\varepsilon Y^{k-2} + \frac{X^{2+\varepsilon}}{T^2} + X^{1-2\delta}Y^{2k(\delta+\lambda-2\lambda\delta)-2+\varepsilon} + X^{1-2\delta}T^{2k(\delta+\lambda-2\lambda\delta)-2+\varepsilon}.$$

Theorem 1.3 now follows from this, (5.36), (5.37), and (5.39).

We next prove Corollary 1.4. If $k=3$, then a theorem of Heath-Brown [6] implies that $\delta = 1/12$ satisfies (1.6). Also, Weyl's bound [24, Theorem 5.5] implies that $\lambda \le 1/6$. Corollary 1.4 follows from this and Theorem 1.3 with $k=3$, $\delta=1/12$, $T = X^{\frac{2}{3}+\varepsilon}$, and $Y = X^{\frac{16}{27}-\varepsilon}$.

To prove Corollary 1.5, we observe that the Lindelöf Hypothesis implies that $\lambda = 0$ and that $\delta = 0$ satisfies (1.6). Thus, Corollary 1.5 for the case $k=3$ follows from Theorem 1.3 with $\delta = 0$, $T = X^{\frac{2}{3}+\varepsilon}$, and $Y = X^{\frac{8}{15}-\varepsilon}$, while Corollary 1.5 for the case $k \ge 4$ follows from Theorem 1.3 with $\delta=0$, $T = X^{\frac{1}{2}+\frac{1}{2k}+\varepsilon}$, and $Y = X^{\frac{k-1}{k(k-2)}-\varepsilon}$.

## 6. The fourth moment of $\Delta_k(x)$

In this section, we shall prove Theorem 1.6. Suppose that $1 \le Y \le T \le X$. We apply Lemma 3.1 with $\delta = 0$ and use the inequality $|a+b|^4 \ll |a|^4+|b|^4$ to write

$$\begin{aligned}&\frac{1}{X}\int_X^{2X}|\Delta_k(x)|^4\,dx\\ &\ll \frac{1}{X}\int_X^{2X}|Q_k(x;Y^k/(2\pi)^k)|^4\,dx + \frac{1}{X}\int_X^{2X}|I_k(x;0,Y,T)|^4\,dx + \frac{1}{X}\int_X^{2X}|E_k(x;0,Y,T)|^4\,dx.\end{aligned}$$

From this, (3.1) with $\delta=\lambda=0$, and Lemma 4.5, we deduce that

$$\begin{aligned}\frac{1}{X}\int_X^{2X}|\Delta_k(x)|^4\,dx \ll \frac{1}{X}\int_X^{2X}|Q_k(x;Y^k/(2\pi)^k)|^4\,dx + \frac{X^{2+\varepsilon}}{Y}\\ + \frac{X^{4+\varepsilon}}{Y^{2k+2}} + X^\varepsilon Y^{2k-4} + \frac{X^{2+\varepsilon}}{Y^4} + \frac{X^{4+\varepsilon}}{T^4}\end{aligned} \tag{6.1}$$

under the assumption of the Lindelöf Hypothesis. To prove Theorem 1.6, our main task in this section is to bound the first term on the right-hand side. For brevity, in this section we set

$$V := \left(\frac{Y}{2\pi}\right)^k, \tag{6.2}$$

$$a_1 = a_1(\mu,\nu,m,n;k) := \frac{d_k(\mu)d_k(\nu)d_k(m)d_k(n)}{(\mu\nu mn)^{\frac12+\frac{1}{2k}}}, \tag{6.3}$$

and

$$X_1 = X_1(\mu,\nu,m,n;V,X) := \min\{2X, V/\mu, V/\nu, V/m, V/n\} \le 2X. \tag{6.4}$$

Use the definition (2.1) of $Q_k$, interchange the order of summation, and repeatedly apply the trigonometric identity $2\cos a\cos b = \cos(a+b)+\cos(a-b)$ to write

$$\frac{1}{X}\int_X^{2X} |Q_k(x;V)|^4\,dx = \frac{1}{\pi^4k^2}\left(\frac38 S_1 + \frac12 S_2 + \frac18 S_3\right), \tag{6.5}$$

where $S_1$, $S_2$, and $S_3$ are defined by

$$S_1 := \frac1X \sum_{\mu,\nu,m,n\le V/X} a_1 \int_X^{X_1} x^{2-\frac2k}\cos\Big(2\pi k x^{1/k}\big(\mu^{1/k}+\nu^{1/k}-m^{1/k}-n^{1/k}\big)\Big)\,dx, \tag{6.6}$$

$$S_2 := \frac1X \sum_{\mu,\nu,m,n\le V/X} a_1 \int_X^{X_1} x^{2-\frac2k}\cos\left(2\pi k x^{1/k}\big(\mu^{1/k}+\nu^{1/k}+m^{1/k}-n^{1/k}\big)+\frac{(k-3)\pi}{2}\right)dx, \tag{6.7}$$

and

$$S_3 := \frac1X \sum_{\mu,\nu,m,n\le V/X} a_1 \int_X^{X_1} x^{2-\frac2k}\cos\Big(2\pi k x^{1/k}\big(\mu^{1/k}+\nu^{1/k}+m^{1/k}+n^{1/k}\big)+(k-3)\pi\Big)\,dx. \tag{6.8}$$

Our first task is to estimate $S_1$, which is defined by (6.6). We bound the right-hand side of (6.6) by taking the absolute value of each term. By symmetry, we may then assume without loss of generality that $\nu\le\mu$, $n\le m$, and $n\le\nu$. We thus arrive at

$$S_1 \ll S_{11} + S_{12}, \tag{6.9}$$

where $S_{11}$ and $S_{12}$ are defined by

$$S_{11} := \frac1X \sum_{\substack{\mu,\nu,m,n\le V/X\\ \nu\le\mu\\ n\le m\\ n=\nu}} a_1 \left| \int_X^{X_1} x^{2-\frac2k}\cos\Big(2\pi k x^{1/k}\big(\mu^{1/k}-m^{1/k}\big)\Big)\,dx\right|$$

and

$$S_{12} := \frac1X \sum_{\substack{\mu,\nu,m,n\le V/X\\ \nu\le\mu\\ n\le m\\ n<\nu}} a_1 \left| \int_X^{X_1} x^{2-\frac2k}\cos\Big(2\pi k x^{1/k}\big(\mu^{1/k}+\nu^{1/k}-m^{1/k}-n^{1/k}\big)\Big)\,dx\right|. \tag{6.10}$$

