IMPROVED BOUNDS FOR
THE TWO-POINT LOGARITHMIC CHOWLA CONJECTURE

CÉDRIC PILATTE

ABSTRACT. Let λ be the Liouville function, defined as $\lambda(n) := (-1)^{\Omega(n)}$ where $\Omega(n)$ is the number of prime factors of n with multiplicity. In 2021, Helfgott and Radziwiłł proved that

$$\sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n+1) \ll \frac{\log x}{(\log\log x)^{1/2}},$$

improving earlier results by Tao and Teräväinen. We prove that

$$\sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n+1) \ll (\log x)^{1-c}$$

for some absolute constant $c > 0$. This appears to be best possible with current methods.

CONTENTS

1. INTRODUCTION

1.1. **Background.** Let $\lambda : \mathbb{N} \to \{-1, +1\}$ be the Liouville function, defined by $\lambda(n) := (-1)^{\Omega(n)}$ where $\Omega(n)$ is the number of prime factors of n, counted with multiplicity. Its statistical properties are closely connected with the distribution of primes. Indeed, the bounds $\frac{1}{x} \sum_{n \leqslant x} \lambda(n) = o_{x \to \infty}(1)$ and

$$\frac{1}{x} \sum_{n \leqslant x} \lambda(n) \ll_{\varepsilon} x^{1/2 + \varepsilon}$$

are equivalent to the Prime Number Theorem and the Riemann Hypothesis respectively, by elementary arguments. These two examples are consistent with the Liouville pseudorandomness principle, a heuristic which suggests that λ should statistically behave like a sequence of independent random variables taking the values -1 and $+1$ with probability $1/2$.

For higher-degree correlations, a well-known conjecture of Chowla [2] asserts that, for any $k \geqslant 1$ and distinct integers $h_1, \ldots, h_k$, one has

$$(1) \qquad \frac{1}{x} \sum_{n \leqslant x} \lambda(n + h_1)\lambda(n + h_2) \cdots \lambda(n + h_k) = o_{x \to \infty}(1).$$

This can be regarded as a multiplicative analogue of the Hardy-Littlewood prime k-tuple conjecture, which predicts an asymptotic formula for correlations of the von Mangoldt function Λ. Chowla's conjecture is subject to the *parity problem*, a major obstacle in analytic number theory (see [3, Section 16.4] for more details). It is open for all $k \geqslant 2$.

Yet, in recent years, remarkable progress has been made on weaker variants of Chowla's conjecture.

In 2015, Matomäki, Radziwiłł and Tao proved that (1) holds *on average* over $h_1, \ldots, h_k$, for every fixed $k \geqslant 2$ [9]. A crucial ingredient in their proof was the groundbreaking work by Matomäki and Radziwiłł [8] on sums of multiplicative functions over short intervals.

One year later, Tao proved a *logarithmic* version of Chowla's conjecture for $k = 2$ [15]. This means that the regular average $\frac{1}{x} \sum_{n \leqslant x} f(n)$ is replaced with the logarithmic average $\frac{1}{\log x} \sum_{n \leqslant x} \frac{1}{n} f(n)$. Fixing $(h_0, h_1) = (0, 1)$ for simplicity, Tao's result thus reads

$$(2) \qquad \frac{1}{\log x} \sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n + 1) = o_{x \to \infty}(1).$$

Tao's proof [15], which used a novel *entropy decrement argument*, was a key step in his resolution of the Erdős discrepancy problem [14]. From his paper [15], it is possible (see [6]) to extract the explicit bound

$$(3) \qquad \sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n + 1) \ll \frac{\log x}{(\log \log \log \log x)^{1/5}}.$$

The logarithmic version of Chowla's conjecture (1) was later proved for all *odd* $k \geqslant 3$, by Tao and Teräväinen [18]. The two authors gave a different proof of that result in [16]. For *even* $k \geqslant 4$, the logarithmically averaged Chowla conjecture is still open. The methods of their paper [16] can be used to obtain the following quantitative refinement of (3): for some small absolute constant $c > 0$,

$$(4) \qquad \sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n + 1) \ll \frac{\log x}{(\log \log \log x)^c}.$$

In 2021, Helfgott and Radziwiłł [5] proved the substantial quantitative improvement

$$(5) \qquad \sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n + 1) \ll \frac{\log x}{(\log \log x)^{1/2}}.$$

They used a very different combinatorial approach, studying the eigenvalues of a certain weighted graph defined in terms of divisibility by small primes. A high-level exposition of their proof is given by Helfgott [4].

In this paper, we improve the approach of Helfgott and Radziwiłł [5] to prove the following.

Theorem 1.1 (Logarithmic two-point Chowla correlations). *For some absolute constant $c > 0$,*

$$\sum_{n \leqslant x} \frac{1}{n} \lambda(n)\lambda(n + 1) \ll (\log x)^{1-c}.$$

It appears that saving a fixed power of the logarithm is the best that is achievable with current techniques. Ultimately, our proof relies on the work of Matomäki and Radziwiłł [8] on multiplicative functions in short intervals, where the current state of the art only allows to save a small power of

$\log x$. The exploitation of multiplicativity using an idea of Tao [15] also separately appears to limit our saving to a small power of $\log x$, because a typical integer has $O(\log x)$ divisors.

The methods of this paper should generalise to a wider class of multiplicative functions through appropriate modifications. The complete multiplicativity of λ is only used in Proposition 2.6 and in the proof that Theorem 2.1 implies Theorem 1.1. The only other property of λ we use is that it is 1-bounded (for Sections 2 and 3), but a weaker ℓ^p bound would suffice.

Our proof also yields an improved bound for the *unweighted* two-point correlations (i.e. without logarithmic averaging) at *almost all scales*, see Remark 2.5.

1.2. **Proof outline.** In this section, we give a very short description of the overall strategy. Fuller explanations are given along the way, at various points in the paper.

Using the multiplicativity of λ, Tao [15] showed that the problem of bounding $\sum_{n \leqslant x} \frac{1}{n} \lambda(n) \lambda(n+1)$ reduces to bounding

$$\mathbb{E}_{p \in P} \sum_{n \leqslant x} \lambda(n) \lambda(n+p) \left(\mathbf{1}_{p|n} - \frac{1}{p} \right)$$

where $P \subset [1, \exp(\sqrt{\log x})]$ is a set of primes.

Helfgott and Radziwiłł [5] interpreted the above expression as the matrix product $\boldsymbol{\lambda}^\top A \boldsymbol{\lambda}$ where $\boldsymbol{\lambda} := (\lambda(1), \ldots, \lambda(x))^\top$ and A is the matrix with entries

$$A_{mn} := \begin{cases} \mathbf{1}_{p|n} - \frac{1}{p} & \text{if } |m - n| = p \in P, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, it is sufficient to bound the eigenvalues of the matrix A, or the eigenvalues of its restriction $A|_X$ to some very dense subset $X \subset \{1, \ldots, x\}$. Using a high trace method, Helfgott and Radziwiłł [5] managed to obtain the bound $\left(\sum_{p \in P} 1/p \right)^{1/2+o(1)}$ for the largest eigenvalue of such a matrix, which is essentially the best possible. Since $\sum_{p \in P} 1/p \ll \log \log x$, this approach cannot yield a saving better than a power of $\log \log x$ over the trivial bound for two-point Chowla correlations.

In our new approach, we replace the average over primes $p \in P$ with an average over integers $d = p_1 \cdots p_k$ that are products of k primes, where $k \asymp \log \log x$. By Tao's argument, we need to bound

$$\mathbb{E}_{d \in D} \sum_{n \leqslant x} \lambda(n) \lambda(n+d) \prod_{p|d} \left(\mathbf{1}_{p|n} - \frac{1}{p} \right)$$

where D is a set of integers with k prime factors. Following the strategy of Helfgott and Radziwiłł [5], it is sufficient to bound the eigenvalues of the matrix $\widetilde{A}|_X$ where $\widetilde{A}$ is the matrix defined by

$$\widetilde{A}_{mn} := \begin{cases} \prod_{p|d} \left(\mathbf{1}_{p|n} - \frac{1}{p} \right) & \text{if } |m - n| = d \in D, \\ 0 & \text{otherwise;} \end{cases}$$

and X is a large subset of $\{1, \ldots, x\}$. We prove that all eigenvalues of $\widetilde{A}|_X$ are $\leqslant \left(\sum_{d \in D} 1/d \right)^{2/3+o(1)}$. Since $k \gg \log \log x$, this is $\gg (\log x)^c$, which produces the exponential improvement in Theorem 1.1.

Unfortunately, working with products of multiple primes rather than single primes introduces new difficulties throughout the argument. It is handling all of these new difficulties which is the key new contribution of our work. We are forced to rework and generalise all the arguments of [5] with the result that our paper is essentially self-contained. One particular new difficulty is in Section 9 where we wish to bound the number of solutions to systems of divisibility constraints. In the prior work this was a linear system, and so could be bounded by a simple lattice point argument. In

our situation this now becomes a polynomial system, and to handle this we require a much more involved argument based on the structure of what we call 'unpredictable words'.

Acknowledgements

1.3. **Structure of the paper.** We now give a broad overview of the structure of the paper. The reader may wish to refer to Fig. 1, which depicts the main propositions of the paper along with their logical dependencies. The paper is designed to be as self-contained as possible. In particular, no prior familiarity with [5] is needed.

FIGURE 1. Dependency graph for the proof of Theorem 1.1 (main propositions only).

In Section 2, we state our main technical estimate, Theorem 2.1. We then reproduce some clever manipulations due to Tao [15] to show how Theorem 2.1 implies our bound for two-point logarithmic Chowla correlations, Theorem 1.1. The first step towards the proof of Theorem 2.1 is Proposition 2.6, which replaces the double sum in Theorem 2.1 with a more convenient 'balanced' version. The proof uses an exponential sum estimate of Matomäki, Radziwiłł and Tao [9].

In Section 3, we begin to implement the elegant strategy of Helfgott and Radziwiłł [5]. The key linear algebra ingredient is Lemma 3.4 on eigenvalues of near-diagonal matrices. It is the same as [5, Proposition 2.4], but we give a very short proof using Cauchy's interlacing theorem.

Certain technical reasons prevent us from working with the matrix $\widetilde{A}$ defined in the previous section, which has some overly large eigenvalues. Proposition 3.5 is the claim that there exists a slight perturbation of $\widetilde{A}$ that does not have any large eigenvalues. The construction of this modification of $\widetilde{A}$ is given in Section 5, following Section 4 which provides some motivation and explanation of the general strategy.

The proof that this new matrix satisfies a suitable high moment bound occupies Sections 6 to 11.

1.4. **Symbols and notations.** For ease of reading, we have provided a table showing the main parameters, their size and a reference to where they are introduced.

