ON SIDON SETS WHICH ARE ASYMPTOTIC BASES OF ORDER 4

Sándor Z. Kiss, Eszter Rozgonyi, Csaba Sándor

Abstract: Let $h \geqslant 2$ be an integer. We say that a set $\mathcal{A}$ of positive integers is an asymptotic basis of order h if every large enough positive integer can be represented as the sum of h terms from $\mathcal{A}$. A set of positive integers $\mathcal{A}$ is called a Sidon set if all the sums $a + b$ with $a, b \in \mathcal{A}$, $a \leqslant b$ are distinct. In this paper we prove the existence of Sidon set $\mathcal{A}$ which is an asymptotic basis of order 4 by using probabilistic methods.

Keywords: additive number theory, representation functions, Sidon set, asymptotic basis.

1. Introduction

Let $\mathbb{N}$ denote the set of nonnegative integers. Let $\mathcal{A} = \{a_1, a_2, \dots\}$ $(a_1 < a_2 < \dots)$ be an infinite sequence of positive integers. For $h \geqslant 2$ integer let $R_h(\mathcal{A}, n)$ denote the number of solutions of the equation

$$a_{i_1} + a_{i_2} + \cdots + a_{i_h} = n, \qquad a_{i_1} \in \mathcal{A}, \dots, a_{i_h} \in \mathcal{A}, \qquad a_{i_1} \leqslant a_{i_2} \leqslant \dots \leqslant a_{i_h}, \ (1)$$

where $n \in \mathbb{N}$. A (finite or infinite) set $\mathcal{A}$ of positive integers is said to be a Sidon set if all the sums $a + b$ with $a, b \in \mathcal{A}$, $a \leqslant b$ are distinct. In other words $\mathcal{A}$ is a Sidon set if for every n positive integer $R_2(\mathcal{A}, n) \leqslant 1$. We say a set $\mathcal{A} \subset \mathbb{N}$ is an asymptotic basis of order h, if every large enough positive integer n can be represented as the sum of h terms from $\mathcal{A}$, i.e., if there exists a positive integer n_0 such that $R_h(\mathcal{A}, n) > 0$ for $n > n_0$. In [3] and [4] P. Erdős, A. Sárközy and V.T. Sós asked if there exists a Sidon set which is an asymptotic basis of order 3. The problem was also appears in [10] (with a typo in it: order 2 is written instead of order 3). It is easy to see [5] that a Sidon set cannot be an asymptotic basis

of order 2. A few years ago J. M. Deshouillers and A. Plagne in [2] constructed a Sidon set which is an asymptotic basis of order at most 7. In [8] S. Kiss proved the existence of a Sidon set which is an asymptotic basis of order 5. In this paper we will improve this result by proving that there exists an asymptotic basis of order 4 which is a Sidon set by using probabilistic methods.

Theorem 1. *There exists an asymptotic basis of order 4 which is a Sidon set.*

Note that at the same time Javier Cilleruelo [1] has proved a slightly stronger result namely the existence of a Sidon set which is an asymptotic basis of order $3+\varepsilon$. He obtained his result independently from our work by using other probabilistic methods. Before we prove the above theorem, we give a short survey of the probabilistic method we are working with.

2. Probabilistic tools

The proof of Theorem 1 is based on the probabilistic method due to Erdős and Rényi. There is an excellent summary of this method in the Halberstam - Roth book [6]. We use the notation and terminology of this book. First we give a survey of the probabilistic tools and notations which we use in the proof of Theorem 1. Let Ω denote the set of strictly increasing sequences of positive integers. In this paper we denote the probability of an event E by $\mathbb{P}(E)$ and the expectation of a random variable ξ by $\mathbb{E}(\xi)$.

Lemma 1. *Let $\theta_1, \theta_2, \theta_3, \ldots$ be real numbers satisfying*

$$0 \leqslant \theta_n \leqslant 1 \qquad (n = 1, 2, \ldots).$$

Then there exists a probability space $(\Omega, X, \mathbb{P})$ with the following two properties:

(i) *For every $n \in \mathbb{N}$, the event $\mathcal{E}^{(n)} = \{\mathcal{A} : \mathcal{A} \in \Omega, n \in \mathcal{A}\}$ is measurable, and $\mathbb{P}(\mathcal{E}^{(n)}) = \theta_n$.*

(ii) *The events $\mathcal{E}^{(1)}$, $\mathcal{E}^{(2)}$, ... are independent.*

See Theorem 13. in [6], p. 142. We denote the characteristic function of the event $\mathcal{E}^{(n)}$ by $t_{(\mathcal{A},n)}$ or we can say the the boolean variable $t_{(\mathcal{A},n)}$ means that:

$$t_{(\mathcal{A},n)} = t_n = \begin{cases} 1, & \text{if } n \in \mathcal{A} \\ 0, & \text{if } n \notin \mathcal{A}. \end{cases}$$

Furthermore, for some $\mathcal{A} = \{a_1, a_2, \ldots\} \in \Omega$ we denote the number of solutions of $a_{i_1} + a_{i_2} + \ldots + a_{i_h} = n$ with $a_{i_1}, \ldots, a_{i_h} \in \mathcal{A}$, $1 \leqslant a_{i_1} < a_{i_2} \ldots < a_{i_h} < n$ by $r_h(\mathcal{A}, n)$.

Let

$$r_h(\mathcal{A}, n) = \sum_{\substack{(a_{i_1}, a_{i_2}, \ldots, a_{i_h}) \\ 1 \leqslant a_{i_1} < \ldots < a_{i_h} < n \\ a_{i_1} + a_{i_2} + \ldots + a_{i_h} = n}} t_{(\mathcal{A}, a_1)} t_{(\mathcal{A}, a_2)} \cdots t_{(\mathcal{A}, a_h)}. \tag{2}$$

Let $r_h^*(\mathcal{A}, n)$ denote the number of those representations of n in the form (1) in which there are at least two equal terms. Thus we have

$$R_h(\mathcal{A}, n) = r_h(\mathcal{A}, n) + r_h^*(\mathcal{A}, n). \tag{3}$$

It is easy to see from (2) that $r_h(\mathcal{A}, n)$ is the sum of random variables. However, for $h > 2$ these variables are not independent because the same $t_{(\mathcal{A}, a_i)}$ may appear in many terms. To overcome this problem we need deeper probabilistic tools. Our proof is based on a method of J. H. Kim and V. H. Vu. We give a short survey of this method. Interested reader can find more details in [7], [11], [12], [13]. Assume that $t_1, t_2, \ldots, t_n$ are independent binary (i.e., all t_i's are in $\{0, 1\}$) random variables. Consider a polynomial $Y = Y(t_1, \ldots t_n)$ in $t_1, t_2, \ldots, t_n$ with degree k (where the degree of this polynomial equals the maximum of the sum of the exponents of the monomials). We say a polynomial Y is totally positive if it can be written in the form $Y = \sum_i e_i \Gamma_i$, where the e_i's are positive and Γ_i is a product of some t_j's. Furthermore, Y is regular if all of its coefficients are between zero and one. We also say Y is simplified, if all of its monomials are square-free (i.e. do not contain any factor of t_i^2), and homogeneous if all the monomials have the same degree. Thus for instance a boolean polynomial is automatically regular and simplified, though not necessarily homogeneous. Given any multi-index $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, we define the partial derivative $\partial^{\underline{\alpha}}(Y)$ of Y as

$$\partial^{\underline{\alpha}}(Y) = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_n}\right)^{\alpha_n} Y(t_1, \ldots t_n),$$

and denote the order of $\underline{\alpha}$ as $|\underline{\alpha}| = \alpha_1 + \cdots + \alpha_n$. For any order $d \geqslant 0$, we denote $\mathbb{E}_d(Y) = \max_{\alpha:|\alpha|=d} \mathbb{E}(\partial^{\underline{\alpha}} Y)$. Thus for instance $\mathbb{E}_0(Y) = \mathbb{E}(Y)$ and $\mathbb{E}_d(Y) = 0$ if d exceeds the degree of Y. We also define $\mathbb{E}_{\geqslant d}(Y) = \max_{d' \geqslant d} \mathbb{E}_{d'}(Y)$. The following result is due to Kim and Vu.