To bound $S_{11}$, we further write

$$S_{11} = S_{111} + S_{112}, \tag{6.11}$$

where $S_{111}$ is the part of $S_{11}$ with $m = \mu$ and $S_{112}$ is the part with $m \neq \mu$. Using the definitions (6.3) and (6.4), we deduce that
(6.12)
$$S_{111} = \frac{1}{X} \sum_{\substack{\mu,\nu,m,n \leq V/X \\ \nu \leq \mu \\ n \leq m \\ n=\nu \\ m=\mu}} a_1 \int_X^{X_1} x^{2-\frac{2}{k}}\, dx \leq \frac{1}{X} \sum_{\substack{m,n \leq V/X \\ n \leq m}} \frac{d_k^2(m) d_k^2(n)}{(mn)^{1+\frac{1}{k}}} \int_X^{2X} x^{2-\frac{2}{k}}\, dx \ll X^{2-\frac{2}{k}}.$$

On the other hand, to bound $S_{112}$, we may assume without loss of generality that $m < \mu$, and integrate by parts to arrive at

$$S_{112} \ll X^{2-\frac{3}{k}} \sum_{\substack{\mu,\nu,m,n \leq V/X \\ \nu \leq \mu \\ n \leq m \\ n=\nu \\ m<\mu}} \frac{a_1}{\mu^{1/k} - m^{1/k}} = X^{2-\frac{3}{k}} \sum_{\substack{\mu,m,n \leq V/X \\ n \leq \mu \\ n \leq m \\ m<\mu}} \frac{d_k(\mu) d_k(m) d_k^2(n)}{n^{1+\frac{1}{k}} (\mu m)^{\frac{1}{2}+\frac{1}{2k}} \left(\mu^{1/k} - m^{1/k}\right)}.$$

Since $\mu^{1/k} - m^{1/k} \gg (\mu - m)\mu^{\frac{1}{k}-1}$ for $\mu > m$ and $d_k(j) \ll j^{\varepsilon}$ for all positive integers $j$, it follows that

$$S_{112} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,m,n \leq V/X \\ n \leq \mu \\ n \leq m \\ m<\mu}} \frac{\mu^{\frac{1}{2}-\frac{3}{2k}}}{n^{1+\frac{1}{k}} m^{\frac{1}{2}+\frac{1}{2k}} (\mu - m)}.$$

The $m$-sum here is $O(1)$ by the Cauchy-Schwarz inequality, and so

$$S_{112} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \left(\frac{V}{X}\right)^{\frac{3}{2}-\frac{3}{2k}} = X^{\frac{1}{2}-\frac{3}{2k}} V^{\frac{3}{2}-\frac{3}{2k}+\varepsilon}.$$

It follows from this, (6.12), and (6.11) that

$$S_{11} \ll X^{2-\frac{2}{k}} + X^{\frac{1}{2}-\frac{3}{2k}} V^{\frac{3}{2}-\frac{3}{2k}+\varepsilon}. \tag{6.13}$$

Having estimated $S_{11}$, we next bound $S_{12}$, which is defined by (6.10). Let $\xi > 0$ be a parameter, to be chosen later, such that $\xi < 1$ and

$$\xi \left(\frac{V}{X}\right)^{1-\frac{1}{k}} = o(1) \tag{6.14}$$

as $X \to \infty$. Define $\Lambda_1$ by

$$\Lambda_1 = \Lambda_1(\mu, \nu, m, n; k) := \mu^{1/k} + \nu^{1/k} - m^{1/k} - n^{1/k}. \tag{6.15}$$

Split the sum $S_{12}$, defined by (6.10), and write

$$S_{12} = S_{121} + S_{122}, \tag{6.16}$$

where $S_{121}$ is the part with $|\Lambda_1| \leq \xi$ and $S_{122}$ is the part with $|\Lambda_1| > \xi$.

To estimate $S_{121}$, we bound the integral in (6.10) trivially using (6.4), and then use (6.3) to deduce that

$$S_{121} \ll X^{2-\frac{2}{k}} \sum_{\substack{\mu,\nu,m,n\le V/X\\ \nu\le\mu\\ n\le m\\ n<\nu\\ |\Lambda_1|\le\xi}} a_1 \ll X^{2-\frac{2}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m,n\le V/X\\ \nu\le\mu\\ n\le m\\ n<\nu\\ |\Lambda_1|\le\xi}} \frac{1}{(\mu\nu m n)^{\frac12+\frac{1}{2k}}}.$$

Note that the summation conditions imply that $\mu > 1$. We partition the range of the summation variable $\mu$ into dyadic intervals $(1,2]$, $(2,4]$, $(4,8]$, $\ldots$ to write

$$S_{121} \ll X^{2-\frac{2}{k}} V^{\varepsilon} \sum_{M} \sum_{\substack{\mu,\nu,m,n\le V/X\\ M<\mu\le 2M\\ \nu\le\mu\\ n\le m\\ n<\nu\\ |\Lambda_1|\le\xi}} \frac{1}{(\mu\nu m n)^{\frac12+\frac{1}{2k}}}, \tag{6.17}$$

where $M \geq 1$ runs through the powers of 2 less than or equal to $V/X$. Our assumption that $\xi < 1$, the definition (6.15) of $\Lambda_1$, and the conditions satisfied by the summation variables in (6.17) imply that $\nu, n, m \ll \mu \ll M$. It follows from this and the polynomial identity $x^k - y^k = (x-y)(x^{k-1} + x^{k-2}y + \cdots + y^{k-1})$ that

$$\left| \left(\mu^{1/k} + \nu^{1/k} - n^{1/k}\right)^k - m \right| \ll_k |\Lambda_1| \mu^{1-\frac{1}{k}} \ll \xi M^{1-\frac{1}{k}}. \tag{6.18}$$