Parameter	Size properties	First appearance
ε_1	$\varepsilon_1 > 0$ sufficiently small	Theorem 2.1
H	H tending to $+\infty$	Theorem 2.1
J	$1 \leqslant J \leqslant \varepsilon_1^2 \log\log H$	Theorem 2.1
H_0	$H_0 = \exp\left((\log H)^{1-\varepsilon_1}\right)$	Theorem 2.1
V_i	$V_i = \sum_{p\in\mathcal{P}_i} 1/p$	Theorem 2.1
V	$V = \max_i V_i$	Lemma 2.4
N	$N \geqslant \exp\left((\log H)^3\right)$	Theorem 2.1
K	$K = 2\lfloor\log H\rfloor$	Proposition 3.5
L	$L = K^{1-10\varepsilon_1}$	Definition 5.2

The following table contains most of the other symbols used repeatedly in the paper.

Notation	Properties	First appearance
$\mathcal{P}_j$	disjoint sets of primes $\subset (H_0, H)$	Theorem 2.1
$\mathcal{P}$	$\mathcal{P} := \bigcup_j \mathcal{P}_j$	Notation 2.7
$\mathcal{D}$	set of all products $\prod_{j\in[\![J]\!]} p_j$ with $p_j \in \mathcal{P}_j$	Notation 2.7
I_N	$\mathbb{N} \cap (N, 2N]$	Notation 2.7
G_0	weighted graph on I_N	Lemma 3.1
$\mathbf{D}_R$	set of $\boldsymbol{d} \in (\pm\mathcal{D})^R$ such that $\sum_i d_i = 0$	Definition 3.6
b_i	partial sums $b_i := \sum_{i' < i} d_{i'}$	Definition 3.6
W	smooth weight supported on $[0, 2JV]$	Definition 5.1
$\mathcal{Y}$	set of all prohibited progressions	Definition 5.3
Y_L	complement of union of all prohibited progressions	Definition 5.3
G	weighted graph on I_N	Definition 5.4
$w_{\boldsymbol{d}}(n)$	weight of a closed walk	Equation (30)
$\mathbf{n}$	random variable uniformly distributed in $\prod_{p\in\mathcal{P}} \mathbb{Z}/p\mathbb{Z}$	Definition 6.1
d_{ij}	unique prime in $\mathcal{P}_j$ dividing d_i	Definition 6.3
$\rho_{\boldsymbol{d}}$	product of all d_{ij}	Definition 6.3
$\rho_{\boldsymbol{d};I}$	product of all d_{ij} with $(i,j) \in I$	Definition 6.3
$\mathcal{S}$	set of single indices	Definition 6.6
$\mathbf{D}_R^{\mathcal{S}}$	set of $\boldsymbol{d} \in \mathbf{D}_R$ with set of single indices $\mathcal{S}$	Definition 6.6
$\mathcal{L}, \mathcal{U}$	sets of lit and unlit indices	Definition 6.8
$\widetilde{\boldsymbol{d}}$	reduced walk	Definition 6.11
$\widetilde{R}$	length of $\widetilde{\boldsymbol{d}}$ if $\boldsymbol{d} \in \mathbf{D}_R$	Definition 6.11
$\mathbf{D}_R^{\mathcal{S},\mathcal{L}}$	set of $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$ satisfying lit indices conditions	Definition 6.15
$\mathcal{S}_{\text{bad}}(\boldsymbol{d})$	set of bad single indices	Definition 7.2
q_R	modulus of the arithmetic progressions R	Proposition 7.3
$A_{\boldsymbol{d}}$	arithmetic progression determined by lit indices	Lemma 7.6
$\mathcal{W}$	set of words with no two identical adjacent letters	Definition 8.1
$\mathcal{W}^{\neq}$	set of words with distinct letters	Definition 8.1
$\widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$	non-backtracking walks $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S},\mathcal{L}}$	Definition 8.6
$v_{j,\boldsymbol{d}}, w_{j,\boldsymbol{d}}$	words associated to $\boldsymbol{d}$, with letters in $\mathcal{P}_j$	Definition 8.7
$\mathbf{P}_R$	set of $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ such that all $w_{j,\boldsymbol{d}}$ are predictable	Definition 8.7
$\mathbf{U}_R$	set of $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ such that some $w_{j,\boldsymbol{d}}$ is unpredictable	Definition 8.7
$C_{I,i_0,j_0,\kappa}(\boldsymbol{d})$	the constraint on $\boldsymbol{d}$ with parameters I, i_0, j_0, z	Definition 9.2
$(J_{\mathcal{N}}, J_{\mathcal{L}}, J_{\mathcal{U}})$	type of an extension	Definition 10.1
τ_h	cyclic permutation with shift h	Definition 10.4

We write $f \ll g$ or $f = O(g)$ if $|f| \leqslant Cg$ for some absolute constant $C > 0$. The notation $f \asymp g$ means that $f \ll g$ and $g \ll f$.

If $a, b \in \mathbb{Z}$, we write $[\![a, b]\!] := \mathbb{Z} \cap [a, b]$ ($= \emptyset$ if $a > b$), and we call a set of this form a *discrete interval*. Its length, or size, is its cardinality ($b - a + 1$ if $a \leqslant b$). For $n \in \mathbb{N}$, we write $[\![n]\!] := \{1, 2, \ldots, n\}$.

If $A \subset \mathbb{R}$, we write $\pm A$ for $\{\sigma a : \sigma \in \{\pm 1\}, a \in A\}$.

In this paper, the term *arithmetic progression* always refers to a 'two-sided infinite' arithmetic progression of the form $a + q\mathbb{Z}$ for some $a \in \mathbb{Z}$ and $q \in \mathbb{N}$.

For $n \in \mathbb{Z}$, we write $\omega(n)$ for the number of distinct prime factors of n. If $\mathcal{P}$ is a set of primes, we let $\omega_{\mathcal{P}}(n)$ be the number of primes in $\mathcal{P}$ that divide n.

Euler's totient function and the divisor sum function are denoted by φ and σ_1, respectively.

A *weighted graph* is a pair (V, w) where V is a set (the vertex set) and $w : V \times V \to \mathbb{C}$ a function ($w(v_1, v_2)$ is the weight of the edge (v_1, v_2)). Thus, we use weight zero edges instead of 'non-existent' edges.

2. MAIN THEOREM, CONSEQUENCES AND REFORMULATIONS

2.1. **Statement of the underlying main theorem.** Our bound for the two-point logarithmic Chowla correlations is a consequence of the following key estimate. To formulate it, we need to define a certain number of parameters.

Theorem 2.1. *Let $\varepsilon_1 > 0$ be a sufficiently small absolute constant. Let $H > 0$ be sufficiently large in terms of ε_1. Let J be a positive integer with $J \leqslant \varepsilon_1^2 \log \log H$, and let $H_0 = \exp\left((\log H)^{1-\varepsilon_1}\right)$.*

Let $C := \exp\left(\varepsilon_1 (\log \log H)/(2J)\right)$. For $1 \leqslant i \leqslant J$, let $\mathcal{P}_i$ be the set of all primes p with

$$C^{2i-2} < \frac{\log p}{\log H_0} < C^{2i-1}$$

and let $V_i := \sum_{p \in \mathcal{P}_i} \frac{1}{p}$.

Let N be an integer such that $\log N \geqslant (\log H)^3$. Then

$$(6) \qquad \sum_{(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J} \sum_{\substack{n \in (N, 2N] \\ p_1 \cdots p_J | n}} \lambda(n) \lambda(n + p_1 \cdots p_J) \ll (V_1 \cdots V_J)^{3/4} N.$$

Remark 2.2. Theorem 2.1 should be compared with the trivial bound $S_1 \ll V_1 \cdots V_J N$.

We stated Theorem 2.1 with the constant $3/4$ in (6), but our proof works for any exponent $> 2/3$. In principle, this exponent could be improved to $1/2 + o(1)$. However, a proof of this would involve combinatorial complications and would not significantly improve the constant c in Theorem 1.1 (which is unspecified anyway).

The lower bound for N in terms of H can be somewhat relaxed, but the proof definitely requires something like $\log N \geqslant (\log H)^{2+o(1)}$.

Remark 2.3. The techniques of this paper actually show the slightly stronger result

$$(7) \qquad \sum_{(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J} \left| \sum_{\substack{n \in (N, 2N] \\ p_1 \cdots p_J | n}} \lambda(n) \lambda(n + p_1 \cdots p_J) \right| \ll (V_1 \cdots V_J)^{3/4} N.$$

To obtain this, all that is required is to reiterate the entire proof, allowing for arbitrary coefficients $c_{p_1, \ldots, p_J} \in \{\pm 1\}$ throughout. No other modifications are necessary, and the result becomes

$$\sum_{(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J} c_{p_1, \ldots, p_J} \sum_{\substack{n \in (N, 2N] \\ p_1 \cdots p_J | n}} \lambda(n) \lambda(n + p_1 \cdots p_J) \ll V^{3J/4} N,$$

from which (7) follows. For the sake of brevity and readability, we will refrain from presenting a detailed proof of (7). Instead, we will concentrate on the seemingly weaker estimate given in Theorem 2.1, which omits absolute values on the left-hand side. In any case, we will see in the next section that Theorem 2.1 suffices to prove Theorem 1.1, which is our primary motivation.

We now prove some technical estimates that will be useful throughout the paper.

Lemma 2.4 (Bounds related to the sets $\mathcal{P}_i$). *Let $\varepsilon_1 > 0$ be a sufficiently small constant. Let $H > 0$ be sufficiently large in terms of ε_1. Let $1 \leqslant J \leqslant \varepsilon_1^2 \log \log H$, and let $H_0 = \exp\left((\log H)^{1-\varepsilon_1}\right)$.*

Let $\mathcal{P}_1, \ldots, \mathcal{P}_J$ be as in Theorem 2.1. Let $V_i := \sum_{p \in \mathcal{P}_i} \frac{1}{p}$ and define $V := \max_{i \in [\![J]\!]} V_i$. Then

(a) *$\mathcal{P}_1, \ldots, \mathcal{P}_J$ are disjoint subsets of (H_0, H),*

(b) *$V_1 V_2 \cdots V_J = (1 + o(1))V^J = (1 + o(1)) \left(\dfrac{\varepsilon_1 \log \log H}{2J}\right)^J$,*

(c) *if $(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J$, then $p_1 p_2 \cdots p_i < p_{i+1}^{1/10}$ for all $1 \leqslant i < J$, and $p_1 p_2 \cdots p_J < H$.*

In particular, $V^J \ll (\log H)^{\varepsilon_1^2 \log(\varepsilon_1^{-1})}$. If, moreover, $J \geqslant \frac{\varepsilon_1^2}{2} \log \log H$, then $V^J \gg (\log H)^{\frac{\varepsilon_1^2}{2} \log(\varepsilon_1^{-1})}$.