Lemma 2 (J.H. Kim and V.H. Vu). *Let $k \geqslant 1$ and $Y = Y(t_1, \ldots, t_n)$ be a totally positive polynomial of n independent boolean variables $t_1, \ldots, t_n$. Then there exists a constant $C_k > 0$ depending only on k (which is the degree of the polynomial) such that*

$$\mathbb{P}\left(|Y - \mathbb{E}(Y)| \geqslant C_k \lambda^{k - \frac{1}{2}} \sqrt{\mathbb{E}_{\geqslant 0}(Y) \mathbb{E}_{\geqslant 1}(Y)}\right) = O_k\left(e^{-\frac{\lambda}{4} + (k-1)\log n}\right)$$

for all $\lambda > 0$.

See [7] for the proof. Informally this theorem asserts that when the derivatives of Y are smaller on average than Y itself, and the degree of Y is small, then Y is concentrated around its mean.

Finally we need the Borel - Cantelli lemma:

Lemma 3 (Borel-Cantelli). *Let $X_1, X_2, \ldots$ be a sequence of events in a probability space. If*

$$\sum_{j=1}^{+\infty} \mathbb{P}(X_j) < \infty,$$

then with probability 1, at most a finite number of the events X_j can occur.

See in [6], p. 135.

3. Proof of Theorem 1

Define the sequence θ_n in Lemma 1 by

$$\theta_n = n^{-\frac{5}{7}}, \tag{4}$$

that is $\mathbb{P}\left(\{\mathcal{A} : \mathcal{A} \in \Omega, n \in \mathcal{A}\}\right) = n^{-\frac{5}{7}}$, for $n \in \mathbb{N}$. For a given set $\mathcal{A} \in \Omega$ let the set $\mathcal{B}$ be the following

$$\mathcal{B} = \{b : b \in \mathcal{A}, \exists a', a'', a''' \in \mathcal{A} : b + a' = a'' + a''', a', a'', a''' < b\}. \tag{5}$$

Thus $\mathcal{A} \setminus \mathcal{B}$ is a Sidon set. We will prove that $\mathcal{A} \setminus \mathcal{B}$ is an asymptotic basis of order 4 with probability 1. This means that there exists integer N_0 such that with probability 1, $r_4(\mathcal{A} \setminus \mathcal{B}, n) > 0$ for $n \geqslant N_0$. Since

$$r_4(\mathcal{A} \setminus \mathcal{B}, n) = r_4(\mathcal{A}, n) - (r_4(\mathcal{A}, n) - r_4(\mathcal{A} \setminus \mathcal{B}, n)),$$

if we get a lower bound for $r_4(\mathcal{A}, n)$ and an upper bound for $(r_4(\mathcal{A}, n) - r_4(\mathcal{A} \setminus \mathcal{B}, n))$ then we will have a lower bound for $r_4(\mathcal{A} \setminus \mathcal{B}, n)$. So formally we will show that there are positive constants C_1 and N_1 such that with probability 1,

$$r_4(\mathcal{A}, n) > C_1 n^{\frac{1}{7}}, \qquad n \geqslant N_1, \tag{6}$$

and there are positive constants C_2 and N_2 such that with probability 1,

$$r_4(\mathcal{A}, n) - r_4(\mathcal{A} \setminus \mathcal{B}, n) < C_2 \left(\log n\right)^{6.5}, \qquad n \geqslant N_2. \tag{7}$$

In order to prove (6) and (7) we use Lemma 2.

We need the following Lemma (see in [9], p. 134., Lemma 5.3). For the sake of completeness we sketch the proof.

Lemma 4. *Let $N \geqslant 3$, $\alpha, \beta > -1$. Then*

$$\sum_{n=1}^{N-1} n^\alpha (N - n)^\beta = \Theta_{\alpha,\beta} \left(N^{\alpha+\beta+1}\right).$$

Proof.

$$\sum_{n=1}^{N-1} n^\alpha (N-n)^\beta = \sum_{1 \leqslant n \leqslant \frac{N}{2}} n^\alpha (N-n)^\beta + \sum_{\frac{N}{2} < n < N} n^\alpha (N-n)^\beta$$

$$= \Theta_{\alpha,\beta} \left(N^\beta \sum_{1 \leqslant n \leqslant \frac{N}{2}} n^\alpha \right) + \Theta_{\alpha,\beta} \left(N^\alpha \sum_{\frac{N}{2} < n < N} (N-n)^\beta \right)$$

$$= \Theta_{\alpha,\beta} \left(N^\beta \int_1^{\frac{N}{2}} x^\alpha \mathrm{d}x \right) + \Theta_{\alpha,\beta} \left(N^\alpha \int_{\frac{N}{2}}^N x^\beta \mathrm{d}x \right)$$

$$= \Theta_{\alpha,\beta} \left(N^{\alpha+\beta+1} \right). \qquad \blacksquare$$

In the first step we prove (6) by using Lemma 2. To do this, we need the following Lemma.

Lemma 5. *Assume that all of the variables y_i's are different and the t_{y_i}'s are random boolean variables.*

1. *For every nonzero integer a_1 and for every integer m*

$$\mathbb{E} \left(\sum_{\substack{y_1 \\ a_1 y_1 = m}} t_{y_1} \right) = O_{a_1}(1).$$

2. *For every nonzero integers a_1, a_2 and for every integer m*

$$\mathbb{E} \left(\sum_{\substack{(y_1, y_2) \\ a_1 y_1 + a_2 y_2 = m}} t_{y_1} t_{y_2} \right) = O_{a_1, a_2}(1).$$

3. *For every nonzero integers a_1, a_2, a_3 and for every integer m*

$$\mathbb{E} \left(\sum_{\substack{(y_1, y_2, y_3) \\ a_1 y_1 + a_2 y_2 + a_3 y_3 = m}} t_{y_1} t_{y_2} t_{y_3} \right) = O_{a_1, a_2, a_3}(1).$$

Proof.

(1)

$$\mathbb{E}\left(\sum_{\substack{y_1 \\ a_1 y_1 = m}} t_{y_1} \right) = \begin{cases} \left(\frac{m}{a_1}\right)^{-\frac{5}{7}} = O_{a_1}(1) & \text{if} \quad \frac{m}{a_1} \in \mathbb{Z}^+ \\ 0 & \text{if} \quad \frac{m}{a_1} \notin \mathbb{Z}^+ \end{cases}$$

(2) We distinguish two different cases.