From this, (6.14), and the fact that $M \leq V/X$, we see for large enough $X$ that, for each triple $\mu, \nu, n$ in (6.17), there is at most one integer $m$ such that $|\Lambda_1| \leq \xi$, and such an $m$ must satisfy

$$m \asymp \left(\mu^{1/k} + \nu^{1/k} - n^{1/k}\right)^k \asymp \mu$$

because $\nu > n$. Furthermore, if such an $m$ exists, then it follows from (6.18) that

$$\left\| \left(\mu^{1/k} + \nu^{1/k} - n^{1/k}\right)^k \right\| \ll_k \xi M^{1-\frac{1}{k}}, \tag{6.19}$$

where $\|x\|$ denotes the distance from $x$ to the nearest integer. These and (6.17) imply that

$$S_{121} \ll X^{2-\frac{2}{k}} V^{\varepsilon} \sum_{M} \sum_{\substack{\mu,\nu,n\le V/X\\ M<\mu\le 2M\\ \nu\le\mu\\ n<\nu\\ (6.19)}} \frac{1}{(\nu n)^{\frac12+\frac{1}{2k}} \mu^{1+\frac{1}{k}}}.$$

From this and Lemma 3.3 with $W = M$, $\rho = O_k(\xi M^{1-\frac{1}{k}})$, and $\alpha = \nu^{1/k} - n^{1/k}$, we arrive at

$$\begin{aligned} S_{121} \ll X^{2-\frac{2}{k}} V^{\varepsilon} \sum_{M} \frac{1}{M^{1+\frac{1}{k}}} \sum_{\substack{\nu,n\le V/X\\ \nu\le 2M\\ n<\nu}} \frac{1}{(\nu n)^{\frac12+\frac{1}{2k}}} \Big( \xi M^{2-\frac{1}{k}} + M^{\frac23-\frac{1}{3k}} \left(\nu^{1/k} - n^{1/k}\right)^{1/3} \\ + M^{\frac12+\frac{1}{2k}} \left(\nu^{1/k} - n^{1/k}\right)^{-1/2} \Big). \end{aligned} \tag{6.20}$$

Recall that, as in (6.17), $M$ runs through the powers of 2 in the interval $[1, V/X]$. Thus

$$\sum_M \frac{1}{M^{1+\frac{1}{k}}} \sum_{\substack{\nu,n\le V/X\\ \nu\le 2M\\ n<\nu}} \frac{1}{(\nu n)^{\frac12+\frac{1}{2k}}}\left(\xi M^{2-\frac{1}{k}}\right) \ll \xi \sum_M M^{1-\frac{2}{k}} \sum_{\nu\le 2M} \frac{1}{\nu^{1/k}} \ll \xi\left(\frac{V}{X}\right)^{2-\frac{3}{k}}. \tag{6.21}$$

Similarly, since $\left(\nu^{1/k}-n^{1/k}\right)^{1/3} \le \nu^{1/(3k)}$, we have

$$\sum_M \frac{1}{M^{1+\frac{1}{k}}} \sum_{\substack{\nu,n\le V/X\\ \nu\le 2M\\ n<\nu}} \frac{1}{(\nu n)^{\frac12+\frac{1}{2k}}}\left(M^{\frac23-\frac{1}{3k}}\left(\nu^{1/k}-n^{1/k}\right)^{1/3}\right) \ll \left(\frac{V}{X}\right)^{\frac23-\frac{2}{k}} \log V \tag{6.22}$$

(the factor $\log V$ is necessary only when $k=3$). To estimate the contribution of the term with $\left(\nu^{1/k}-n^{1/k}\right)^{-1/2}$ in (6.20), we use the bound $\nu^{1/k}-n^{1/k} \gg (\nu-n)\nu^{\frac{1}{k}-1}$ to deduce that

$$\begin{aligned}\sum_{n<\nu} \frac{\left(\nu^{1/k}-n^{1/k}\right)^{-1/2}}{n^{\frac12+\frac{1}{2k}}} = \sum_{n<\nu/2} + \sum_{\nu/2<n<\nu} &\ll \frac{1}{\nu^{1/(2k)}} \sum_{n<\nu/2} \frac{1}{n^{\frac12+\frac{1}{2k}}} + \frac{1}{\nu^{1/k}} \sum_{\nu/2<n<\nu} (\nu-n)^{-1/2}\\ &\ll \nu^{\frac12-\frac1k}.\end{aligned}$$

Hence

$$\begin{aligned}\sum_M \frac{1}{M^{1+\frac{1}{k}}} \sum_{\substack{\nu,n\le V/X\\ \nu\le 2M\\ n<\nu}} \frac{1}{(\nu n)^{\frac12+\frac{1}{2k}}}\left(M^{\frac12+\frac{1}{2k}}\left(\nu^{1/k}-n^{1/k}\right)^{-1/2}\right) &\ll \sum_M \frac{1}{M^{\frac12+\frac{1}{2k}}} \sum_{\nu\le 2M} \frac{1}{\nu^{3/(2k)}}\\ &\ll \sum_M M^{\frac12-\frac2k} \ll \max\left\{\log V, (V/X)^{\frac12-\frac2k}\right\}.\end{aligned}$$

From this, (6.22), (6.21), and (6.20), we arrive at

$$S_{121} \ll X^{2-\frac{2}{k}} V^{\varepsilon}\left(\xi\left(\frac{V}{X}\right)^{2-\frac{3}{k}} + \left(\frac{V}{X}\right)^{\frac23-\frac2k}\log V + \max\left\{\log V, (V/X)^{\frac12-\frac2k}\right\}\right). \tag{6.23}$$