Proof. Let $C := \exp\left(\varepsilon_1 (\log \log H)/(2J)\right) \geqslant 20$, so that $\mathcal{P}_i$ is the set of all primes in the interval

$$\left(\exp\left(C^{2i-2} \log H_0\right), \ \exp\left(C^{2i-1} \log H_0\right)\right).$$

Property (a) is clear. By Mertens' second estimate, we have

$$(8) \qquad\qquad V_i = \frac{\varepsilon_1 \log \log H}{2J} + O\left(\frac{1}{\log H_0}\right).$$

This implies property (b). If $(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J$, then for all $i \in [\![J]\!]$ we have

$$p_1 p_2 \cdots p_i < \exp\left(\sum_{1 \leqslant j \leqslant i} C^{2j-1} \log H_0\right) \leqslant \exp\left(2C^{2i-1} \log H_0\right) \leqslant \exp\left(\frac{C^{2i}}{10} \log H_0\right).$$

The right-hand side is $\leqslant p_{i+1}^{1/10}$ if $i < J$, and equals $H^{1/10}$ if $i = J$. This proves (c). Finally, the last two bounds for V^J follow from (b) and the fact that the function $J \mapsto (A/J)^J$ is increasing on $[0, A/e]$, for any $A > 0$. $\qquad\square$

2.2. **Proof of the two-point logarithmic Chowla bound.** In this section, we show how a bound on the double sum

$$(9) \qquad S_1 := \sum_{\substack{(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J}} \sum_{\substack{n \in (N, 2N] \\ p_1 \cdots p_J \mid n}} \lambda(n) \lambda(n + p_1 \cdots p_J)$$

implies a bound on the two-point logarithmically averaged Chowla conjecture. This step is due to Tao [15], and crucially relies on the multiplicativity of λ. With the proof of Proposition 2.6, this is the only place where the multiplicativity of λ is used – the rest of the paper will only use that λ is 1-bounded.

Proof of Theorem 1.1, assuming Theorem 2.1. Let $\varepsilon_1 > 0$ be a sufficiently small constant. Let H be a real number, chosen sufficiently large in terms of ε_1 so that Lemma 2.4 applies. We define $H_0 := \exp\left((\log H)^{1-\varepsilon_1}\right)$ and $x := \exp\left((\log H)^6\right)$. Choose J to be an integer of the form $c \log \log H$ where $\varepsilon_1^2/2 \leqslant c \leqslant \varepsilon_1^2$. Let $(\mathcal{P}_i)$ and (V_i) be as in Theorem 2.1. Let $V := \max_i V_i$. In particular, $(\log x)^{4\varepsilon_1^2} \ll V^J \ll (\log x)^{\varepsilon_1}$ by Lemma 2.4.

By Theorem 2.1, we know that

$$(10) \qquad S_1 = \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{n\in(N,2N] \\ p_1\cdots p_J|n}} \lambda(n)\lambda(n+p_1\cdots p_J) \ll V^{3J/4}N$$

whenever $\log N \geqslant (\log H)^3 = \sqrt{\log x}$. Moreover, when $\log N \leqslant \sqrt{\log x}$, trivially bounding $|\lambda| \leqslant 1$ we have

$$(11) \qquad \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{n\in(N,2N] \\ p_1\cdots p_J|n}} \lambda(n)\lambda(n+p_1\cdots p_J) \ll V^J N.$$

By a suitable dyadic decomposition, (10) and (11) together give, for all $M \geqslant 1$, the bound

$$(12) \qquad \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{m\leqslant M \\ p_1\cdots p_J|m}} \lambda(m)\lambda(m+p_1\cdots p_J) \ll V^J \min\!\left(M, e^{\sqrt{\log x}}\right) + V^{3J/4}M.$$

By partial summation, (12) and the bound $V^J \ll (\log x)^{\varepsilon_1}$ imply that

$$(13) \qquad \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{m\leqslant x \\ p_1\cdots p_J|m}} \frac{1}{m}\lambda(m)\lambda(m+p_1\cdots p_J) \ll V^{3J/4}\log x.$$

Let us now relate this estimate (13) to the expression $\sum_{n\leqslant x}\frac{1}{n}\lambda(n)\lambda(n+1)$ we are interested in. For any $(p_1,\ldots,p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J$, since λ is completely multiplicative and $\lambda^2 = 1$, we may rewrite

$$\sum_{n\leqslant x}\frac{1}{n}\lambda(n)\lambda(n+1) = \sum_{n\leqslant x}\frac{1}{n}\lambda(p_1\cdots p_J n)\lambda(p_1\cdots p_J n + p_1\cdots p_J)$$

$$= p_1\cdots p_J \sum_{\substack{m\leqslant p_1\cdots p_J x \\ p_1\cdots p_J|m}} \frac{1}{m}\lambda(m)\lambda(m+p_1\cdots p_J).$$

Dividing by $p_1\cdots p_J$ and summing over $(p_1,\ldots,p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J$ yields

$$(14) \qquad V_1\cdots V_J \sum_{n\leqslant x}\frac{1}{n}\lambda(n)\lambda(n+1) = \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{m\leqslant p_1\cdots p_J x \\ p_1\cdots p_J|m}} \frac{1}{m}\lambda(m)\lambda(m+p_1\cdots p_J).$$

This is almost the expression in (13), up to an error

$$\left| \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{x<m\leqslant p_1\cdots p_J x \\ p_1\cdots p_J|m}} \frac{1}{m}\lambda(m)\lambda(m+p_1\cdots p_J) \right| \leqslant \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \sum_{\substack{x<m\leqslant p_1\cdots p_J x \\ p_1\cdots p_J|m}} \frac{1}{m}$$

$$\ll \sum_{(p_1,\ldots,p_J)\in\mathcal{P}_1\times\cdots\times\mathcal{P}_J} \frac{\log(p_1\cdots p_J)}{p_1\cdots p_J}$$

which is $\ll V^J \log H$. Hence, by (13) and (14) we conclude that

$$\sum_{n\leqslant x}\frac{1}{n}\lambda(n)\lambda(n+1) \ll \frac{1}{V^J}\left(V^{3J/4}\log x + V^J\log H\right) \ll (\log x)^{1-\varepsilon_1^2}. \qquad \square$$

Remark 2.5 (Two-point Chowla at almost all scales)**.** Our main result also implies an improved quantitative version of Chowla's conjecture for two-point correlations at *almost all scales*. Namely,

for all $e < w \leqslant X$, we have

$$(15) \qquad \frac{1}{\log w} \int_{X/w}^{X} \left| \frac{1}{x} \sum_{n \leqslant x} \lambda(n)\lambda(n+1) \right| \frac{dx}{x} \ll \frac{1}{(\log w)^c},$$

where $c > 0$ is an absolute constant. In particular, we get

$$\frac{1}{x} \sum_{n \leqslant x} \lambda(n)\lambda(n+1) \ll \frac{1}{(\log X)^{c/2}}$$

for all $x \in [1, X]$ outside of a set E_X of logarithmic density $O((\log X)^{-c/2})$.[1]

This almost all scales result (15) follows from (7) by a straightforward adaptation of the proof in [5, Section 8] (which is itself inspired from [17]) to our setting.

2.3. Balanced sum. We define the 'balanced' double sum

$$(16) \qquad S_2 := \sum_{(p_1,\dots,p_J) \in \mathcal{P}_1 \times \dots \times \mathcal{P}_J} \sum_{n \in (N, 2N]} \left(\mathbf{1}_{p_1|n} - \frac{1}{p_1} \right) \cdots \left(\mathbf{1}_{p_J|n} - \frac{1}{p_J} \right) \lambda(n)\lambda(n + p_1 \cdots p_J).$$

Of course, S_1 is the same expression, but with $\mathbf{1}_{p_1|n} \cdots \mathbf{1}_{p_J|n}$ in place of $\left(\mathbf{1}_{p_1|n} - \frac{1}{p_1} \right) \cdots \left(\mathbf{1}_{p_J|n} - \frac{1}{p_J} \right)$. Working with S_1 or S_2 is essentially equivalent, as the following proposition shows.

Proposition 2.6. *Let ε_1, H, J, H_0, $(\mathcal{P}_i)$ and (V_i) be as in Theorem 2.1. Let $V := \max_i V_i$. Let $N \geqslant \exp\left((\log H)^2\right)$. With S_1 and S_2 as in (9) and (16), we have*

$$|S_1 - S_2| \ll \frac{N}{(\log H)^{1/2500}}.$$

Proposition 2.6 is proved in Appendix B, using the circle method and an estimate of Matomäki-Radziwiłł-Tao [9]. We will now focus on bounding S_2.

Notation 2.7. We define $\mathcal{P} := \bigsqcup_{j=1}^{J} \mathcal{P}_j$. To shorten the expressions, we define $\mathcal{D}$ to be the set of all products $p_1 p_2 \cdots p_J$ with $p_i \in \mathcal{P}_i$ for all $i \in [\![J]\!]$. We also write $I_N := \mathbb{N} \cap (N, 2N]$.

Thus, S_1 and S_2 may be rewritten more concisely as

$$S_1 := \sum_{n \in I_N} \sum_{\substack{d \in \mathcal{D} \\ d|n}} \lambda(n)\lambda(n+d) \quad \text{and} \quad S_2 = \sum_{n \in I_N} \sum_{d \in \mathcal{D}} \lambda(n)\lambda(n+d) \prod_{p|d} \left(\mathbf{1}_{p|n} - \frac{1}{p} \right).$$

3. A linear-algebraic approach

The purpose of this section is to simplify the analysis of the balanced expression S_2 by studying a certain weighted graph and its weighted adjacency matrix, which will effectively suppress the role of the Liouville function in the problem.

---

[1]This means that $\dfrac{1}{\log X} \displaystyle\int_{1}^{X} \mathbf{1}_{E_X}(x) \dfrac{dx}{x} \ll (\log X)^{-c/2}$.

3.1. **The original weighted graph G_0.** In this section, the vectors of size N and the $N \times N$ matrices will be indexed by the elements of $I_N = [\![N+1, 2N]\!]$ (instead of $[\![N]\!]$, as is standard).