Case 1: Assume, that $a_1 > 0$, $a_2 > 0$, thus $m > 0$. (Since y_1, y_2, a_1, a_2 are nonnegative, therefore m can not be negative, at this case.) Thus applying Lemma 4 we get

$$\mathbb{E}\left(\sum_{\substack{(y_1,y_2) \\ a_1 y_1 + a_2 y_2 = m}} t_{y_1} t_{y_2} \right) = O_{a_1,a_2}\left(\sum_{y_1=1}^{\frac{m}{a_1}} y_1^{-\frac{5}{7}} \left(\frac{m - a_1 y_1}{a_2}\right)^{-\frac{5}{7}} \right)$$

$$O_{a_1,a_2}\left(\sum_{y_1=1}^{\frac{m}{a_1}} (a_1 y_1)^{-\frac{5}{7}} (m - a_1 y_1)^{-\frac{5}{7}} \right) = O_{a_1,a_2}\left(\sum_{y=1}^{m-1} y^{-\frac{5}{7}} (m - y)^{-\frac{5}{7}} \right)$$

$$= O_{a_1,a_2}\left(m^{-\frac{3}{7}} \right)$$

$$= O_{a_1,a_2}(1).$$

Case 2: Now assume that $a_1 > 0$, $a_2 < 0$ and $m \geqslant 0$. (If m is negative, then consider the equation $-a_1 y_1 - a_2 y_2 = m$.) We apply Lemma 4 again.

$$\mathbb{E}\left(\sum_{\substack{(y_1,y_2) \\ a_1 y_1 + a_2 y_2 = m}} t_{y_1} t_{y_2} \right) = O_{a_1,a_2}\left(\sum_{y_2=1}^{\infty} y_2^{-\frac{5}{7}} \left(\frac{m - a_2 y_2}{a_1}\right)^{-\frac{5}{7}} \right)$$

$$= O_{a_1,a_2}\left(\sum_{y=1}^{\infty} y^{-\frac{10}{7}} \right) = O_{a_1,a_2}(1).$$

The other cases can be deduced from the aboves. So we leave the details to the reader. (The case $a_1 < 0$, $a_2 > 0$, either $m \geqslant 0$ or $m < 0$ is almost the same like *Case 2*, we have to change the role of a_1 and a_2. The case $a_1 < 0$, $a_2 < 0$, thus $m < 0$ is almost the same like *Case 1*, we have to get $-a_1$ and $-a_2$ instead of a_1 and a_2.)

(3) We distinguish three different cases.

Case 1: Assume, that $a_1 > 0$, $a_2 > 0$, $a_3 > 0$ thus $m > 0$. (Since y_1, y_2, y_3, a_1, a_2, a_3 are nonnegative, therefore m can not be negative, at this case.) Thus applying Lemma 4 we get

$$\mathbb{E}\left(\sum_{\substack{(y_1,y_2,y_3)\\a_1y_1+a_2y_2+a_3y_3=m}} t_{y_1}t_{y_2}t_{y_3}\right)$$

$$= O_{a_1,a_2,a_3}\left(\sum_{y_1=1}^{\frac{m}{a_1}} y_1^{-\frac{5}{7}} \sum_{y_2=1}^{\frac{m-a_1y_1}{a_2}} y_2^{-\frac{5}{7}}\left(\frac{m-a_1y_1-a_2y_2}{a_3}\right)^{-\frac{5}{7}}\right)$$

$$= O_{a_1,a_2,a_3}\left(\sum_{y_1=1}^{\frac{m}{a_1}} y_1^{-\frac{5}{7}} \sum_{y_2=1}^{\frac{m-a_1y_1}{a_2}} (a_2y_2)^{-\frac{5}{7}}(m-a_1y_1-a_2y_2)^{-\frac{5}{7}}\right)$$

$$= O_{a_1,a_2,a_3}\left(\sum_{y_1=1}^{\frac{m}{a_1}} y_1^{-\frac{5}{7}}(m-a_1y_1)^{-\frac{3}{7}}\right) = O_{a_1,a_2,a_3}\left(m^{-\frac{1}{7}}\right)$$

$$= O_{a_1,a_2,a_3}(1).$$

Case 2: Now assume that $a_1 > 0$, $a_2 > 0$, $a_3 < 0$ and $m \geqslant 0$. Thus applying Lemma 4 again we get

$$\mathbb{E}\left(\sum_{\substack{(y_1,y_2,y_3)\\a_1y_1+a_2y_2+a_3y_3=m}} t_{y_1}t_{y_2}t_{y_3}\right)$$

$$= O_{a_1,a_2,a_3}\left(\sum_{y_3=1}^{\infty} y_3^{-\frac{5}{7}} \sum_{y_1=1}^{\frac{m-a_3y_3}{a_1}} y_1^{-\frac{5}{7}}\left(\frac{m-a_3y_3-a_1y_1}{a_2}\right)^{-\frac{5}{7}}\right)$$

$$= O_{a_1,a_2,a_3}\left(\sum_{y_3=1}^{\infty} y_3^{-\frac{5}{7}} \sum_{y_1=1}^{\frac{m-a_3y_3}{a_1}} (a_1y_1)^{-\frac{5}{7}}(m-a_3y_3-a_1y_1)^{-\frac{5}{7}}\right)$$

$$= O_{a_1,a_2,a_3}\left(\sum_{y_3=1}^{\infty} y_3^{-\frac{5}{7}}(m-a_3y_3)^{-\frac{3}{7}}\right) = O_{a_1,a_2,a_3}\left(\sum_{y=1}^{\infty} y^{-\frac{8}{7}}\right)$$

$$= O_{a_1,a_2,a_3}(1).$$

Case 3: Now assume that $a_1 > 0$, $a_2 > 0$, $a_3 < 0$ and $m < 0$. By Lemma 4 we get

$$
\mathbb{E}\left(\sum_{\substack{(y_1, y_2, y_3) \\ a_1 y_1 + a_2 y_2 + a_3 y_3 = m}} t_{y_1} t_{y_2} t_{y_3} \right)
$$

$$
= O_{a_1, a_2, a_3}\left(\sum_{y_3 = \lfloor \frac{m}{a_3} \rfloor + 1}^{\infty} y_3^{-\frac{5}{7}} \sum_{y_1 = 1}^{\frac{m - a_3 y_3}{a_1}} y_1^{-\frac{5}{7}} \left(\frac{m - a_3 y_3 - a_1 y_1}{a_2} \right)^{-\frac{5}{7}} \right)
$$

$$
= O_{a_1, a_2, a_3}\left(\sum_{y_3 = \lfloor \frac{m}{a_3} \rfloor}^{\infty} y_3^{-\frac{5}{7}} \sum_{y_1 = 1}^{\frac{m - a_3 y_3}{a_1}} (a_1 y_1)^{-\frac{5}{7}} (m - a_3 y_3 - a_1 y_1)^{-\frac{5}{7}} \right)
$$

$$
= O_{a_1, a_2, a_3}\left(\sum_{y_3 = \lfloor \frac{m}{a_3} \rfloor}^{\infty} y_3^{-\frac{5}{7}} (m - a_3 y_3)^{-\frac{3}{7}} \right) = O_{a_1, a_2, a_3}\left(\sum_{y=1}^{\infty} y^{-\frac{8}{7}} \right)
$$

$$
= O_{a_1, a_2, a_3}(1).
$$

The other cases can be deduced from the aboves again. We leave the details to the reader. (The case $a_1 > 0$, $a_2 < 0$, $a_3 < 0$, either $m \geqslant 0$ or $m < 0$ is almost the same as *Case 2:* and as Case 3. The case $a_1 < 0$, $a_2 < 0$, $a_3 < 0$ thus $m < 0$ is almost the same as *Case 1*. In both we have to get $-a_1$, $-a_2$ and $-a_3$ instead of a_1, a_2 and a_2.) ■

Now we are ready to prove (6). In view of (2) define Y by

$$
Y = r_4(\mathcal{A}, n) = \sum_{\substack{(x_1, x_2, x_3, x_4) \\ 1 \leqslant x_1 < \cdots < x_4 \\ x_1 + x_2 + x_3 + x_4 = n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4}.
$$

We want to use Lemma 2. To do this we have to estimate the expectation of the variable Y and its partial derivatives. Let $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)$ be a multi-index. In the first step we prove that for $\underline{\alpha} = \underline{0}$

$$
\mathbb{E}\left(\partial^{\underline{\alpha}} Y \right) = \mathbb{E}(Y) = \Theta(n^{\frac{1}{7}}), \tag{8}
$$

and for $\underline{\alpha} \neq \underline{0}$

$$
\mathbb{E}\left(\partial^{\underline{\alpha}} Y \right) = O(1). \tag{9}
$$

Let $\underline{\alpha} = \underline{0}$. By using Lemma 4 we have

$$
\mathbb{E}\left(\partial^{\underline{\alpha}} Y\right) = \mathbb{E}(Y) = \mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ 1 \leqslant x_1 < \cdots < x_4 \\ x_1 + x_2 + x_3 + x_4 = n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \right)
$$

$$
= \Theta\left(\sum_{x_1=1}^{n-3} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1-2} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{n-x_1-x_2-1} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{5}{7}} \right)
$$

$$
= \Theta\left(\sum_{x_1=1}^{n-3} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1-2} x_2^{-\frac{5}{7}} (n - x_1 - x_2)^{-\frac{3}{7}} \right)
$$

$$
= \Theta\left(\sum_{x_1=1}^{n-3} x_1^{-\frac{5}{7}} (n - x_1)^{-\frac{1}{7}} \right) = \Theta\left(n^{\frac{1}{7}} \right),
$$

which shows (8).