We may assume that $V \ge X$ since otherwise $S_{121}=0$ by (6.17). Thus $(V/X)^{\frac12-\frac2k} \le (V/X)^{\frac23-\frac2k}$, and (6.23) simplifies to

$$S_{121} \ll \xi X^{1/k} V^{2-\frac{3}{k}+\varepsilon} + X^{4/3} V^{\frac23-\frac2k+\varepsilon}. \tag{6.24}$$

Having bounded the sum $S_{121}$ in (6.16), we next estimate $S_{122}$, which is the part of (6.10) that has $|\Lambda_1| > \xi$. Recalling the definitions (6.4) of $X_1$ and (6.15) of $\Lambda_1$, we estimate the integral in (6.10) via integration by parts and then use (6.3) to arrive at

$$S_{122} \ll X^{2-\frac{3}{k}} \sum_{\substack{\mu,\nu,m,n\le V/X\\ \nu\le\mu\\ n\le m\\ n<\nu\\ |\Lambda_1|>\xi}} \frac{a_1}{|\Lambda_1|} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m,n\le V/X\\ \nu\le\mu\\ n\le m\\ n<\nu\\ |\Lambda_1|>\xi}} \frac{1}{(\mu\nu m n)^{\frac12+\frac{1}{2k}}|\Lambda_1|}. \tag{6.25}$$

We split the range of $|\Lambda_1|$ dyadically to deduce from (6.25) that

$$S_{122} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{L > \xi/2} \frac{1}{L} \sum_{\substack{\mu,\nu,m,n \leq V/X \\ \nu \leq \mu \\ n \leq m \\ n < \nu \\ L < |\Lambda_1| \leq 2L}} \frac{1}{(\mu\nu m n)^{\frac{1}{2}+\frac{1}{2k}}}, \tag{6.26}$$

where $L$ runs through the numbers $2^j$ with $j \in \mathbb{Z}$. Now if $n < \nu \leq \mu$ and $\Lambda_1 \ll \mu^{1/k}$, then the definition (6.15) of $\Lambda_1$ and the binomial theorem imply

$$m = \big(\mu^{1/k} + \nu^{1/k} - n^{1/k} - \Lambda_1\big)^k = \big(\mu^{1/k} + \nu^{1/k} - n^{1/k}\big)^k + O_k\big(|\Lambda_1| \mu^{1-\frac{1}{k}}\big). \tag{6.27}$$

Let $\varepsilon_k > 0$ be a small enough constant, depending only on $k$, such that if $|\Lambda_1| \leq 2\varepsilon_k \mu^{1/k}$, then the error term in (6.27) has absolute value $\leq \mu/2$. Split the $L$-sum in (6.26) and write

$$S_{122} \ll \Sigma_1 + \Sigma_2, \tag{6.28}$$

where $\Sigma_1$ is the part with $L \leq \varepsilon_k \mu^{1/k}$ and $\Sigma_2$ is the part with $L > \varepsilon_k \mu^{1/k}$. To bound $\Sigma_1$, observe that if $n < \nu \leq \mu$ and $L \leq \varepsilon_k \mu^{1/k}$, then (6.27) implies that there are at most $1 + O_k(L\mu^{1-\frac{1}{k}})$ integers $m$ satisfying $|\Lambda_1| \leq 2L$. Moreover, each such $m$ satisfies $m \asymp \mu$ by (6.27), the definition of $\varepsilon_k$ below (6.27), and the fact that $n < \nu \leq \mu$. Thus

$$\Sigma_1 \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,n \leq V/X \\ \nu \leq \mu \\ n < \nu}} \frac{1}{(\nu n)^{\frac{1}{2}+\frac{1}{2k}} \mu^{1+\frac{1}{k}}} \sum_{\xi/2 < L \leq \varepsilon_k \mu^{1/k}} \frac{1}{L}\Big(1 + O_k\big(L\mu^{1-\frac{1}{k}}\big)\Big). \tag{6.29}$$

Recall that, as in (6.26), $L$ runs through powers of 2. Thus the number of terms in the $L$-sum in (6.29) is $\ll V^{\varepsilon}|\log \xi|$, and so

$$\begin{aligned} \Sigma_1 &\ll X^{2-\frac{3}{k}} V^{\varepsilon} |\log \xi| \sum_{\substack{\mu,\nu,n \leq V/X \\ \nu \leq \mu \\ n < \nu}} \frac{1}{(\nu n)^{\frac{1}{2}+\frac{1}{2k}} \mu^{1+\frac{1}{k}}} \left(\frac{1}{\xi} + \mu^{1-\frac{1}{k}}\right) \\ &\ll X^{2-\frac{3}{k}} V^{\varepsilon} |\log \xi| \sum_{\substack{\mu,\nu \leq V/X \\ \nu \leq \mu}} \frac{1}{\nu^{1/k} \mu^{1+\frac{1}{k}}} \left(\frac{1}{\xi} + \mu^{1-\frac{1}{k}}\right) \\ &\ll \xi^{-1} X^{1-\frac{1}{k}} V^{1-\frac{2}{k}+\varepsilon} |\log \xi| + V^{2-\frac{3}{k}+\varepsilon} |\log \xi|. \end{aligned} \tag{6.30}$$

To bound the sum $\Sigma_2$ in (6.28), ignore the conditions $L < |\Lambda_1| \leq 2L$ and $n \leq m$, and then evaluate the $L$-sum as a geometric series to deduce that

$$\begin{aligned} \Sigma_2 &\ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m,n \leq V/X \\ \nu \leq \mu \\ n < \nu}} \frac{1}{(\mu\nu m n)^{\frac{1}{2}+\frac{1}{2k}}} \sum_{L > \varepsilon_k \mu^{1/k}} \frac{1}{L} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m,n \leq V/X \\ \nu \leq \mu \\ n < \nu}} \frac{1}{(\nu m n)^{\frac{1}{2}+\frac{1}{2k}} \mu^{\frac{1}{2}+\frac{3}{2k}}} \\ &\ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m \leq V/X \\ \nu \leq \mu}} \frac{1}{\nu^{1/k} m^{\frac{1}{2}+\frac{1}{2k}} \mu^{\frac{1}{2}+\frac{3}{2k}}} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\mu,m \leq V/X} \frac{\mu^{\frac{1}{2}-\frac{5}{2k}}}{m^{\frac{1}{2}+\frac{1}{2k}}} \ll V^{2-\frac{3}{k}+\varepsilon}. \end{aligned}$$