Lemma 3.1. *Define the weighted graph $G_0 = (I_N, w_0)$, where the edge between $n \in I_N$ and $m \in I_N$ has weight*

$$w_0(m, n) = \begin{cases} \prod_{p|d}\left(\mathbf{1}_{p|n} - \frac{1}{p}\right) & \text{if } |m - n| = d \in \mathcal{D}, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathrm{Ad}_{G_0} = (w_0(m, n))_{m,n \in I_N}$ be the weighted adjacency matrix of G_0. Let $\boldsymbol{\lambda}$ be the column vector $(\lambda(n))_{n \in I_N}$. We have

$$S_2 = \tfrac{1}{2}\langle \boldsymbol{\lambda}, \mathrm{Ad}_{G_0}\boldsymbol{\lambda}\rangle + O(H^2).$$

Proof. By definition of Ad_{G_0}, we have

$$\langle \boldsymbol{\lambda}, \mathrm{Ad}_{G_0}\boldsymbol{\lambda}\rangle = \sum_{n \in I_N} \sum_{\substack{d \in \pm\mathcal{D} \\ n+d \in I_N}} \prod_{p|d}\left(\mathbf{1}_{p|n} - \frac{1}{p}\right)\lambda(n)\lambda(n+d) = 2S_2 + O\left(\sum_{n \in I_N}\sum_{d \in \pm\mathcal{D}} \mathbf{1}_{n+d \notin I_N}\right).$$

Recalling that $\pm\mathcal{D} \subset [-H, H]$, the error term is $\ll \sum_{n \in I_N} |\mathcal{D}|\, \mathbf{1}_{\min(n-N, 2N-n) \leqslant H} \ll H^2$. $\square$

Expressing S_2 in terms of the inner product $\langle \boldsymbol{\lambda}, \mathrm{Ad}_{G_0}\boldsymbol{\lambda}\rangle$ enables us to focus on the matrix Ad_{G_0} and remove the function λ from consideration. If we could show that every eigenvalue of Ad_{G_0} is $\ll V^{3J/4}$, we would be able to conclude that $S_2 \ll V^{3J/4}N$, as desired. Unfortunately, Ad_{G_0} itself does not satisfy such an eigenvalue bound. The strategy will thus be to cleverly modify G_0 in order to obtain a weighted graph whose weighted adjacency matrix has all its eigenvalues $\ll V^{3J/4}$.

3.2. **The high trace method for localised matrices.** The high trace method is a standard technique designed to control the eigenvalues of a Hermitian matrix A. Given an inequality of the form $\mathrm{Tr}(A^R) \leqslant C$ where R is an even integer, we can deduce that every eigenvalue α of A satisfies $|\alpha| \leqslant C^{1/R}$. This bound is weak when the dimension of the matrix is much larger than R. Fortunately, a stronger variant can be obtained for matrices whose non-zero entries all lie near the diagonal.

Let us recall Cauchy's interlacing theorem.

Notation 3.2 (Submatrix). Given $A = (a_{m,n})_{m,n \in I_N}$ and a subset $X \subset I_N$, we write $A|_X$ for the principal submatrix $(a_{m,n})_{m,n \in X}$ obtained by deleting all rows and columns at indices not in X.

Lemma 3.3 (Cauchy's interlacing theorem). *Let $A = (a_{m,n})_{m,n \in I_N}$ be a Hermitian matrix with eigenvalues $\alpha_1 \geqslant \ldots \geqslant \alpha_N$. Let $X \subset I_N$ and let $\beta_1 \geqslant \ldots \geqslant \beta_{N-|X|}$ be the eigenvalues of $A|_X$. Then, for $j \in [\![N - |X|]\!]$, we have*

$$\alpha_j \geqslant \beta_j \geqslant \alpha_{j+|X|}.$$

Proof. This is [1, Corollary III.1.5]. $\square$

Lemma 3.4. *Let $A = (a_{m,n})_{m,n \in I_N}$ be a Hermitian matrix such that $a_{m,n} = 0$ whenever $|m - n| > H$. Let $\alpha > 0$, $\varepsilon \in (0, 1)$, and suppose that $N \geqslant 10H/\varepsilon^2$. Then at least one of the following holds.*

(1) *There is a subset $E \subset I_N$ with $|E| \ll \varepsilon N$ such that every eigenvalue of $A|_{I_N \setminus E}$ has absolute value $\leqslant \alpha$.*

(2) *For any even integer $R \geqslant 2$,*

$$\mathrm{Tr}(A^R) \geqslant \frac{\varepsilon^2}{H}\alpha^R N.$$

Proof. Divide I_N into a sequence of disjoint discrete intervals $B_0, E_0, B_1, E_1, \ldots, B_{q-1}, E_{q-1}, B_q$ (lying in this order in I_N), such that the E_i have size H and the B_i have size $\lfloor 1/\varepsilon \rfloor H$, except for the last interval B_q which contains the $< H(1 + \lfloor 1/\varepsilon \rfloor)$ remaining elements. Since $N \geqslant 10H/\varepsilon^2$, we have $q \asymp \varepsilon N/H \gg 1/\varepsilon$.

Let $B = \bigsqcup_{i=0}^q B_i$. Since $B_0, \ldots, B_q$ have pairwise distance $> H$, the property of A in the statement implies that the submatrix $A|_B$ is block-diagonal with blocks $A|_{B_i}$ for $0 \leqslant i \leqslant q$.

Let $\mathcal{I}$ be the set of all indices $0 \leqslant i \leqslant q$ such that $A|_{B_i}$ has an eigenvalue $> \alpha$ in absolute value.

If $|\mathcal{I}| \leqslant \varepsilon^2 N/H$, we are in the first case of the conclusion. Indeed, we can take $E = \bigcup_{i=0}^{q-1} E_i \cup \bigcup_{i \in \mathcal{I}} B_i$ – clearly, all eigenvalues of $A|_{I_N \setminus E} = A|_{\bigsqcup_{i \notin \mathcal{I}} B_i}$ have absolute value $\leqslant \alpha$, and

$$|E| \leqslant qH + (\varepsilon^2 N/H)(2H/\varepsilon) \ll \varepsilon N.$$

Otherwise, $A|_B$ has at least $\varepsilon^2 N/H$ eigenvalues with absolute value $> \alpha$, counted with multiplicity. By Cauchy's interlacing theorem, the same is true for A. Thus, if $(\lambda_i)_{1 \leqslant i \leqslant N}$ are the eigenvalues of A, we have, for every even integer R,

$$\mathrm{Tr}\big(A^R\big) = \sum_i \lambda_i^R \geqslant \frac{\varepsilon^2 N}{H} \alpha^R$$

as all λ_i are real. $\qquad\square$

3.3. **Proof of main theorem assuming a high trace bound.**
We cannot use Lemma 3.4 with $A = \mathrm{Ad}_{G_0}$ directly, as the trace $\mathrm{Tr}\big((\mathrm{Ad}_{G_0})^R\big)$ turns out to be too large to yield any useful result. Instead, we will construct a close approximation G of the weighted graph G_0, whose weighted adjacency matrix Ad_G does satisfy a suitable high trace bound.

Proposition 3.5. *There exists a weighted graph $G = (I_N, w)$ with $\|w\|_\infty \leqslant 1$ such that*

(1) *(close to G_0)* $\|w - w_0\|_1 \ll N$;

(2) *(localised near the diagonal)* $w(m, n) = 0$ *whenever* $|m - n| > H$;

(3) *(small trace)* $\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant \big(e^{O(J)} V^{2J/3}\big)^K N$, *where* $K = 2\lfloor \log H \rfloor$.

Here $\|f\|_1 := \sum_{m,n} |f(m, n)|$ *for* $f : I_N \times I_N \to \mathbb{C}$.

With Proposition 3.5 at our disposal, it is straightforward to deduce Theorem 2.1.

Proof of Theorem 2.1, assuming Proposition 2.6 and Proposition 3.5. By Proposition 2.6, it suffices to prove that $S_2 \ll NV^{3J/4}$. By Lemma 3.1, it suffices to prove the same bound for $\langle \boldsymbol{\lambda}, \mathrm{Ad}_{G_0}\boldsymbol{\lambda} \rangle$.

Let G be the graph given by Proposition 3.5. Since $\|w_0 - w\|_1 \ll N$, we have

$$\langle \boldsymbol{\lambda}, \mathrm{Ad}_{G_0}\boldsymbol{\lambda} \rangle = \langle \boldsymbol{\lambda}, \mathrm{Ad}_G\boldsymbol{\lambda} \rangle + O(N). \tag{17}$$

We now apply Lemma 3.4 with $A = \mathrm{Ad}_G$, $\varepsilon = 1/H$ and $\alpha = V^{3J/4}$. The second case of Lemma 3.4 cannot hold, since otherwise we would have

$$\frac{1}{H^3}\big(V^{3J/4}\big)^K N \leqslant \mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant \big(e^{O(J)} V^{2J/3}\big)^K N.$$

This implies $V \ll 1$, but $V \gg \varepsilon_1^{-1}$ by part (b) of Lemma 2.4, so we obtain a contradiction provided that ε_1 is sufficiently small.

Thus, the first case holds and there is a subset $E \subset I_N$ of size $|E| \ll N/H$ such that every eigenvalue of $(\mathrm{Ad}_G)|_{I_N \setminus E}$ has absolute value $\leqslant V^{3J/4}$. The bound on the size of E implies that

$\left\| w - w|_{I_N \setminus E} \right\|_1 \ll N$. Hence, writing $\boldsymbol{\lambda}|_{I_N \setminus E}$ for the vector $(\lambda(n))_{n \in I_N \setminus E}$, we have

$$(18) \qquad \langle \boldsymbol{\lambda}, \mathrm{Ad}_G \boldsymbol{\lambda} \rangle = \langle \boldsymbol{\lambda}|_{I_N \setminus E}, (\mathrm{Ad}_G)|_{I_N \setminus E} \boldsymbol{\lambda}|_{I_N \setminus E} \rangle + O(N).$$

Since $(\mathrm{Ad}_G)|_{I_N \setminus E}$ is a Hermitian matrix with all eigenvalues $\ll V^{3J/4}$, we conclude that

$$\langle \boldsymbol{\lambda}|_{I_N \setminus E}, (\mathrm{Ad}_G)|_{I_N \setminus E} \boldsymbol{\lambda}|_{I_N \setminus E} \rangle \leqslant \left\| \boldsymbol{\lambda}|_{I_N \setminus E} \right\|_2 \left\| (\mathrm{Ad}_G)|_{I_N \setminus E} \boldsymbol{\lambda}|_{I_N \setminus E} \right\|_2 \ll N V^{3J/4}.$$

By (17) and (18), Theorem 2.1 follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The remainder of this paper devoted to the proof of Proposition 3.5.

3.4. **High trace as a sum over closed walks.** The first step to get a high trace bound is to use the following well-known fact. Let G be a finite weighted graph. For any $R \geqslant 1$, the trace of the R-fold composition $(\mathrm{Ad}_G)^R$ is given by the sum of the weights of all closed walks of length R in G, where the weight of a walk is the product of the weights of its edges.

Definition 3.6. Let $\mathbf{D}_R$ be the set of all $\boldsymbol{d} \in (\pm \mathcal{D})^R$ such that $\sum_{i=1}^R d_i = 0$. For $\boldsymbol{d} \in \mathbf{D}_R$, we define the partial sums

$$(19) \qquad b_i(\boldsymbol{d}) := \sum_{1 \leqslant i' < i} d_{i'}.$$

When $\boldsymbol{d}$ is clear from the context, we will write b_i instead of $b_i(\boldsymbol{d})$.