Now assume that $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n) \neq \underline{0}$. If there exists an index i such that $\alpha_i \geqslant 2$ or $|\underline{\alpha}| = \sum_{i=1}^{n} \alpha_i \geqslant 5$, then $\partial^{\underline{\alpha}} Y = 0$. It means that in this case $\mathbb{E}(\partial^{\underline{\alpha}} Y) = 0$. So we may assume that for every index i, $\alpha_i \leqslant 1$ and $|\underline{\alpha}| \leqslant 4$. Let $l_1 < l_2 < \cdots < l_w$, for which $\alpha_{l_1} = \ldots = \alpha_{l_w} = 1$, $1 \leqslant w \leqslant 4$. If $1 \leqslant \kappa_1 < \ldots < \kappa_{4-w} \leqslant 4$, $\kappa_j \in \mathbb{N}$, then $\{1, 2, 3, 4\} \setminus \{\kappa_1, \ldots, \kappa_{4-w}\} = \{r_1, \ldots, r_w\}$, where $r_1 < \cdots < r_w$ and $x_{r_1} = l_1, \ldots, x_{r_w} = l_w$. It means that the variables $x_{r_1}, \ldots, x_{r_w}$ occur in the partial derivative of Y and the variables $x_{\kappa_1}, \ldots, x_{\kappa_{4-w}}$ do not. Thus

$$
\mathbb{E}\left(\partial^{\underline{\alpha}} Y\right) = \sum_{\substack{(\kappa_1, \ldots, \kappa_{4-w}) \\ 1 \leqslant \kappa_1 < \ldots < \kappa_{4-w} \leqslant 4}} \mathbb{E}\left(\sum_{\substack{(x_{\kappa_1}, \ldots, x_{\kappa_{4-w}}) \\ \sum_{j=1}^{4-w} x_{\kappa_j} = n - \sum_{j=1}^{w} l_j}} t_{x_{\kappa_1}} \ldots t_{x_{\kappa_{4-w}}} \right).
$$

Since $w \geqslant 1$, thus $4 - w \leqslant 3$. Since the number of the tuples $(\kappa_1, \ldots, \kappa_{4-w})$ is bounded, by Lemma 5 part (2) we get that the expectation is

$$
\mathbb{E}\left(\partial^{\underline{\alpha}} Y\right) = \sum_{\substack{(\kappa_1, \ldots, \kappa_{4-w}) \\ 1 \leqslant \kappa_1 < \ldots < \kappa_{4-w} \leqslant 4}} O(1) = O(1),
$$

which proves (9).

From (8) and (9) we get that

$$
\mathbb{E}_{\geqslant 0}(Y) = \max_{d' \geqslant 0} \mathbb{E}_{d'}(Y) = \mathbb{E}(Y) = \Theta(n^{\frac{1}{7}}),
$$

and

$$
\mathbb{E}_{\geqslant 1}(Y) = \max_{d' \geqslant 1} \mathbb{E}_{d'}(Y) = O(1).
$$

Now apply Lemma 2 with $\lambda = 20 \log n$ and $k = 4$. We get that

$$\mathbb{P}\left(|Y - \mathbb{E}(Y)| \geqslant C_4 (20 \log n)^{3.5} \sqrt{\mathbb{E}_{\geqslant 0}(Y)\mathbb{E}_{\geqslant 1}(Y)}\right) = O\left(\frac{1}{n^2}\right). \qquad (10)$$

Thus by (10) and Lemma 3 we get that if n is large enough, with probability 1,

$$Y = \mathbb{E}(Y) + O\left((\log n)^{3.5}\sqrt{\mathbb{E}_{\geqslant 0}(Y)\mathbb{E}_{\geqslant 1}(Y)}\right) = \mathbb{E}(Y) + O\left(n^{\frac{1}{14}}(\log n)^{3.5}\right) = \Theta(n^{\frac{1}{7}}),$$

which means, that (6) holds.

In the next step we will prove (7), which shows us that the number of those representations in which there is at least one element from the set $\mathcal{B}$ is not too big. Using the definition of the representation functions and the definitions of the sets $\mathcal{A}$ and $\mathcal{B}$ we get

$$r_4(\mathcal{A}, n) - r_4(\mathcal{A} \setminus \mathcal{B}, n) = \sum_{\substack{(a_i, a_j, a_k, a_l) \\ a_i < a_j < a_k < a_l \\ a_i + a_j + a_k + a_l = n \\ \exists m \in \{i,j,k,l\}, \exists a_u, a_v, a_z \in \mathcal{A} \\ a_u, a_v, a_z < a_m \\ a_m + a_u = a_v + a_z}} t_{a_i} t_{a_j} t_{a_k} t_{a_l} \qquad (11)$$

To make the analytic calculations easier we estimate (11) and we have that

$$r_4(\mathcal{A}, n) - r_4(\mathcal{A} \setminus \mathcal{B}, n) = \frac{1}{24} \sum_{\substack{(a_i, a_j, a_k, a_l) \\ a_i, a_j, a_k, a_l \in \mathcal{A} \text{ are distinct} \\ a_i + a_j + a_k + a_l = n \\ \exists m \in \{i,j,k,l\}, \exists a_u, a_v, a_z \in \mathcal{A} \\ a_u, a_v, a_z < a_m \\ a_m + a_u = a_v + a_z}} t_{a_i} t_{a_j} t_{a_k} t_{a_l}$$

$$= \frac{4}{24} \sum_{\substack{(a_i, a_j, a_k, a_l) \\ a_i, a_j, a_k, a_l \in \mathcal{A} \text{ are distinct} \\ a_i + a_j + a_k + a_l = n \\ \exists a_u, a_v, a_z \in \mathcal{A} \\ a_u, a_v, a_z < a_l \\ a_l + a_u = a_v + a_z}} t_{a_i} t_{a_j} t_{a_k} t_{a_l} \qquad (12)$$

$$\leqslant \frac{1}{6} \sum_{\substack{(a_i, a_j, a_k, a_l, a_u, a_v, a_z) \\ a_i, a_j, a_k, a_l, a_u, a_v, a_z \in \mathcal{A} \text{ are distinct} \\ a_i + a_j + a_k + a_l = n \\ a_u, a_v, a_z < a_l \\ a_l + a_u = a_v + a_z}} t_{a_i} t_{a_j} t_{a_k} t_{a_l} t_{a_u} t_{a_v} t_{a_z}$$

Using the variables x_i-s we can write this in the following form

$$r_4(\mathcal{A}, n) - r_4(\mathcal{A} \setminus \mathcal{B}, n) \leqslant \sum_{\substack{(x_1, x_2, x_3, x_4, x_5, x_6, x_7) \\ x_1 + x_2 + x_3 + x_4 = n \\ x_1, x_2, x_3, x_4 \text{ are distinct} \\ x_4 + x_5 = x_6 + x_7 \\ x_5, x_6, x_7 < x_4}} t_{x_1} \dots t_{x_7}. \qquad (13)$$