From this, (6.30), and (6.28), we arrive at

$$S_{122} \ll \xi^{-1} X^{1-\frac{1}{k}} V^{1-\frac{2}{k}+\varepsilon} |\log \xi| + V^{2-\frac{3}{k}+\varepsilon}(1 + |\log \xi|). \tag{6.31}$$

This, (6.24), (6.16), (6.13), and (6.9) now imply

$$\begin{aligned} S_1 \ll X^{2-\frac{2}{k}} + X^{\frac{1}{2}-\frac{3}{2k}} V^{\frac{3}{2}-\frac{3}{2k}+\varepsilon} + \xi X^{1/k} V^{2-\frac{3}{k}+\varepsilon} + X^{4/3} V^{\frac{2}{3}-\frac{2}{k}+\varepsilon} \\ + \xi^{-1} X^{1-\frac{1}{k}} V^{1-\frac{2}{k}+\varepsilon} |\log \xi| + V^{2-\frac{3}{k}+\varepsilon}(1+|\log \xi|). \end{aligned} \tag{6.32}$$

This completes our estimation of $S_1$.

Our next task is to bound $S_2$, which is defined by (6.7). The procedure is similar to our estimation of $S_{12}$, which starts with (6.16), and so we only present a sketch. Define $\Lambda_2$ by

$$\Lambda_2 = \Lambda_2(\mu, \nu, m, n; k) := \mu^{1/k} + \nu^{1/k} + m^{1/k} - n^{1/k} \tag{6.33}$$

and let $\xi$ be as in (6.14). Split the sum $S_2$ in (6.7) to write

$$S_2 = S_{21} + S_{22}, \tag{6.34}$$

where $S_{21}$ is the part with $|\Lambda_2| \le \xi$ and $S_{22}$ is the part with $|\Lambda_2| > \xi$. To bound $S_{21}$, we may assume that $m \le \nu \le \mu$. We bound the integral trivially and partition the range of $\mu$ into dyadic intervals to deduce that, similarly to (6.17), we have

$$S_{21} \ll X^{2-\frac{2}{k}} V^{\varepsilon} \sum_{M} \sum_{\substack{\mu,\nu,m,n \le V/X \\ M<\mu\le 2M \\ m\le\nu\le\mu \\ |\Lambda_2|\le\xi}} \frac{1}{(\mu\nu m n)^{\frac{1}{2}+\frac{1}{2k}}},$$

where $M$ runs through the powers of 2 in the interval $[1/2, V/X]$. For each triple $m, \nu, \mu$ in this sum, the condition (6.14) ensures that there is at most one integer $n$ such that $|\Lambda_2| \le \xi$, and such an $n$ satisfies $n \asymp \mu$. If such an $n$ exists, then

$$\left\| \left(\mu^{1/k} + \nu^{1/k} + m^{1/k}\right)^k \right\| \ll_k \xi M^{1-\frac{1}{k}}.$$

It follows from these and Lemma 3.3 that

$$\begin{aligned} S_{21} \ll X^{2-\frac{2}{k}} V^{\varepsilon} \sum_{M} \frac{1}{M^{1+\frac{1}{k}}} \sum_{\substack{\nu,m\le V/X \\ m\le\nu\le 2M}} \frac{1}{(\nu m)^{\frac{1}{2}+\frac{1}{2k}}} \Big( \xi M^{2-\frac{1}{k}} + M^{\frac{2}{3}-\frac{1}{3k}} \left(\nu^{1/k} + m^{1/k}\right)^{1/3} \\ + M^{\frac{1}{2}+\frac{1}{2k}} \left(\nu^{1/k} + m^{1/k}\right)^{-1/2} \Big) \end{aligned}$$

(to handle the case $M = 1/2$, we note that the conclusion of Lemma 3.3 holds trivially for $W = 1/2$). By an argument similar to our proof that (6.20) implies (6.24), we arrive at

$$S_{21} \ll \xi X^{1/k} V^{2-\frac{3}{k}+\varepsilon} + X^{4/3} V^{\frac{2}{3}-\frac{2}{k}+\varepsilon}. \tag{6.35}$$

Next, to estimate the sum $S_{22}$ in (6.34), we bound the integral in (6.7) via integration by parts and split the range of $|\Lambda_2|$ dyadically to deduce that, similarly to (6.26), we have

$$S_{22} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{L>\xi/2} \frac{1}{L} \sum_{\substack{\mu,\nu,m,n\le V/X \\ m\le\nu\le\mu \\ L<|\Lambda_2|\le 2L}} \frac{1}{(\mu\nu m n)^{\frac{1}{2}+\frac{1}{2k}}}.$$

If $m \le \nu \le \mu$ and $|\Lambda_2| \ll \mu^{1/k}$, then the definition (6.33) of $\Lambda_2$ implies that

$$n = \left(\mu^{1/k} + \nu^{1/k} + m^{1/k}\right)^k + O_k\left(|\Lambda_2| \mu^{1-\frac{1}{k}}\right).$$

Hence, as in our arguments below (6.27), there exists a constant $\varepsilon_k > 0$ such that if $m \le \nu \le \mu$ and $L \le \varepsilon_k \mu^{1/k}$, then there are at most $1 + O_k(L\mu^{1-\frac{1}{k}})$ integers $n$ satisfying $|\Lambda_2| \le 2L$, and each such $n$ satisfies $n \asymp \mu$. The estimations leading up to (6.31) then show that

$$S_{22} \ll \xi^{-1} X^{1-\frac{1}{k}} V^{1-\frac{2}{k}+\varepsilon} |\log \xi| + V^{2-\frac{3}{k}+\varepsilon}(1 + |\log \xi|).$$

From this, (6.35), and (6.34), we arrive at

$$S_2 \ll \xi X^{1/k} V^{2-\frac{3}{k}+\varepsilon} + X^{4/3} V^{\frac{2}{3}-\frac{2}{k}+\varepsilon} + \xi^{-1} X^{1-\frac{1}{k}} V^{1-\frac{2}{k}+\varepsilon} |\log \xi| + V^{2-\frac{3}{k}+\varepsilon}(1 + |\log \xi|). \tag{6.36}$$

This finishes our estimation of $S_2$.