In our graph $G_0 = (I_N, w_0)$, the closed walks (with non-zero weight) of length K are of the form

$$\underbrace{n + b_1}_{=n} \overset{+d_1}{\curvearrowright} n + b_2 \overset{+d_2}{\curvearrowright} n + b_3 \overset{+d_3}{\curvearrowright} \cdots \curvearrowright n + b_K \overset{+d_K}{\curvearrowright} \underbrace{n + b_{K+1}}_{=n}$$

for some $\boldsymbol{d} \in \mathbf{D}_K$. The above fact about the trace of powers of adjacency matrices implies that

$$(20) \qquad \mathrm{Tr}\big((\mathrm{Ad}_{G_0})^K\big) = \sum_{\boldsymbol{d} \in \mathbf{D}_K} \sum_{n \in I_N} \prod_{i \in [\![K]\!]} w_0(n + b_i, n + b_{i+1}),$$

with the convention that $w_0(m, n) := 0$ if m or n is not in I_N.

4. Heuristics for the definition of G

This section serves purely as motivation and is separate from the actual proof. The aim is to explain why G_0 needs to be replaced with a smoothed out graph G.

4.1. **Cancellation from the balanced weights.** By definition of w_0, (20) can be rewritten as

$$(21) \qquad \mathrm{Tr}\big((\mathrm{Ad}_{G_0})^K\big) = \sum_{\boldsymbol{d} \in \mathbf{D}_K} \sum_{\substack{n \in I_N \\ \forall i, \, n + b_i \in I_N}} \prod_{i \in [\![K]\!]} \prod_{p | d_i} \left(\mathbf{1}_{p | n + b_i} - \frac{1}{p} \right).$$

We may divide the long sum over $n \in I_N$ into arithmetic progressions of modulus $d_1 \cdots d_K$ (note that N is much larger than the product $d_1 \cdots d_K$). Ignoring the error terms for this sketch, we obtain

$$(22) \qquad \mathrm{Tr}\big((\mathrm{Ad}_{G_0})^K\big) \approx \sum_{\boldsymbol{d} \in \mathbf{D}_K} \frac{N}{d_1 \cdots d_K} \sum_{n \,(\mathrm{mod}\, d_1 \cdots d_K)} \prod_{i \in [\![K]\!]} \prod_{p | d_i} \left(\mathbf{1}_{p | n + b_i} - \frac{1}{p} \right).$$

For $\boldsymbol{d} \in \mathbf{D}_K$, define

$$F(\boldsymbol{d}) := \frac{1}{d_1 \cdots d_K} \sum_{n \,(\mathrm{mod}\, d_1 \cdots d_K)} \prod_{i \in [\![K]\!]} \prod_{p \mid d_i} \left(\mathbf{1}_{p \mid n + b_i} - \frac{1}{p} \right).$$

By the Chinese remainder theorem, $F(\boldsymbol{d})$ admits a factorisation into terms corresponding to the primes dividing $d_1 \cdots d_K$. More precisely, we have

$$(23) \qquad\qquad F(\boldsymbol{d}) = \prod_{p \mid d_1 \cdots d_K} F_p(\boldsymbol{d}),$$

where

$$(24) \qquad\qquad F_p(\boldsymbol{d}) := \frac{1}{p} \sum_{n \,(\mathrm{mod}\, p)} \prod_{\substack{i \in [\![K]\!] \\ p \mid d_i}} \left(\mathbf{1}_{p \mid n + b_i} - \frac{1}{p} \right).$$

Let $\boldsymbol{d} \in \mathbf{D}_K$ and suppose that there is a prime $p \in \mathcal{P}$ dividing exactly one of $d_1, \ldots, d_K$, say $p \mid d_{i_0}$. Then, we have perfect cancellation

$$F_p(\boldsymbol{d}) = \frac{1}{p} \sum_{n \,(\mathrm{mod}\, p)} \left(\mathbf{1}_{p \mid n + b_{i_0}} - \frac{1}{p} \right) = 0,$$

and hence $F(\boldsymbol{d}) = 0$. This means that those $\boldsymbol{d}$ having a prime $p \mid d_1 \cdots d_K$ with $p^2 \nmid d_1 \cdots d_K$ do not contribute to the expression (22). This is an important observation as the vast majority of $\boldsymbol{d} \in \mathbf{D}_K$ have this property.

Therefore, it only remains to consider the $\boldsymbol{d} \in \mathbf{D}_K$ such that, for every $p \in \mathcal{P}$, having $p \mid d_1 \cdots d_K$ implies that $p^2 \mid d_1 \cdots d_K$.

4.2. **Repeated prime divisors.** Let $\boldsymbol{d} \in \mathbf{D}_K$ and $p \in \mathcal{P}$. Suppose that there are exactly two indices $i \in [\![K]\!]$ such that $p \mid d_i$, say i_1 and i_2. Then

$$F_p(\boldsymbol{d}) = \frac{1}{p} \sum_{n \,(\mathrm{mod}\, p)} \left(\mathbf{1}_{p \mid n + b_{i_1}} - \frac{1}{p} \right) \left(\mathbf{1}_{p \mid n + b_{i_2}} - \frac{1}{p} \right) = \begin{cases} 1/p - 1/p^2 & \text{if } b_{i_1} \equiv b_{i_2} \pmod{p}, \\ -1/p^2 & \text{otherwise.} \end{cases}$$

Observe that $|F_p(\boldsymbol{d})|$ is as large as what would be obtained by replacing the weights w_0 by their absolute values, so there is no cancellation from the balanced weights. Moreover, the size of $F_p(\boldsymbol{d})$ depends on whether $b_{i_2} - b_{i_1}$ is divisible by p or not.

   (1) If $p \mid b_{i_2} - b_{i_1}$, we have $|F_p(\boldsymbol{d})| \asymp 1/p$.
   (2) If $p \nmid b_{i_2} - b_{i_1}$, we have $|F_p(\boldsymbol{d})| \asymp 1/p^2$.

Recall that all primes $p \in \mathcal{P}$ are $\geqslant H_0$, where H_0 is a rather large parameter. Hence, in the second case, we have $F_p(\boldsymbol{d}) \ll 1/(H_0 p)$ and we save a factor H_0 compared with the first case.

The main takeaway is the following. Let $\boldsymbol{d} \in \mathbf{D}_K$ and suppose that there are many primes $p \in \mathcal{P}$ such that case (2) holds. Then $F_p(\boldsymbol{d}) \ll 1/(H_0 p)$ for all these p, which implies that $F(\boldsymbol{d})$ is small and has a negligible contribution to the trace (21).

A similar reasoning applies where there are more than two indices $i \in [\![K]\!]$ such that $p \mid d_i$, and the size of $F_p(\boldsymbol{d})$ depends on whether the corresponding shifts b_i are all congruent modulo p or not.

We still have to examine the walks $\boldsymbol{d} \in \mathbf{D}_K$ where all the primes p dividing $d_1 \cdots d_K$ are repeated and most of them satisfy case (1).

4.3. **Problematic walks.** We already mentioned that the graph G_0 does not satisfy the third property of Proposition 3.5, i.e. a suitable high trace bound. Let us explain why this is the case.

It is possible to exhibit a family of $\boldsymbol{d} \in \mathbf{D}_K$ for which $F(\boldsymbol{d})$ is rather large. Let $e_1, \ldots, e_{K/2}$ be arbitrary elements of $\mathcal{D}$ and consider the vector

(25) $$\boldsymbol{d} := (e_1, -e_1, e_2, -e_2, \ldots, e_{K/2}, -e_{K/2}) \in \mathbf{D}_K.$$

Note that all the primes dividing $d_1 \cdots d_K$ are repeated, as $p \mid d_{2i-1} = e_i$ if and only if $p \mid d_{2i} = -e_i$. Moreover, whenever a prime p divides two coordinates d_{i_1} and d_{i_2}, we have $p \mid b_{i_2} - b_{i_1}$. This immediately follows from the fact that $b_{2i-1} = 0$ and $b_{2i} = e_i$ for all i. Therefore, case (1) of Section 4.2 applies, which means that $|F_p(\boldsymbol{d})| \approx 1/p$ and thus

$$F(\boldsymbol{d}) \approx \prod_{p \mid d_1 \cdots d_K} \frac{1}{p}.$$

To obtain the total contribution of those $\boldsymbol{d}$ of the form given by (25), one would need to sum $F(\boldsymbol{d})$ over all possible choices of $e_1, \ldots, e_{K/2}$. This is a fairly straightforward computation – very similar to Lemma 6.4, so we shall not repeat it here. In the end, one finds that the contribution of these $\boldsymbol{d}$ to the trace (21) is much greater than what is allowed by Proposition 3.5.

It is instructive to interpret this issue in terms of 'back-and-forth' walks on I_N. Let $n \in I_N$ and consider the family of walks

$$n \overset{+e_1}{\curvearrowright} n + e_1 \overset{-e_1}{\curvearrowright} n \overset{+e_2}{\curvearrowright} n + e_2 \curvearrowright \ldots \curvearrowright n \overset{+e_{K/2}}{\curvearrowright} n + e_{K/2} \overset{-e_{K/2}}{\curvearrowright} n$$

where the e_i range over the set $\{d \in \mathcal{D} : d \mid n\}$. Since we restrict the e_i to be divisors of n, we have $d_i \mid n + b_i$ for all i (indeed, this just means that $e_i \mid n$ and $-e_i \mid n + e_i$). Hence, the weight of this walk is

$$\prod_{i \in [\![K]\!]} \prod_{p \mid d_i} \left(\mathbf{1}_{p \mid n + b_i} - \frac{1}{p} \right) = \prod_{i \in [\![K]\!]} \prod_{p \mid d_i} \left(1 - \frac{1}{p} \right) \approx 1$$

(since this is only a sketch, we ignore the fact that these walks can escape I_N if n is very close to the boundary of that interval). Let $\tau_{\mathcal{D}}(n)$ be the number of divisors of n in the set $\mathcal{D}$. Since there are $\tau_{\mathcal{D}}(n)$ choices for every e_i, the contribution of these back-and-forth walks to the trace (21) is

$$\approx \sum_{n \in I_N} \tau_{\mathcal{D}}(n)^{K/2}.$$

On average, the number of divisors $d \in \mathcal{D}$ of an element of I_N is $\approx V^J$. If all $n \in I_N$ satisfied $\tau_{\mathcal{D}}(n) \ll V^J$, the contribution to (21) of these back-and-forth walks would roughly be

$$\sum_{n \in I_N} \tau_{\mathcal{D}}(n)^{K/2} \ll e^{O(K)} V^{KJ/2} N.$$

This contribution would be acceptable as it is smaller than the bound in Proposition 3.5. Unfortunately, it is not true that all $n \in I_N$ have $\tau_{\mathcal{D}}(n) \ll V^J$. In fact, since K is quite large, the high moment $\sum_{n \in I_N} \tau_{\mathcal{D}}(n)^{K/2}$ is dominated by the contribution of those $n \in I_N$ with a lot more than V^J divisors from $\mathcal{D}$. Because of this, the contribution of these back-and-forth walks vastly exceeds the required trace upper bound.