So this estimation means, that always x_4 will be the element in the 4-tuple (x_1, x_2, x_3, x_4) which hurts the Sidon property. It is easy to see that in the product $t_{x_1} \ldots t_{x_7}$ the t_{x_i} variables are not necessarily independent. So we need to transform (13). For any 7-tuple $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$ with condition $x_5, x_6, x_7 < x_4$ let $\{x_1, x_2, x_3\} \cap \{x_5, x_6, x_7\} = \{x_{i_1}, \ldots, x_{i_s}\}$, where $1 \leqslant i_1 < \cdots < i_s \leqslant 3$. Let $\{x_5, x_6, x_7\} \setminus \{x_{i_1}, \ldots, x_{i_s}\} = \{x_{h_1}, \ldots, x_{h_u}\}$, where $5 \leqslant h_1 < \cdots < h_u \leqslant 7$ and $u \leqslant 3 - s$. Then for every fixed s-tuple $(i_1, \ldots, i_s)$ there exist $s + u$ tuple $(d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u})$ such that we can write the condition $x_4 + x_5 - x_6 - x_7 = 0$ in the following form:

$$x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0, \tag{14}$$

where $x_4, x_{i_1}, \ldots, x_{i_s}, x_{h_1}, \ldots, x_{h_u}$ are different. In (14) $d_{i_j} \neq 0, j = 1, \ldots, s$, $b_{h_j} \neq 0, j = 1, \ldots, u$, there is only one positive coefficients, which is equal to 1 and the sum of the negative coefficients is equal to -2. Since $t_x^k = t_x$, if $k \geqslant 1$ then $t_{x_1} \ldots t_{x_7} = t_{x_1} t_{x_2} t_{x_3} t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}}$.

Thus (13) is equal to the following

$$\sum_{\substack{(i_1, \ldots, i_s) \\ 1 \leqslant i_1 < \ldots < i_s \leqslant 3}} \sum_{\substack{(d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}) \\ \text{only one term is positive and} = 1 \\ \text{the sum of the negative terms is} -2}}$$

$$\times \sum_{\substack{(x_1, x_2, x_3, x_4, x_{h_1}, \ldots, x_{h_u}) \\ x_1 + x_2 + x_3 + x_4 = n \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_1, x_2, x_3, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} t_{x_1} \ldots t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}}.$$

Let the inner sum be

$$Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}} = \sum_{\substack{(x_1, x_2, x_3, x_4, x_{h_1}, \ldots, x_{h_u}) \\ x_1 + x_2 + x_3 + x_4 = n \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_1, x_2, x_3, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} t_{x_1} \ldots t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}}.$$

Since the number of the variables $Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}}$ is bounded, it is enough to show that for every $s + u$ tuple $(d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u})$ with probability 1,

$$Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}} = O\left((\log n)^{6.5}\right). \tag{15}$$

Let's fix an $s+u$ tuple $(d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u})$. We will use Lemma 2. So we have to estimate both the expectation of $Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}}$ and its partial derivatives. First we will show that for every $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)$,

$$\mathbb{E}\left(\partial^{\underline{\alpha}} Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}}\right) = O(1) \tag{16}$$

holds.

Let $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)$. If there exists an index i such that $\alpha_i \geqslant 2$ or $|\underline{\alpha}| = \sum_{i=1}^n \alpha_i \geqslant 5+u$, then $\partial^{\underline{\alpha}} Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}} = 0$. So we may assume that for every index i, $\alpha_i \leqslant 1$ and $|\underline{\alpha}| \leqslant 4+u$. Let's fix $\underline{\alpha} = (\alpha_1, \ldots, \alpha_n)$ and let $\underline{\beta} = (\beta_1, \ldots, \beta_n)$, $\underline{\gamma} = (\gamma_1, \ldots, \gamma_n)$, where $\beta_i \in \mathbb{N}$, $\gamma_i \in \mathbb{N}$ and $\underline{\alpha} = \underline{\beta} + \underline{\gamma}$. Here $\underline{\beta}$ shows the partial derivatives of the variables x_1, x_2, x_3, x_4 and $\underline{\gamma}$ shows the derivatives of the variables $x_{h_1}, \ldots, x_{h_u}$. So we can write the partial derivatives of $Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}}$ in the following form

$$\partial^{\underline{\alpha}} Y_{d_{i_1}, \ldots, d_{i_s}, b_{h_1}, \ldots, b_{h_u}} = \sum_{\substack{(\underline{\beta}, \underline{\gamma}) \\ \underline{\beta} + \underline{\gamma} = \underline{\alpha}}} \sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1 + x_2 + x_3 + x_4 = n \\ x_1, x_2, x_3, x_4 \text{ are distinct}}}$$

$$\times \left(\partial^{\underline{\beta}} t_{x_1} \ldots t_{x_4}\right) \left(\sum_{\substack{(x_{h_1}, \ldots, x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_1, x_2, x_3, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} \partial^{\underline{\gamma}} t_{x_{h_1}} \ldots t_{x_{h_u}} \right).$$

Since the number of pairs $(\underline{\beta}, \underline{\gamma})$ is bounded, it is enough to show that for every fixed pair $(\underline{\beta}, \underline{\gamma})$, where $0 \leqslant \beta_i \leqslant 1$, $0 \leqslant \gamma_i \leqslant 1$, $|\underline{\beta}| + |\underline{\gamma}| = \sum_{i=1}^n (\beta_i + \gamma_i) \leqslant 4 + u$,

$$\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1 + x_2 + x_3 + x_4 = n \\ x_1, x_2, x_3, x_4 \text{ are distinct}}} \right.$$
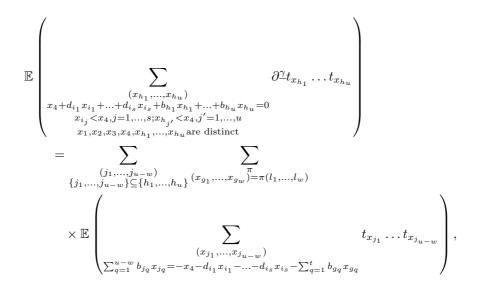
$$\left. \times \left(\partial^{\underline{\beta}} t_{x_1} \ldots t_{x_4}\right) \left(\sum_{\substack{(x_{h_1}, \ldots, x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_1, x_2, x_3, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} \partial^{\underline{\gamma}} t_{x_{h_1}} \ldots t_{x_{h_u}} \right) \right) = O(1)$$

$$\tag{17}$$

holds. Let's fix now a pair $(\underline{\beta}, \underline{\gamma})$. We will show that for every 4-tuple (x_1, x_2, x_3, x_4)

$$
\mathbb{E}\left(\sum_{\substack{(x_{h_1},\ldots,x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j=1,\ldots,s; x_{h_{j'}} < x_4, j'=1,\ldots,u \\ x_1,x_2,x_3,x_4,x_{h_1},\ldots,x_{h_u} \text{ are distinct}}} \partial^{\underline{\gamma}} t_{x_{h_1}} \ldots t_{x_{h_u}} \right) = O(1). \qquad (18)
$$

Let $|\underline{\gamma}| = w$, $\gamma_{l_1} = \ldots = \gamma_{l_w} = 1$, $1 \leqslant l_1 < \ldots < l_w \leqslant n$. Let $\{g_1, \ldots, g_w\} \subseteq \{h_1, \ldots h_u\}$ and $\{j_1, \ldots, j_{u-w}\} = \{h_1, \ldots, h_u\} \setminus \{g_1, \ldots, g_w\}$. These sets shows us that the variables $x_{g_1}, \ldots, x_{g_w}$ occurs in the partial derivative of $Y_{d_{i_1},\ldots,d_{i_s},b_{h_1},\ldots,b_{h_u}}$ and the variables $x_{j_1}, \ldots, x_{j_{u-w}}$ do not. Thus we have

$$
\mathbb{E}\left(\sum_{\substack{(x_{h_1},\ldots,x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j=1,\ldots,s; x_{h_{j'}} < x_4, j'=1,\ldots,u \\ x_1,x_2,x_3,x_4,x_{h_1},\ldots,x_{h_u} \text{ are distinct}}} \partial^{\underline{\gamma}} t_{x_{h_1}} \ldots t_{x_{h_u}} \right)
$$

$$
= \sum_{\substack{(j_1,\ldots,j_{u-w}) \\ \{j_1,\ldots,j_{u-w}\} \subseteq \{h_1,\ldots,h_u\}}} \sum_{\substack{(x_{g_1},\ldots,x_{g_w}) = \pi(l_1,\ldots,l_w) \\ \pi}}
$$

$$
\times \mathbb{E}\left(\sum_{\substack{(x_{j_1},\ldots,x_{j_{u-w}}) \\ \sum_{q=1}^{u-w} b_{j_q} x_{j_q} = -x_4 - d_{i_1} x_{i_1} - \ldots - d_{i_s} x_{i_s} - \sum_{q=1}^{t} b_{g_q} x_{g_q}}} t_{x_{j_1}} \ldots t_{x_{j_{u-w}}} \right),
$$

where $\pi(l_1, \ldots, l_w)$ denotes the permutations of $(l_1, \ldots, l_w)$.