It is left to estimate $S_3$, which is defined by (6.8). We bound the right-hand side of (6.8) by taking the absolute value of each term. By symmetry, we may then assume without loss of generality that $n \le m \le \nu \le \mu$. Recalling the definition (6.4) of $X_1$, we estimate the integral in (6.8) via integration by parts and then use (6.3) to deduce that

$$S_3 \ll X^{2-\frac{3}{k}} \sum_{\substack{\mu,\nu,m,n \le V/X \\ n \le m \le \nu \le \mu}} \frac{a_1}{\mu^{1/k}} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m,n \le V/X \\ n \le m \le \nu \le \mu}} \frac{1}{(\nu m n)^{\frac{1}{2}+\frac{1}{2k}} \mu^{\frac{1}{2}+\frac{3}{2k}}}.$$

We estimate the $n$-sum, $m$-sum, $\nu$-sum, and $\mu$-sum, in that order, to arrive at

$$S_3 \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu,m \le V/X \\ m \le \nu \le \mu}} \frac{1}{m^{1/k} \nu^{\frac{1}{2}+\frac{1}{2k}} \mu^{\frac{1}{2}+\frac{3}{2k}}} \ll X^{2-\frac{3}{k}} V^{\varepsilon} \sum_{\substack{\mu,\nu \le V/X \\ \nu \le \mu}} \frac{\nu^{\frac{1}{2}-\frac{3}{2k}}}{\mu^{\frac{1}{2}+\frac{3}{2k}}} \ll V^{2-\frac{3}{k}+\varepsilon}. \tag{6.37}$$

Now from (6.1), (6.5), (6.32), (6.36), (6.37), we conclude that if $1 \le Y \le T \le X$, $V$ is defined by (6.2), and $0 < \xi < 1$ such that (6.14) holds, then

$$\begin{aligned} \frac{1}{X} \int_X^{2X} |\Delta_k(x)|^4 \, dx \ll & X^{2-\frac{2}{k}} + X^{\frac{1}{2}-\frac{3}{2k}} V^{\frac{3}{2}-\frac{3}{2k}+\varepsilon} + \xi X^{1/k} V^{2-\frac{3}{k}+\varepsilon} + X^{4/3} V^{\frac{2}{3}-\frac{2}{k}+\varepsilon} \\ & + \xi^{-1} X^{1-\frac{1}{k}} V^{1-\frac{2}{k}+\varepsilon} |\log \xi| + V^{2-\frac{3}{k}+\varepsilon}(1 + |\log \xi|) + \frac{X^{2+\varepsilon}}{Y} \\ & + \frac{X^{4+\varepsilon}}{Y^{2k+2}} + X^{\varepsilon} Y^{2k-4} + \frac{X^{2+\varepsilon}}{Y^4} + \frac{X^{4+\varepsilon}}{T^4} \end{aligned} \tag{6.38}$$

under the assumption of the Lindelöf Hypothesis. We now choose $\xi = X^{-\frac{1}{k}-\varepsilon}$, $T = X^{\frac{1}{2}+\frac{1}{2k}+\varepsilon}$, and $Y = X^{1/(k-1)}$, so that (6.2) gives $V \ll X^{k/(k-1)}$, and the conditions $1 \le Y \le T \le X$ and (6.14) are satisfied. With these choices for the parameters, (6.38) gives

$$\frac{1}{X} \int_X^{2X} |\Delta_k(x)|^4 \, dx \ll X^{2-\frac{1}{k-1}+\varepsilon}.$$

This completes the proof of Theorem 1.6.

## 7. Intervals containing no sign changes

We now have all the ingredients to prove Theorems 1.1 and 1.2 using the method of Heath-Brown and Tsang [7]. Let $\eta > 0$ be an arbitrarily small (fixed) constant. Define $G_k(x)$ by

$$G_k(x) := |\Delta_k(x)| - \left(\frac{1}{2} C_k - \eta\right) x^{\frac{1}{2}-\frac{1}{2k}}, \tag{7.1}$$

where the constant $C_k$ is defined by (1.3). Then Tong's formula (1.2) and the Cauchy-Schwarz inequality imply

$$\begin{aligned}\int_X^{2X}\Big(G_k(x)\Big)^2\,dx \geq &\int_X^{2X}|\Delta_k(x)|^2\,dx + \left(\frac{1}{2}C_k-\eta\right)^2\int_X^{2X}x^{1-\frac{1}{k}}\,dx\\ &-2\left(\int_X^{2X}|\Delta_k(x)|^2\,dx\right)^{1/2}\left(\int_X^{2X}\left(\frac{1}{2}C_k-\eta\right)^2x^{1-\frac{1}{k}}\,dx\right)^{1/2}\\ \geq&(1+o(1))\left(\frac{1}{2}C_k+\eta\right)^2\int_X^{2X}x^{1-\frac{1}{k}}\,dx.\end{aligned} \tag{7.2}$$

To examine the change in $G_k$ over a short interval, observe that the mean value theorem of differential calculus implies

$$(x+h)^{\frac{1}{2}-\frac{1}{2k}}-x^{\frac{1}{2}-\frac{1}{2k}}\ll_k hx^{-\frac{1}{2}-\frac{1}{2k}}$$

for $h\geq 0$. It follows from this, the definition (7.1) of $G_k$, and the inequality $|a+b|^2\ll|a|^2+|b|^2$ that

$$\sup_{0\leq h\leq H}\Big(G_k(x+h)-G_k(x)\Big)^2\ll\sup_{0\leq h\leq H}\Big(\Delta_k(x+h)-\Delta_k(x)\Big)^2+H^2x^{-1-\frac{1}{k}}.$$