To resolve this issue, we will remove from the vertex set of G_0 all integers $n \in I_N$ having an unusual number of prime factors in $\mathcal{P}$. This modification will reduce the contribution of the above back-and-forth walks (and more generally, the contribution of *backtracking* walks) within acceptable bounds.

4.4. **General strategy.** In Section 5, we will replace G_0 with a better-behaved weighted graph by suppressing certain undesired integers n from the vertex set I_N. As we have seen in Section 4.3, restricting to a suitable subset of I_N is necessary to be able to prove an acceptable high trace bound.

Removing those $n \in I_N$ with too many prime factors in $\mathcal{P}$ allows us to control walks that retrace their steps. We treat these walks in Section 10. However, this first change causes additional technical difficulties. Unlike for the naive graph G_0, the weight of a walk does not perfectly cancel when there is a prime p dividing exactly one of $d_1, \ldots, d_K$. Rather, we will be able to obtain a little saving from each such prime. These savings accumulate, and we will obtain an acceptable bound if there are many such unrepeated primes. This is the content of Section 7.

It remains to deal with the walks $\boldsymbol{d}$ having many repeated primes, i.e. primes p dividing several of $d_1, \ldots, d_K$. As we have seen in Section 4.2, their contribution is small unless certain divisibility relations hold. These divisibilities are of the form $p \mid b_{i_2} - b_{i_1}$, where p is a common prime factor of d_{i_1} and d_{i_2}. The hope would be to show there can only be very few $\boldsymbol{d}$ which satisfy many such divisibility relations. Doing so turns out to be a complicated combinatorial problem.

To simplify this task, we further restrict the vertex set of our weighted graph: we remove certain $n \in I_N$ satisfying some unexpected divisibility conditions. Just like the integers with too many prime factors from $\mathcal{P}$, these special n form a sparse subset of I_N, but could potentially boost the contribution of certain bad walks. With this second modification of the weighted graph, we are able to deal with walks having many repeated primes in Sections 8 and 9.

Putting everything together, we will obtain the desired bound for the trace of a high power of the weighted adjacency matrix of the modified graph.

5. The smoothed weighted graph G

In this section, we define the weighted graph $G = (I_N, w)$ and prove that it satisfies the first property of Proposition 3.5. To construct it, we will make two modifications to G_0. Although these changes affect few entries of Ad_{G_0}, they become significant when we raise this matrix to a large power K.

5.1. **Discarding integers with too many prime factors.** An integer $n \in I_N$ typically has about JV prime factors in $\mathcal{P}$. However, a few exceptional integers have a lot more prime factors in $\mathcal{P}$. As we hinted in Section 4.3, this is the main reason why $\mathrm{Tr}\big((\mathrm{Ad}_{G_0})^K\big)$ is exceedingly large.

For $n \in \mathbb{Z}$, recall that $\omega_{\mathcal{P}}(n)$ denotes the number of distinct prime factors of n in $\mathcal{P}$. We will restrict the vertex set of our weighted graph to only contain integers n having $\omega_{\mathcal{P}}(n) \approx JV$. For technical reasons, we do so by introducing a smooth cut-off (this will be useful in the proof of Proposition 7.3.). We need a smooth approximation to the indicator function of the interval $\big[\frac{1}{2}JV, \frac{3}{2}JV\big]$. The properties that we need are summarised in Lemma C.1, which we reproduce here for convenience.

Lemma C.1. *There exists a C^{∞} function $W : \mathbb{R} \to [0,1]$ such that*

- $W(x) = 1$ *for $x \in \big[\frac{1}{2}JV, \frac{3}{2}JV\big]$;*
- $W(x) = 0$ *for $x \notin \big[0, 2JV\big]$;*
- *(Bound a-th derivative of m-th power) For any integers $a \geqslant 1$ and $m \geqslant 1$,*

$$\left\| (W^m)^{(a)} \right\|_{\infty} \leqslant 2^m \left(\frac{Ca}{JV} \right)^a,$$

where C is an absolute constant.

Definition 5.1. We define the weighted graph $G_1 = (I_N, w_1)$, where the edge between $n \in I_N$ and $m \in I_N$ has weight

$$w_1(m, n) = \begin{cases} \left[\prod_{p|d} \left(\mathbf{1}_{p|n} - \frac{1}{p} \right) \right] W\left(\omega_{\mathcal{P}}(n) \right)^{1/2} W\left(\omega_{\mathcal{P}}(m) \right)^{1/2} & \text{if } |m - n| = d \in \mathcal{D}, \\ 0 & \text{otherwise.} \end{cases}$$

5.2. Excluding some special divisibility patterns. We mentioned in Section 4.4 that certain integers $n \in I_N$ satisfying some unexpected congruence conditions would also need to be removed from the vertex set. This modification is required for our methods to be able to handle the walks with many repeated primes: it will be crucial for Section 9.

While this is a necessary step for our methods, it does lead to technical obstacles in Section 7; these are overcome in Lemma 7.6 (which is proved in Section 11).

We now give the definition of these exceptional integers. The details are not too important for now as we only really need this definition for Lemmas 9.20 and 9.21, as well as Sections 11.3 and 11.4.

Definition 5.2. Let $L := K^{1-10\varepsilon_1}$.

A *prohibited sequence* is a sequence $(d_1, \ldots, d_\ell)$ of ℓ elements of $\pm\mathcal{D}$, for some $2 < \ell \leqslant L$, with the following properties:

- (non-backtracking) $d_{i+1} \neq -d_i$ for all $1 \leqslant i < \ell$, and;
- (consecutiveness) for every prime q, the set $\{i \in [\![\ell]\!] : q \mid d_i\}$ is a discrete interval, and;
- (prohibited pattern) there is a prime p and some $1 < \ell_0 < \ell$ such that $p \mid d_1$, $p \nmid d_\ell$ and

$$(26) \qquad\qquad p \; \Big| \sum_{\ell_0 \leqslant i \leqslant \ell} d_i.$$

A prohibited sequence $(d_1, \ldots, d_\ell)$ is *primitive* if there is no consecutive[2] subsequence of $(d_1, \ldots, d_\ell)$ or of $(d_\ell, \ldots, d_1)$, of length $< \ell$, which is also prohibited.

A key difference with [5] is that, in their situation, the authors can restrict themselves to the case $\ell_0 = 1$. This is not possible here, and leads to additional complications in the proof of Lemma 7.6 (due to the fact that the constraint (26) only involves a subset of the prime factors of the d_i). Having defined prohibited sequences, we may now turn to the exceptional integers that need to be removed from the vertex set.

Definition 5.3. The *prohibited (arithmetic) progression* associated with a primitive prohibited sequence $(d_1, \ldots, d_\ell)$ is the set of all integers $n \in \mathbb{Z}$ such that

$$d_1 \mid n, \quad d_2 \mid n + d_1, \quad \ldots \quad d_\ell \mid n + d_1 + \cdots + d_{\ell-1}.$$

It is an arithmetic progression of square-free modulus $\mathrm{lcm}(d_1, \ldots, d_\ell)$.

Let $\mathcal{Y}$ be the set of all prohibited progressions associated with some primitive prohibited sequence. We define $Y_L := \mathbb{Z} \setminus \cup\mathcal{Y}$, the set of all integers that do not belong to any prohibited progression.

We are ready to define the announced weighted graph $G = (I_N, w)$.

---

[2]By 'consecutive subsequence of $(d_1, \ldots, d_\ell)$', we mean a sequence of the form $(d_{k_1}, d_{k_1+1}, \ldots, d_{k_2})$ for some $1 \leqslant k_1 < k_2 \leqslant \ell$.

Definition 5.4. Let G be the weighted graph with vertex set I_N where the edge between $n \in I_N$ and $m \in I_N$ has weight

$$w(m,n) = \begin{cases} \left[\prod_{p|d} \left(\mathbf{1}_{p|n} - \frac{1}{p} \right) \right] W\left(\omega_{\mathcal{P}}(n) \right)^{1/2} W\left(\omega_{\mathcal{P}}(m) \right)^{1/2} \mathbf{1}_{Y_L}(n) \mathbf{1}_{Y_L}(m) & \text{if } |m-n| = d \in \mathcal{D}, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, G can be identified with the weighted graph $(I_N \cap Y_L, w_1|_{(I_N \cap Y_L) \times (I_N \cap Y_L)})$.

5.3. **Comparison of the two weighted graphs.** The weighted graph G just defined clearly satisfies the second property of Proposition 3.5. We now prove the first property, which states that G is a close approximation to G_0. Note that the weight functions of G_0 and G only differ for edges (m,n) where one of the endpoints m,n either has an atypical number of prime factors from $\mathcal{P}$, or does not lie in Y_L.

Lemma 5.5. $|I_N \setminus Y_L| \ll H_0^{-1/3} N.$

Lemma 5.5 is not hard to show, but we defer the proof of this fact to Section 11, where we will prove many other bounds of a similar type. Assuming Lemma 5.5, it is easy to prove the following lemma.