Since the number of the $u - w$ tuples $(j_1, \ldots, j_{u-w})$ and the permutations are bounded and since $u - w \leqslant 3$, thus from Lemma 4 we get that this expectation is

bounded i.e.,

$$\mathbb{E}\left(\sum_{\substack{(x_{j_1},\ldots,x_{j_{u-w}}) \\ \sum_{q=1}^{u-w} b_{j_q} x_{j_q} = -x_4 - d_{i_1} x_{i_1} - \ldots d_{i_s} x_{i_s} - \sum_{q=1}^{w} b_{g_q} x_{g_q}}} t_{x_{j_1}} \ldots t_{x_{j_{u-w}}}\right) = O(1). \quad (19)$$

So from (19) we get that the left hand side of the equation (17) is equal to the following

$$\sum_{\substack{(x_1,x_2,x_3,x_4) \\ x_1+x_2+x_3+x_4=n \\ x_1,x_2,x_3,x_4 \text{ are distinct}}} \mathbb{E}\left(\partial^{\underline{\beta}} t_{x_1} \ldots t_{x_4}\right)$$

$$\times \mathbb{E}\left(\sum_{\substack{(x_{h_1},\ldots,x_{h_u}) \\ x_4+d_{i_1}x_{i_1}+\ldots+d_{i_s}x_{i_s}+b_{h_1}x_{h_1}+\ldots+b_{h_u}x_{h_u}=0 \\ x_{i_j}<x_4, j=1,\ldots,s; x_{h_{j'}}<x_4, j'=1,\ldots,u \\ x_1,x_2,x_3,x_4,x_{h_1},\ldots,x_{h_u} \text{ are distinct}}} \partial^{\underline{\gamma}} t_{x_{h_1}} \ldots t_{x_{h_u}}\right)$$

$$= O\left(\sum_{\substack{(x_1,x_2,x_3,x_4) \\ x_1+x_2+x_3+x_4=n \\ x_1,x_2,x_3,x_4 \text{ are distinct}}} \mathbb{E}\left(\partial^{\underline{\beta}} t_{x_1} \ldots t_{x_4}\right)\right).$$

By (9) we get that if $\underline{\beta} \neq 0$ then this term is equal to $O(1)$. So we may assume that $\underline{\beta} = 0$. This means, that we have to prove that

$$\mathbb{E}\left(\sum_{\substack{(x_1,x_2,x_3,x_4) \\ x_1+x_2+x_3+x_4=n \\ x_1,x_2,x_3,x_4 \text{ are distinct}}} t_{x_1} \ldots t_{x_4}\right.$$

$$\left.\times \left(\sum_{\substack{(x_{h_1},\ldots,x_{h_u}) \\ x_4+d_{i_1}x_{i_1}+\ldots+d_{i_s}x_{i_s}+b_{h_1}x_{h_1}+\ldots+b_{h_u}x_{h_u}=0 \\ x_{i_j}<x_4, j=1,\ldots,s; x_{h_{j'}}<x_4, j'=1,\ldots,u \\ x_1,x_2,x_3,x_4,x_{h_1},\ldots,x_{h_u} \text{ are distinct}}} \partial^{\underline{\gamma}} t_{x_{h_1}} \ldots t_{x_{h_u}}\right)\right) = O(1). \quad (20)$$

If $1 \leqslant i_1 < \ldots < i_s \leqslant 3$ then let $\{1,2,3\} \setminus \{i_1, \ldots i_s\} = \{f_1, \ldots, f_{3-s}\}$. Thus (20) is equivalent to the following

$$
\mathbb{E}\left(\sum_{\substack{(x_{i_1}, \ldots, x_{i_s}, x_4, x_{h_1}, \ldots, x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} \partial^\gamma t_{x_{i_1}} \ldots t_{x_{i_s}} t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}} \right.
$$

$$
\left. \times \left(\sum_{\substack{(x_{f_1}, \ldots, x_{f_{3-s}}) \\ x_{f_1} + \ldots + x_{f_{3-s}} = n - x_{i_1} - \ldots - x_{i_s} - x_4 \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{f_1}, \ldots, x_{f_{3-s}} \text{ are distinct}}} t_{x_{f_1}} \ldots t_{x_{f_{3-s}}} \right) \right)
$$

$$
= \sum_{\substack{(x_{i_1}, \ldots x_{i_s}, x_4, x_{h_1}, \ldots, x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} \left(\mathbb{E}\left(\partial^\gamma t_{x_{i_1}} \ldots t_{x_{i_s}} t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}} \right) \right.
$$

$$
\left. \times \mathbb{E}\left(\sum_{\substack{(x_{f_1}, \ldots, x_{f_{3-s}}) \\ x_{f_1} + \ldots + x_{f_{3-s}} = n - x_{i_1} - \ldots - x_{i_s} - x_4 \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{f_1}, \ldots, x_{f_{3-s}} \text{ are distinct}}} t_{x_{f_1}} \ldots t_{x_{f_{3-s}}} \right) \right). \tag{21}
$$

By using Lemma 5 we have

$$
\mathbb{E}\left(\sum_{\substack{(x_{f_1}, \ldots, x_{f_{3-s}}) \\ x_{f_1} + \ldots + x_{f_{3-s}} = n - x_{i_1} - \ldots - x_{i_s} - x_4 \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{f_1}, \ldots, x_{f_{3-s}} \text{ are distinct}}} t_{x_{f_1}} \ldots t_{x_{f_{3-s}}} \right) = O(1).
$$

It follows that (21) is equal to

$$
O\left(\mathbb{E}\left(\partial^\gamma \sum_{\substack{(x_{i_1}, \ldots x_{i_s}, x_4, x_{h_1}, \ldots x_{h_u}) \\ x_4 + d_{i_1} x_{i_1} + \ldots + d_{i_s} x_{i_s} + b_{h_1} x_{h_1} + \ldots + b_{h_u} x_{h_u} = 0 \\ x_{i_j} < x_4, j = 1, \ldots, s; x_{h_{j'}} < x_4, j' = 1, \ldots, u \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} t_{x_{i_1}} \ldots t_{x_{i_s}} t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}} \right) \right).
$$