We integrate both sides from $x=X$ to $x=2X$ and then apply Corollaries 1.4 and 1.5 to deduce that

$$\int_X^{2X}\sup_{0\leq h\leq H}\Big(G_k(x+h)-G_k(x)\Big)^2\,dx\ll X^{2-\frac{1}{k}-\varepsilon} \tag{7.3}$$

unconditionally for $k=3$ and $H=X^{\frac{13}{27}-\varepsilon}$, or assuming the Lindelöf Hypothesis (LH) for $k=3$ and $H=X^{\frac{3}{5}-\varepsilon}$ and for $k\geq 4$ and $H=X^{\frac{3}{4}+\frac{1}{4k-8}-\frac{1}{k}-\varepsilon}$. Now define $W_k(x)$ by

$$W_k(x)=W_k(x;H):=G_k^2(x)-\sup_{0\leq h\leq H}\Big(G_k(x+h)-G_k(x)\Big)^2-\left(\frac{1}{2}C_kx^{\frac{1}{2}-\frac{1}{2k}}\right)^2, \tag{7.4}$$

and let $\mathcal{S}$ be the set

$$\mathcal{S}:=\{x\in[X,2X]:W_k(x)>0\}. \tag{7.5}$$

By the definition (7.4) of $W_k$, if $x\in\mathcal{S}$, then

(i) $|G_k(x)|>\sup\limits_{0\leq h\leq H}|G_k(x+h)-G_k(x)|$, and

(ii) $|G_k(x)|>\frac{1}{2}C_kx^{\frac{1}{2}-\frac{1}{2k}}$.

Property (i) implies that $G_k(x)$ has the same sign as $G_k(y)$ for all $y\in[x,x+H]$. Property (ii) implies that $G_k(x)>0$, since otherwise the definition (7.1) of $G_k$ would imply

$$|G_k(x)|=\left(\frac{1}{2}C_k-\eta\right)x^{\frac{1}{2}-\frac{1}{2k}}-|\Delta_k(x)|<\frac{1}{2}C_kx^{\frac{1}{2}-\frac{1}{2k}},$$

which negates (ii). Thus, if $x\in\mathcal{S}$, then $G_k(y)>0$ for all $y\in[x,x+H]$. By (7.1), this means that if $x\in\mathcal{S}$, then

$$|\Delta_k(y)|>\left(\frac{1}{2}C_k-\eta\right)y^{\frac{1}{2}-\frac{1}{2k}} \tag{7.6}$$

for all $y \in [x, x+H]$. To find a lower bound for the Lebesgue measure of $\mathcal{S}$, first observe that (7.2), (7.3), and (7.4) imply

$$\int_X^{2X} W_k(x)\,dx \geq (1+o_\eta(1))\eta(C_k+\eta)\int_X^{2X} x^{1-\frac{1}{k}}\,dx \tag{7.7}$$

unconditionally for $k=3$ and $H = X^{\frac{13}{27}-\varepsilon}$, or assuming LH for $k=3$ and $H = X^{\frac{3}{5}-\varepsilon}$ and for $k \geq 4$ and $H = X^{\frac{3}{4}+\frac{1}{4k-8}-\frac{1}{k}-\varepsilon}$. On the other hand, it follows from the definitions (7.1) of $G_k$ and (7.5) of $\mathcal{S}$ and the Cauchy-Schwarz inequality that

$$\int_X^{2X} W_k(x)\,dx \leq \int_{\mathcal{S}} W_k(x)\,dx \leq \int_{\mathcal{S}} G_k^2(x)\,dx \leq \mathcal{M}^{1/2}\left(\int_X^{2X} G_k^4(x)\,dx\right)^{1/2}, \tag{7.8}$$

where $\mathcal{M}$ is the Lebesgue measure of $\mathcal{S}$. The definition (7.1) of $G_k$, the inequality $|a+b|^4 \ll |a|^4+|b|^4$, Theorem 1.6, and (1.9) give

$$\int_X^{2X} G_k^4(x)\,dx \ll \begin{cases} X^{\frac{235}{96}+\varepsilon} & \text{if } k=3, \text{unconditionally} \\ X^{3-\frac{1}{k-1}+\varepsilon} & \text{if } k\geq 4, \text{ assuming LH.} \end{cases} \tag{7.9}$$

From (7.7), (7.8), and (7.9), we deduce that

$$(1+o_\eta(1))\eta(C_3+\eta)\int_X^{2X} x^{2/3}\,dx \ll \mathcal{M}^{1/2}X^{\frac{235}{192}+\varepsilon} \tag{7.10}$$

unconditionally for $k=3$ and $H = X^{\frac{13}{27}-\varepsilon}$, or assuming LH for $k=3$ and $H = X^{\frac{3}{5}-\varepsilon}$, while

$$(1+o_\eta(1))\eta(C_k+\eta)\int_X^{2X} x^{1-\frac{1}{k}}\,dx \ll \mathcal{M}^{1/2}X^{\frac{3}{2}-\frac{1}{2k-2}+\varepsilon}, \tag{7.11}$$

assuming LH, for $k \geq 4$ and $H = X^{\frac{3}{4}+\frac{1}{4k-8}-\frac{1}{k}-\varepsilon}$. For large enough $X$, the inequalities (7.10) and (7.11) imply the lower bound

$$\mathcal{M} \gg \begin{cases} X^{\frac{85}{96}-\varepsilon} & \text{if } k=3 \\ X^{1+\frac{1}{k-1}-\frac{2}{k}-\varepsilon} & \text{if } k\geq 4 \end{cases}$$

for the Lebesgue measure $\mathcal{M}$ of $\mathcal{S}$. Now since each element $x \in \mathcal{S}$ has the property that (7.6) holds for all $y \in [x, x+H]$, there are at least $\gg \mathcal{M}/H$ disjoint subintervals of $[X, 2X]$ of length $H$ such that (7.6) holds for all $y$ in the subinterval. If (7.6) holds for all $y \in [x, x+H]$, then $\Delta_k$ does not change sign in $[x, x+H]$ because if $\Delta_k$ has a jump discontinuity at $y$, then the jump has size $d_k(y) \ll y^\varepsilon$. This completes the proof of Theorems 1.1 and 1.2.