Lemma 5.6. *We have* $\|w_0 - w\|_1 \ll N$, *where* $\|f\|_1 := \sum_{m,n} |f(m,n)|.$

Proof of Lemma 5.6, assuming Lemma 5.5. Let

$$\xi(n) := \mathbf{1}_{|\omega_{\mathcal{P}}(n) - JV| \geqslant JV/2} + \mathbf{1}_{n \notin Y_L}.$$

Since $\|W\|_\infty \leqslant 1$, we have

$$\|w_0 - w\|_1 \leqslant \sum_{n \in I_N} \sum_{\substack{d \in \pm\mathcal{D} \\ n+d \in I_N}} \left[\prod_{p|d} \left(\mathbf{1}_{p|n} + \frac{1}{p} \right) \right] (\xi(n) + \xi(n+d)) \ll \sum_{n \in I_N} \xi(n) \prod_{i=1}^{J} \left(\omega_{\mathcal{P}_i}(n) + V_i \right).$$

Hence, by Cauchy-Schwarz,

$$(27) \qquad \|w_0 - w\|_1 \ll \left(\sum_{n \in I_N} \xi(n)^2 \right)^{1/2} \left(\sum_{n \in I_N} \prod_{i=1}^{J} \left(\omega_{\mathcal{P}_i}(n) + V_i \right)^2 \right)^{1/2}.$$

Let us bound the first sum on the right-hand side. By [11, Eq. (1.11)], we know that

$$(28) \qquad \sum_{n \in I_N} \mathbf{1}_{|\omega_{\mathcal{P}}(n) - JV| \geqslant JV/2} \ll e^{-JV/50} N.$$

By Lemma 5.5, we have

$$(29) \qquad \sum_{n \in I_N} \mathbf{1}_{n \notin Y_L} \ll H_0^{-1/3} N.$$

Together, (28) and (29) give

$$\sum_{n \in I_N} \xi(n)^2 \ll \left(H_0^{-1/3} + e^{-JV/12} \right) N.$$

For the second sum on the right-hand side of (27), we have, by the AM-GM inequality,

$$\sum_{n \in I_N} \prod_{i=1}^{J} \left(\omega_{\mathcal{P}_i}(n) + V_i \right)^2 \leqslant \sum_{n \in I_N} \left(\frac{\omega_{\mathcal{P}}(n)}{J} + V \right)^{2J} \leqslant 2^{2J} V^{2J} N + 2^{2J} \sum_{n \in I_N} \left(\frac{\omega_{\mathcal{P}}(n)}{J} \right)^{2J}.$$

Using that $(a/n)^n \leqslant e^a$ for $a \geqslant 0$, we obtain that

$$\sum_{n \in I_N} \left(\frac{\omega_{\mathcal{P}}(n)}{J} \right)^{2J} = e^{O(J)} \sum_{n \in I_N} \left(\frac{\omega_{\mathcal{P}}(n)/200}{2J} \right)^{2J} \leqslant e^{O(J)} \sum_{n \in I_N} e^{\omega_{\mathcal{P}}(n)/200} \leqslant e^{O(J)} N e^{JV/100},$$

the last inequality being a consequence of [11, Lemma (3.10)].

Putting everything together, we conclude that

$$\|w_0 - w\|_1 \ll e^{O(J)} \left(H_0^{-1/3} + e^{-JV/50} \right)^{1/2} \left(V^{2J} + e^{JV/100} \right)^{1/2} N.$$

By our choices of parameters (see Lemma 2.4), we have $H_0^{-1/3} \ll e^{-JV/50}$ (as $JV \ll (\log \log H_0)^2$) and $V^{2J} \ll e^{JV/100}$ (as $V \gg \varepsilon_1^{-1}$). Thus, we conclude that

$$\|w_0 - w\|_1 \ll e^{O(J)} e^{-JV/100} N,$$

which is $\ll N$ if ε_1 is sufficiently small. $\qquad \square$

Hence, G satisfies the first two hypotheses of Proposition 3.5. The remaining sections are devoted to the proof of the high trace bound $\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant \big(e^{O(J)} V^{2J/3}\big)^K N$.

6. THE THREE TYPES OF INDICES

Now that we have defined our weighted graph G, we start our analysis of the trace of $(\mathrm{Ad}_G)^K$. The main statement summarising the results of this section is Proposition 6.16.

6.1. **Rewriting the trace.** We have seen at the end of Section 3 that the trace of a power of the adjacency matrix of a weighted graph can be expanded in terms of closed walks on that graph. For $\boldsymbol{d} \in \mathbf{D}_R$ and $n \in I_N$, let

$$(30) \qquad w_{\boldsymbol{d}}(n) := \prod_{i \in [\![R]\!]} w(n + b_i, n + b_{i+1}),$$

where $b_i = b_i(\boldsymbol{d}) = \sum_{i' < i} d_{i'}$ as before. Similarly to (20), we have

$$(31) \qquad \mathrm{Tr}\big((\mathrm{Ad}_G)^R\big) = \sum_{\boldsymbol{d} \in \mathbf{D}_R} \sum_{\substack{n \in I_N \\ \forall i,\, n + b_i \in I_N}} w_{\boldsymbol{d}}(n).$$

Observe that the term

$$w_{\boldsymbol{d}}(n) = \prod_{i \in [\![R]\!]} W\big(\omega_{\mathcal{P}}(n + b_i)\big) \mathbf{1}_{n + b_i \in Y_L} \prod_{p \mid d_i} \left(\mathbf{1}_{p \mid n + b_i} - \frac{1}{p} \right)$$

only depends on the congruence class of n modulo every $p \in \mathcal{P}$ (or more precisely, on the set of prime factors in $\mathcal{P}$ of each $n + b_i$). For our study of the cancellations arising from these balanced weights (see Section 7), it will be convenient to adopt a probabilistic viewpoint.

Definition 6.1. Let $\mathbf{n}$ be a random variable taking values in $\prod_{p \in \mathcal{P}} \mathbb{Z}/p\mathbb{Z}$ with the uniform distribution. If $f : \mathbb{Z} \to \mathbb{C}$ is a function such that $f(n)$ only depends on the congruence class of n modulo each prime $p \in \mathcal{P}$, we still write $f(\mathbf{n})$ for the random variable defined in the obvious way.

The following lemma says that we may replace, in (31), the uniform probability measure on I_N with the uniform probability measure on $\prod_{p \in \mathcal{P}} \mathbb{Z}/p\mathbb{Z}$. This step corresponds to Equation (22) in the outline given in Section 4.

Lemma 6.2. *We have*

(32) $$\text{Tr}\big((\text{Ad}_G)^K\big) = N \sum_{\boldsymbol{d}\in\mathbf{D}_K} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n})\right] + O\left(N + Ne^{-\sqrt{\log N}} \sum_{\boldsymbol{d}\in\mathbf{D}_K} \mathbb{E}\left[|w_{\boldsymbol{d}}(\mathbf{n})|\right]\right).$$

Lemma 6.2 is proved in Appendix D, using the Fundamental Lemma of sieve theory.

Let us simplify the error term in Lemma 6.2.

Definition 6.3. Let $\boldsymbol{d} \in \mathbf{D}_R$. For $(i,j) \in [\![R]\!] \times [\![J]\!]$, we write d_{ij} for the unique prime in $\mathcal{P}_j$ that divides d_i. Thus $|d_i| = \prod_{j\in[\![J]\!]} d_{ij}$.

For any subset $I \subset [\![R]\!] \times [\![J]\!]$, we set

$$\rho_{\boldsymbol{d};I} := \prod_{(i,j)\in I} d_{ij}.$$

In the special case $I = [\![R]\!] \times [\![J]\!]$, we will write $\rho_{\boldsymbol{d}}$ instead of $\rho_{\boldsymbol{d};[\![R]\!]\times[\![J]\!]}$ to shorten notation.

Lemma 6.4. *We have*

$$\sum_{\boldsymbol{d}\in\mathbf{D}_K} \prod_{p|\rho_{\boldsymbol{d}}} \frac{2}{p} \ll K^{2KJ}.$$

Proof. Any $\boldsymbol{d} \in \mathbf{D}_K$ induces a partition of $[\![K]\!] \times [\![J]\!]$, where (i,j) and (i',j') are in the same class if and only if $d_{ij} = d_{i'j'}$. Every class α of the partition is contained in $[\![K]\!] \times \{j_\alpha\}$ for some $j_\alpha \in [\![J]\!]$, because the sets $\mathcal{P}_j$ are disjoint. Observe that $\boldsymbol{d}$ is fully determined by a sequence of K signs (the signs of the d_i), such a partition of $[\![K]\!] \times [\![J]\!]$ and the assignment of a prime in $\mathcal{P}_{j_\alpha}$ to every class α of this partition (the prime factors of the d_i).

Summing over all sequences of signs σ, suitable partitions Π of $[\![K]\!] \times [\![J]\!]$ and primes in $\mathcal{P}$, we have

$$\sum_{\boldsymbol{d}\in\mathbf{D}_K} \prod_{p|\rho_{\boldsymbol{d}}} \frac{1}{p} \leqslant \sum_{\sigma\in\{\pm1\}^K} \sum_{\Pi} \prod_{\alpha\in\Pi} \sum_{p_\alpha\in\mathcal{P}_{j_\alpha}} \frac{1}{p_\alpha} \leqslant 2^K (KJ)^{KJ} V^{KJ},$$

where we used that the number of partitions of $[\![K]\!] \times [\![J]\!]$ is $\leqslant (KJ)^{KJ}$.

By property (b) of Lemma 2.4 and the simple bound $(a/n)^n \leqslant e^a$ for $a \geqslant 0$, we have $V^J \leqslant K$. Therefore, the sum in the statement is $\ll 2^{KJ}(2K)^K(KJ)^{KJ} \ll K^{2KJ}$. $\qquad\square$

Lemma 6.5. *We have*

$$\text{Tr}\big((\text{Ad}_G)^K\big) = N \sum_{\boldsymbol{d}\in\mathbf{D}_K} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n})\right] + O(N).$$

Proof. By the triangle inequality,

$$\mathbb{E}\left[|w_{\boldsymbol{d}}(\mathbf{n})|\right] \leqslant \mathbb{E}\left[\prod_{i\in[\![K]\!]} \prod_{p|d_i} \left|\mathbf{1}_{p|\mathbf{n}+b_i} - \frac{1}{p}\right|\right] \leqslant \prod_{p|\rho_{\boldsymbol{d}}} \frac{2}{p},$$

which is $\ll K^{2KJ}$ by Lemma 6.4. Plugging this into Lemma 6.2, the corollary follows. $\qquad\square$

6.2. **Single, lit and unlit indices.** The weight $w_{\boldsymbol{d}}(\mathbf{n})$ contains a factor $\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}$ for every $(i,j) \in [\![K]\!] \times [\![J]\!]$ (in addition to some W and $\mathbf{1}_{Y_L}$ factors). As we have discussed in Section 4.1, some factors $\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}$ induce cancellation in the expected value. This happens exactly for those primes d_{ij} that are not repeated in the array $(d_{ij})_{(i,j)\in[\![K]\!]\times[\![J]\!]}$.

Definition 6.6 (Single indices)**.** Let $\boldsymbol{d} \in \mathbf{D}_R$. We say that an index $(i, j) \in [\![R]\!] \times [\![J]\!]$ is *single* if $d_{ij}^2 \nmid \rho_{\boldsymbol{d}}$, i.e. the prime d_{ij} does not appear at any index other than (i, j).

Given $\mathcal{S} \subset [\![R]\!] \times [\![J]\!]$, we let $\mathbf{D}_R^{\mathcal{S}}$ be the set of all $\boldsymbol{d} \in \mathbf{D}_R$ whose set of single indices is $\mathcal{S}$.

We now put the single indices aside, and divide the remaining indices into two classes, in order to replace the random factor $\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}$ by a deterministic factor $1 - \frac{1}{d_{ij}}$ or $-\frac{1}{d_{ij}}$.