Let $\underline{\gamma} = (\gamma_1, \ldots, \gamma_n)$, $|\underline{\gamma}| = w$, $\gamma_{l_1} = \ldots = \gamma_{l_w} = 1$, $1 \leqslant l_1 < \ldots < l_w \leqslant n$. If $\{v_1, \ldots, v_w\} \subset \{h_1, \ldots, h_u\}$, then let $\{e_1, \ldots, e_{u-w}\} = \{h_1, \ldots, h_u\} \setminus \{v_1, \ldots, v_w\}$. Thus we have

$$
\mathbb{E}\left(\partial^{\underline{\gamma}} \sum_{\substack{(x_{i_1}, \ldots x_{i_s}, x_4, x_{h_1}, \ldots x_{h_u}) \\ x_4 + d_{i_1}x_{i_1} + \ldots + d_{i_s}x_{i_s} + b_{h_1}x_{h_1} + \ldots + b_{h_u}x_{h_u} = 0 \\ x_{i_j} < x_4, j=1,\ldots,s; x_{h_{j'}} < x_4, j'=1,\ldots,u}} t_{x_{i_1}} \ldots t_{x_{i_s}} t_{x_4} t_{x_{h_1}} \ldots t_{x_{h_u}} \right)
$$

$$
= \sum_{\{v_1, \ldots v_w\} \subset \{h_1, \ldots, h_u\}} \sum_{\substack{\pi \\ (x_{v_1}, \ldots, x_{v_w}) = \pi(l_1, \ldots, l_w)}} \tag{22}
$$

$$
\times \mathbb{E}\left(\sum_{\substack{(x_{i_1}, \ldots, x_{i_s}, x_4, x_{e_1} \ldots x_{e_{u-w}}) \\ x_4 + d_{i_1}x_{i_1} + \ldots + d_{i_s}x_{i_s} + b_{h_1}x_{h_1} + \ldots + b_{h_u}x_{h_u} = 0 \\ x_{i_j} < x_4, j=1,\ldots,s; x_{h_{j'}} < x_4, j'=1,\ldots,u \\ x_{i_1}, \ldots, x_{i_s}, x_4, x_{h_1}, \ldots, x_{h_u} \text{ are distinct}}} t_{x_{i_1}} \ldots t_{x_{i_s}} t_{x_4} t_{x_{e_1}} \ldots t_{x_{e_{u-w}}} \right),
$$

where $\pi(l_1, \ldots, l_w)$ denotes the permutations of $(l_1, \ldots, l_w)$.

Since $s + u + 1$ is the number of the variables in the equation $x_4 + d_{i_1}x_{i_1} + \ldots + d_{i_s}x_{i_s} + b_{h_1}x_{h_1} + \ldots + b_{h_u}x_{h_u} = 0$, it is clear, that $s + u + 1$ is equal to 3 or 4. So $s + u + 1 - w \leqslant 3$, except if $s + u + 1 = 4$ and $w = 0$. If $s + u + 1 - w \leqslant 3$ then we may use Lemma 4 again, which implies (20). So we can only show the case $s + u + 1 = 4$ and $w = 0$, which is equivalent to $\underline{\gamma} = 0$. To do this, it remains to prove the following Lemma.

Lemma 6. *The following expectation are bounded.*

1. $\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1+x_2+x_3+x_4=n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \left(\sum_{\substack{(x_5, x_6, x_7) \\ x_4+x_5=x_6+x_7}} t_{x_5} t_{x_6} t_{x_7} \right) \right) = O(1)$

2. $\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1+x_2+x_3+x_4=n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \left(\sum_{\substack{(x_5, x_6) \\ x_4+x_3=x_5+x_6}} t_{x_5} t_{x_6} \right) \right) = O(1)$

3. $\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1+x_2+x_3+x_4=n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \left(\sum_{\substack{(x_5, x_6) \\ x_4+x_5=x_3+x_6}} t_{x_5} t_{x_6} \right) \right) = O(1)$

4. $\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1+x_2+x_3+x_4=n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \left(\sum_{\substack{x_5 \\ x_4+x_3=x_2+x_5}} t_{x_5} \right) \right) = O(1)$

5. $\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1+x_2+x_3+x_4=n}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \left(\sum_{\substack{x_5 \\ x_4+x_5=x_2+x_3}} t_{x_5} \right) \right) = O(1)$

6. $\mathbb{E}\left(\sum_{\substack{(x_1, x_2, x_3, x_4) \\ x_1+x_2+x_3+x_4=n \\ x_4+x_3=x_2+x_1}} t_{x_1} t_{x_2} t_{x_3} t_{x_4} \right) = O(1)$

Proof.

(1) Using the definition of the variables t_i's get that the order of this expectation is

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{n-x_1-x_2-1} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{5}{7}}\right.$$

$$\left. \sum_{x_5=1}^{x_4-1} x_5^{-\frac{5}{7}} \sum_{x_6=1}^{x_4+x_5-1} x_6^{-\frac{5}{7}} (x_4 + x_5 - x_6)^{-\frac{5}{7}}\right). \quad (23)$$

Applying Lemma 4 to the last sum it follows that (23) is equivalent to the following

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{n-x_1-x_2-1} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{5}{7}}\right.$$

$$\left. \sum_{x_5=1}^{x_4-1} x_5^{-\frac{5}{7}} (x_4 + x_5)^{-\frac{3}{7}}\right). \quad (24)$$

Since

$$\sum_{x_5=1}^{x_4-1} x_5^{-\frac{5}{7}} (x_4 + x_5)^{-\frac{3}{7}} = O\left(x_4^{-\frac{3}{7}} \sum_{x_5=1}^{x_4-1} x_5^{-\frac{5}{7}}\right) = O\left(x_4^{-\frac{1}{7}}\right)$$

$$= O\left((n - x_1 - x_2 - x_3)^{-\frac{1}{7}}\right),$$

it follows that (24) is equal to

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{n-x_1-x_2-1} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{6}{7}}\right)$$

$$= O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} (n - x_1 - x_2)^{-\frac{4}{7}}\right)$$

$$= O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} (n - x_1)^{-\frac{2}{7}}\right) = O(1). \quad (25)$$

(2) Using again the definition of the variables t_i's this expression is equivalent with the following

$$\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{n-x_1-x_2-1} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{5}{7}}$$

$$\sum_{x_5=1}^{x_4+x_3-1} x_5^{-\frac{5}{7}} (x_4 + x_3 - x_5)^{-\frac{5}{7}}. \quad (26)$$

Since

$$\sum_{x_5=1}^{x_4+x_3-1} x_5^{-\frac{5}{7}}(x_4+x_3-x_5)^{-\frac{5}{7}} = O\left((x_4+x_3)^{-\frac{3}{7}}\right) = O\left(x_4^{-\frac{3}{7}}\right) = O\left(x_4^{-\frac{1}{7}}\right)$$

$$= O\left((n-x_1-x_2-x_3)^{-\frac{1}{7}}\right),$$

it follows that (26) is equal to

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{n-x_1-x_2-1} x_3^{-\frac{5}{7}}(n-x_1-x_2-x_3)^{-\frac{6}{7}}\right).$$

It follows from (25) that this is equal to $O(1)$.

(3) Since $x_3 < x_4$, thus $x_3 < \frac{x_3+x_4}{2} = \frac{n-x_1-x_2}{2}$. This expectation is

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{\frac{n-x_1-x_2}{2}} x_3^{-\frac{5}{7}}(n-x_1-x_2-x_3)^{-\frac{5}{7}}\right.$$

$$\left. \times \sum_{x_5=1}^{x_4-1} x_5^{-\frac{5}{7}}(x_4+x_5-x_3)^{-\frac{5}{7}}\right), \quad (27)$$

where

$$\sum_{x_5=1}^{x_4-1} x_5^{-\frac{5}{7}}(x_4+x_5-x_3)^{-\frac{5}{7}} = \sum_{x_5=1}^{x_4-x_3} x_5^{-\frac{5}{7}}(x_4+x_5-x_3)^{-\frac{5}{7}}$$

$$+ \sum_{x_5=x_4-x_3+1}^{x_4-1} x_5^{-\frac{5}{7}}(x_4+x_5-x_3)^{-\frac{5}{7}}$$

$$= O\left((x_4-x_3)^{-\frac{5}{7}} \sum_{x_5=1}^{x_4-x_3} x_5^{-\frac{5}{7}}\right)$$

$$+ O\left(\int_{x_4-x_3}^{\infty} x^{-\frac{10}{7}}\,dx\right)$$

$$= O\left((x_4-x_3)^{-\frac{3}{7}}\right).$$

Equation (27) is equal to

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{n-x_1} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{\frac{n-x_1-x_2}{2}} (2x_3)^{-\frac{5}{7}}(n-x_1-x_2-2x_3)^{-\frac{6}{7}}\right).$$

It follows from (25) that this is equal to $O(1)$.