## Acknowledgments

The first author gratefully acknowledges funding from the National Science Foundation grant DMS-1854398. The second author is grateful for a Graduate College Master's Fellowship from the University of Illinois and a Sloan Scholarship from the Alfred P. Sloan Foundation.

## References


[1] Sandro Bettin and J. Brian Conrey. *Averages of long Dirichlet polynomials.* Riv. Mat. Univ. Parma **12** (2021), no. 1, 1–27.

[2] J. Bourgain. *Decoupling, exponential sums and the Riemann zeta function.* J. Amer. Math. Soc. **30** (2017), 205–224.

[3] Harald Cramér. *Über zwei Sätze des Herrn G. H. Hardy.* Math. Z. **15** (1922), no. 1, 201–210.

[4] Fan Ge and Steven M. Gonek. *The mean square discrepancy in the divisor problem.* Trans. Amer. Math. Soc. **373** (2020), 4713–4734.

[5] J. L. Hafner. *New omega results for two classical lattice point problems.* Invent. Math. **63** (1981), 181–186.

[6] D. R. Heath-Brown. *Mean values of the zeta function and divisor problems.* Recent Progress in Analytic Number Theory, Vol. 1 (Durham, 1979), Academic Press, London-New York, 1981, pp. 115–119.

[7] D. R. Heath-Brown and K. Tsang. *Sign changes of $E(T)$, $\Delta(x)$, and $P(x)$.* J. Number Theory **49** (1994), 73–83.

[8] M. N. Huxley. *Exponential sums and lattice points III.* Proc. London Math. Soc. (3) **87** (2003), 591–609.

[9] M. N. Huxley. *Exponential sums and the Riemann zeta-function V.* Proc. London Math. Soc. (3) **90** (2005), 1–41.

[10] Aleksandar Ivić. *The Riemann Zeta-Function: Theory and Applications.* John Wiley & Sons, New York, 1985 (reprinted by Dover Publications, Inc., Mineola, New York, 2003).

[11] Aleksandar Ivić. *On the mean square of the divisor function in short intervals.* J. Théor. Nombres Bordeaux **21** (2009), no. 2, 251–261.

[12] Aleksandar Ivić. *On the divisor function and the Riemann zeta-function in short intervals.* Ramanujan J. **19** (2009), no. 2, 207–224.

[13] Matti Jutila. *On the divisor problem for short intervals.* Ann. Univ. Turku. Ser. A I **186** (1984), 23–30. Studies in honour of Arto Kustaa Salomaa on the occasion of his fiftieth birthday.

[14] J. P. Keating, B. Rodgers, E. Roditty-Gershon, and Z. Rudnick. *Sums of divisor functions in $\mathbb{F}_q[t]$ and matrix integrals.* Math. Z. **288** (2018), 167–198.

[15] G. Kolesnik. *On the estimation of multiple exponential sums.* Recent Progress in Analytic Number Theory, Vol. 1 (Durham, 1979), Academic Press, London-New York, 1981, pp. 231–246.

[16] E. Landau. *Über die $\zeta$-Funktion und die L-Funktionen.* Math. Z. **20** (1924), 105–125.

[17] Y.-K. Lau and K.-M. Tsang. *Mean square of the remainder term in the Dirichlet divisor problem.* J. Théor. Nombres Bordeaux **7** (1995), 75–92.

[18] Y.-K. Lau and K.-M. Tsang. *On the mean square formula of the error term in the Dirichlet divisor problem.* Math. Proc. Camb. Philos. Soc. **146** (2009), 277–287.

[19] Stephen Lester. *On the variance of sums of divisor functions in short intervals.* Proc. Amer. Math. Soc. **144** (2016), no. 12, 5015–5027.

[20] Micah B. Milinovich and Caroline L. Turnage-Butterbaugh. *Moments of products of automorphic L-functions.* J. Number Theory **139** (2014), 175–204.

[21] Hugh L. Montgomery. *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis.* CBMS Reg. Conf. Ser. Math. 84, Amer. Math. Soc., Providence, 1994.

[22] K. Soundararajan. *Omega results for the divisor and circle problems.* Int. Math. Res. Not. **2003** (2003), no. 36, 1987–1998.

[23] Elias M. Stein and Rami Shakarchi. *Functional Analysis: Introduction to Further Topics in Analysis.* Princeton University Press, 2011.

[24] E. C. Titchmarsh. *The Theory of the Riemann Zeta-function.* Second edition. Revised by D. R. Heath-Brown. The Clarendon Press, Oxford University Press, New York, 1986.

[25] K. C. Tong. *On division problems. I.* Acta. Math. Sinica **5** (1955), 313–324.

[26] K. C. Tong. *On divisor problems, III.* Acta. Math. Sinica **5** (1956), 515–541.

[27] Kai-Man Tsang. *Higher-power moments of $\Delta(x)$, $E(t)$ and $P(x)$.* Proc. London Math. Soc. (3) **65** (1992), 65–84.

[28] K.-M. Tsang. *Mean square of the remainder term in the Dirichlet divisor problem II.* Acta Arith. **71** (1995), 279–299.

[29] Kai-Man Tsang. *Recent progress on the Dirichlet divisor problem and the mean square of the Riemann zeta-function*. Sci. China Math. **53** (2010), 2561–2572.

American Institute of Mathematics, 600 East Brokaw Road, San Jose, CA 95112
*Email address*: sbaluyot@aimath.org

Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green Street, Urbana, IL 61801
*Email address*: ccasti30@illinois.edu