(4) Note, that here $n = x_1 + 2x_2 + x_5$. This expectation is

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{\frac{n-x_1}{2}} x_2^{-\frac{5}{7}} \sum_{x_3=1}^{\frac{n-x_1-x_2}{2}} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{5}{7}}\right.$$

$$\left. \times (x_4 + x_3 - x_2)^{-\frac{5}{7}}\right), \quad (28)$$

where

$$(x_4 + x_3 - x_2)^{-\frac{5}{7}} = (n - x_1 - x_2 - x_2)^{-\frac{5}{7}} = (n - x_1 - 2x_2)^{-\frac{5}{7}}.$$

Equation (28) is equal to

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{\frac{n-x_1}{2}} x_2^{-\frac{5}{7}} (n - x_1 - 2x_2)^{-\frac{5}{7}}\right.$$

$$\left. \times \sum_{x_3=1}^{\frac{n-x_1-x_2}{2}} x_3^{-\frac{5}{7}} (n - x_1 - x_2 - x_3)^{-\frac{5}{7}}\right)$$

$$= O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{\frac{n-x_1}{2}} x_2^{-\frac{5}{7}} (n - x_1 - 2x_2)^{-\frac{5}{7}} (n - x_1 - x_2)^{-\frac{3}{7}}\right)$$

$$= O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_2=1}^{\frac{n-x_1}{2}} (2x_2)^{-\frac{5}{7}} (n - x_1 - 2x_2)^{-\frac{5}{7}}\right)$$

$$= O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} (n - x_1)^{-\frac{3}{7}}\right) = O\left(n^{-\frac{1}{7}}\right) = O(1).$$

(5) It is clear that the $n = x_1 + x_2 + x_3 + x_4$, $x_4 + x_5 = x_2 + x_3$ equation-system is equivalent with $n = x_1 + 2x_4 + x_5$, $x_4 + x_5 = x_2 + x_3$. Thus this expectation is

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_4=1}^{\frac{n-x_1}{2}} x_4^{-\frac{5}{7}} (n - x_1 - 2x_4)^{-\frac{5}{7}} \sum_{x_3=1}^{x_4+x_5-1} x_3^{-\frac{5}{7}} (x_4 + x_5 - x_3)^{-\frac{5}{7}}\right),$$

$$(29)$$

$$O\left(\sum_{x_3=1}^{x_4+x_5-1} x_3^{-\frac{5}{7}} (x_4 + x_5 - x_3)^{-\frac{5}{7}}\right) = O\left((x_4 + x_5)^{-\frac{3}{7}}\right) = O(1).$$

So (29) is equal to

$$O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}} \sum_{x_4=1}^{\frac{n-x_1}{2}} (2x_4)^{-\frac{5}{7}}(n-x_1-2x_4)^{-\frac{5}{7}}\right) = O\left(\sum_{x_1=1}^{n} x_1^{-\frac{5}{7}}(n-x_1)^{-\frac{3}{7}}\right)$$

$$= O\left(n^{-\frac{1}{7}}\right) = O(1).$$

(6) It is clear that the $n = x_1+x_2+x_3+x_4$, $x_1+x_2 = x_3+x_4$ equation-system is equivalent with $\frac{n}{2} = x_1+x_2 = x_3+x_4$. Thus this expectation is

$$O\left(\sum_{x_1=1}^{\frac{n}{2}} x_1^{-\frac{5}{7}}\left(\frac{n}{2}-x_1\right)^{-\frac{5}{7}} \sum_{x_3=1}^{\frac{n}{2}} x_3^{-\frac{5}{7}}\left(\frac{n}{2}-x_1\right)^{-\frac{5}{7}}\right) \tag{30}$$

$$= O\left(\sum_{x_1=1}^{\frac{n}{2}} x_1^{-\frac{5}{7}}\left(\frac{n}{2}-x_1\right)^{-\frac{5}{7}} n^{-\frac{3}{7}}\right) = O\left(n^{-\frac{6}{7}}\right) = O(1). \quad \blacksquare$$

So from (16) we get that

$$\mathbb{E}_{\geqslant 0}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}}) = \max_{d'\geqslant 0}\mathbb{E}_{d'}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}}) = O(1),$$

and

$$\mathbb{E}_{\geqslant 1}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}}) = \max_{d'\geqslant 1}\mathbb{E}_{d'}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}}) = O(1).$$

In Lemma 2, we have $k \leqslant 7$ and $\lambda = 32\log n$. Thus

$$\mathbb{P}\left(\left|Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}} - \mathbb{E}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}})\right|\right.$$

$$\geqslant C_k(32\log n)^{k-\frac{1}{2}}\sqrt{\mathbb{E}_{\geqslant 0}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}})\mathbb{E}_{\geqslant 1}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}})}\right)$$

$$= O_k\left(e^{-8\log n + (k-1)\log n}\right) = O_k\left(\frac{1}{n^2}\right). \tag{31}$$

Thus by (16) and Lemma 2 we get that with probability 1,

$$Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}} = \mathbb{E}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}})$$

$$+ O\left((\log n)^{k-\frac{1}{2}}\sqrt{\mathbb{E}_{\geqslant 0}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}})\mathbb{E}_{\geqslant 1}(Y_{d_{i_1},\dots,d_{i_s},b_{h_1},\dots,b_{h_u}})}\right)$$

$$= O\left((\log n)^{6.5}\right),$$

which shows (16), which proves (7). And this completes the proof.

References

[1] J. Cilleruelo, *Sidon basis*, arXiv: 1304.5351.
[2] J.M. Deshouillers, A. Plagne, *A Sidon basis*, Acta Mathematica Hungarica **123** (2009), 233–238.

[3] P. Erdős, A. Sárközy, V.T. Sós, *On additive properties of general sequences*, Discrete Mathematics **136** (1994), 75–99.

[4] P. Erdős, A. Sárközy, V.T. Sós, *On sum sets of Sidon sets I.*, Journal of Number Theory **47** (1994), 329–347.

[5] G. Grekos, L. Haddad, C. Helou, J. Pihko, *Representation functions, Sidon sets and bases*, Acta Arithmetica **130**(2) (2007), 149–156.

[6] H. Halberstam, K.F. Roth, *Sequences*, Springer–Verlag, New York, 1983.

[7] J.H. Kim and V.H. Vu, *Concentration of multivariate polynomials and its applications*, Combinatorica **20** (2000), 417–434.

[8] S.Z. Kiss, *On Sidon sets which are asymptotic bases*, Acta Mathematica Hungarica **128** (2010), 46–58.

[9] M.B. Nathanson, *Additive Number Theory The Classical Bases*, Springer, 1996

[10] A. Sárközy, *Unsolved problems in number theory*, Periodica Mathematica Hungarica **42**(1-2) (2001), 17–35.

[11] T. Tao, V.H. Vu, *Additive Combinatorics*, Cambridge University Press, 2006.

[12] V.H. Vu, *On the concentration of multi-variate polynomials with small expectation*, Random Structures and Algorithms **16** (2000), 344–363.

[13] V.H. Vu, *Chernoff type bounds for sum of dependent random variables and applications in additive number theory*, Number theory for the millennium, III (Urbana, IL, 2000), 341–356, A K Peters, Natick, MA, 2002.

Addresses: Sándor Z. Kiss: Institute of Mathematics, Budapest University of Technology and Economics, MTA-BME Lendület Future Internet Research Group, H-1529 B.O. Box, Hungary; Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest H-1111, Lágymányosi street 1,1 Hungary;
Eszter Rozgonyi and Csaba Sándor: Institute of Mathematics, Budapest University of Technology and Economics, H-1529 B.O. Box, Hungary.

E-mail: kisspest@cs.elte.hu, reszti@math.bme.hu, csandor@math.bme.hu