Lemma 6.7. *We have*

$$\sum_{\boldsymbol{d} \in \mathbf{D}_K} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n})\right] \leqslant \sum_{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!]} \left| \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij}|\mathbf{n}+b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i \, \forall (i,j) \in \mathcal{U}}}\right] \right|.$$

We denote the inner sum (over $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$*) by* $\Sigma_{\mathcal{S},\mathcal{L},\mathcal{U}}$*.*

Proof. Summing over all possible sets of single indices, we have

$$\sum_{\boldsymbol{d} \in \mathbf{D}_K} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n})\right] = \sum_{\mathcal{S} \subset [\![K]\!] \times [\![J]\!]} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n})\right].$$

Let us 'condition' on the value of the sequence $\left(\mathbf{1}_{d_{ij}|\mathbf{n}+b_i}\right)_{(i,j) \in ([\![K]\!] \times [\![J]\!]) \setminus \mathcal{S}}$. We do this by summing over all possible decompositions of $([\![K]\!] \times [\![J]\!]) \setminus \mathcal{S}$ as a disjoint union $\mathcal{L} \sqcup \mathcal{U}$, which gives

$$\sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n})\right] = \sum_{([\![K]\!] \times [\![J]\!]) \setminus \mathcal{S} = \mathcal{L} \sqcup \mathcal{U}} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij}|\mathbf{n}+b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i \, \forall (i,j) \in \mathcal{U}}}\right].$$

The result now follows from the triangle inequality. $\qquad\square$

Note that we had to leave the single indices $\mathcal{S}$ aside in order to exploit the cancellation from the factors $\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}$ when $(i, j) \in \mathcal{S}$.

Definition 6.8 (Lit and unlit indices)**.** In the expression $\Sigma_{\mathcal{S},\mathcal{L},\mathcal{U}}$, we call $\mathcal{L}$ the set of *lit* indices and $\mathcal{U}$ the set of *unlit* indices. By construction, $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. In particular, the primes d_{ij} with $(i, j) \in \mathcal{L} \sqcup \mathcal{U}$ are all repeated in the array $(d_{ij})_{(i,j) \in [\![K]\!] \times [\![J]\!]}$.

6.3. **Walks with many unlit indices.** The next lemma shows that $\Sigma_{\mathcal{S},\mathcal{L},\mathcal{U}}$ is small when there are many unlit indices.

Lemma 6.9. *Let* $\mathcal{S}$*,* $\mathcal{L}$*,* $\mathcal{U}$ *be sets such that* $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ *and* $|\mathcal{U}| \geqslant K^{2\varepsilon_1}$*. Then*

$$\Sigma_{\mathcal{S},\mathcal{L},\mathcal{U}} \ll 1.$$

Proof. Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$. We start by using the trivial bound

$$\mathbb{E}\left[\left|w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij}|\mathbf{n}+b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i \, \forall (i,j) \in \mathcal{U}}}\right|\right] \leqslant \mathbb{E}\left[\mathbf{1}_{d_{ij}|\mathbf{n}+b_i \, \forall (i,j) \in \mathcal{L}} \prod_{(i,j) \in \mathcal{S}} \left|\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}\right| \prod_{(i,j) \in \mathcal{U}} \frac{1}{d_{ij}}\right]$$

$$\leqslant \prod_{p|\rho_{\boldsymbol{d};\mathcal{S} \sqcup \mathcal{L}}} \frac{2}{p} \prod_{(i,j) \in \mathcal{U}} \frac{1}{d_{ij}}.$$

Next, we observe that, by definition of single, lit and unlit indices,

$$\mathbf{1}_{p|\rho_{\boldsymbol{d};\mathcal{S} \sqcup \mathcal{L}}} + \tfrac{1}{2} |\{(i, j) \in \mathcal{U} : d_{ij} = p\}| \geqslant \mathbf{1}_{p|\rho_{\boldsymbol{d}}},$$

for all $p \in \mathcal{P}$. Indeed, if $p \mid \rho_{\boldsymbol{d}}$ and $p \nmid \rho_{\boldsymbol{d};\mathcal{S} \sqcup \mathcal{L}}$, there are at least two indices (i, j) such that $p = d_{ij}$, which must be unlit.

Since all primes in $\mathcal{P}$ are $\geqslant H_0$, this implies that

$$\prod_{p \mid \rho_{\boldsymbol{d}; \mathcal{S} \sqcup \mathcal{L}}} \frac{2}{p} \prod_{(i,j) \in \mathcal{U}} \frac{1}{d_{ij}} \leqslant H_0^{-|\mathcal{U}|/2} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{2}{p}.$$

Therefore,

$$\Sigma_{\mathcal{S}, \mathcal{L}, \mathcal{U}} \leqslant H_0^{-|\mathcal{U}|/2} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{2}{p},$$

which is $\ll H_0^{-|\mathcal{U}|/2} K^{2KJ}$ by Lemma 6.4. Recall that $\log H_0 \gg K^{1-\varepsilon_1}$, while $|\mathcal{U}| \geqslant K^{2\varepsilon_1}$ by assumption. Thus, $H_0^{|\mathcal{U}|/2} \gg \exp\left(K^{1+\frac{\varepsilon_1}{2}}\right) \gg K^{2KJ}$ and the conclusion follows. $\qquad\square$

6.4. Strategy for single and repeated primes.

By Lemma 6.7, our task is reduced to showing that $\Sigma_{\mathcal{S}, \mathcal{L}, \mathcal{U}} \ll \left(e^{O(J)} V^{2J/3}\right)^K$ for every possible decomposition of $[\![K]\!] \times [\![J]\!]$ into three sets $\mathcal{S}$, $\mathcal{L}$ and $\mathcal{U}$. We just dealt with the case where there are many unlit indices. Let us briefly outline how we plan to handle the single and lit indices.

For single indices (i, j) we want to exploit the fact that each factor $\mathbf{1}_{d_{ij} \mid \mathbf{n} + b_i} - \frac{1}{d_{ij}}$ appearing in $w_{\boldsymbol{d}}(\mathbf{n})$ has mean zero and is more or less independent from the other factors. Recall that

$$w_{\boldsymbol{d}}(\mathbf{n}) = \prod_{i \in [\![K]\!]} W\left(\omega_{\mathcal{P}}(\mathbf{n} + b_i)\right) \mathbf{1}_{\mathbf{n} + b_i \in Y_L} \prod_{p \mid d_i} \left(\mathbf{1}_{p \mid \mathbf{n} + b_i} - \frac{1}{p}\right).$$

If the terms $W\left(\omega_{\mathcal{P}}(\mathbf{n} + b_i)\right)$ and $\mathbf{1}_{\mathbf{n} + b_i \in Y_L}$ were not there, the factor $\mathbf{1}_{d_{ij} \mid \mathbf{n} + b_i} - \frac{1}{d_{ij}}$ would be genuinely independent from the rest of the expression, if $(i, j) \in \mathcal{S}$. However, this is not exactly the case here. Instead of obtaining full cancellation as in Section 4.1, we will obtain a smaller amount of cancellation using a Laplace transform computation.

If there are many lit indices, we will show that there are only a small number of $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$ such that the conditions $\{d_{ij} \mid n + b_i : (i, j) \in \mathcal{L}\}$ can be simultaneously satisfied. Thus, the terms

$$\mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n} + b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n} + b_i \, \forall (i,j) \in \mathcal{U}}}\right]$$

can be close to 1 for some $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$, but for most $\boldsymbol{d}$ they will actually vanish, and $\Sigma_{\mathcal{S}, \mathcal{L}, \mathcal{U}}$ will be sufficiently small as a result. To be able to show this, the extra terms $W\left(\omega_{\mathcal{P}}(\mathbf{n} + b_i)\right)$ and $\mathbf{1}_{\mathbf{n} + b_i \in Y_L}$ will be essential – in fact, we have already seen in Section 4.3 that the conclusion would not hold if the $W\left(\omega_{\mathcal{P}}(\mathbf{n} + b_i)\right)$ terms were removed.

6.5. Divisibility conditions from lit indices.

In this section, we show that the divisibilities $d_{ij} \mid \mathbf{n} + b_i$, for $(i, j) \in \mathcal{L}$, induce conditions on $\boldsymbol{d}$ that are actually independent of $\mathbf{n}$. It is these conditions that will later allow us to bound the contribution of the walks with many lit indices.

Lemma 6.10. *Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Suppose that $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$ is such that*

$$(33) \qquad \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n} + b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n} + b_i \, \forall (i,j) \in \mathcal{U}}}\right] \neq 0.$$

Then the following hold.

(1) *Whenever two indices $(i, j), (i', j) \in \mathcal{L}$ are such that $d_{ij} = d_{i'j}$, we have*

$$d_{ij} \mid b_{i'} - b_i.$$

(2) *For every $k \in [\![R]\!]$, there are at most $2JV$ distinct primes $p \mid \rho_{\boldsymbol{d}}$ for which there exists an index $(i, j) \in \mathcal{L}$ such that $p = d_{ij}$ and $p \mid b_i - b_k$.*

Proof. Let $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$ be such that (33) holds. In particular, there exists some $n \in \mathbb{Z}$ such that

- $d_{ij} \mid n + b_i$ for all $(i, j) \in \mathcal{L}$,
- and $\omega_{\mathcal{P}}(n + b_i) \leqslant 2JV$ for all $i \in [\![R]\!]$.

Suppose first that there are two indices $(i, j), (i', j) \in \mathcal{L}$ such that $d_{ij} = d_{i'j}$. Since $(i, j), (i', j) \in \mathcal{L}$ we have $d_{ij} \mid n + b_i$ and $d_{ij} = d_{i'j} \mid n + b_{i'}$, and thus $d_{ij} \mid b_i - b_{i'}$. Hence (1) is satisfied.

Let $k \in [\![R]\!]$. On the one hand, for all $(i, j) \in \mathcal{L}$, having $d_{ij} \mid b_i - b_k$ implies that $d_{ij} \mid n + b_k$, because we also know that $d_{ij} \mid n + b_i$ as $(i, j) \in \mathcal{L}$. On the other hand, by assumption we know that $n + b_k$ has at most $2JV$ prime factors in $\mathcal{P}$. Therefore, there can be at most $2JV$ distinct primes p such that $p = d_{ij}$ for some $(i, j) \in \mathcal{L}$ and $p \mid b_i - b_k$, which proves (2). $\square$

In addition to properties (1) and (2) of Lemma 6.10, there is one more condition that comes from the terms $\mathbf{1}_{Y_L}$ in $w_{\boldsymbol{d}}(\mathbf{n})$. To state it, we need to define the non-backtracking part of a walk, also known as the reduced walk. Roughly speaking, backtracking is when a walk retraces its steps.

Definition 6.11. Let $\boldsymbol{d} \in \mathbb{Z}^R$. We define the *reduced walk* to be the vector $\widetilde{\boldsymbol{d}}$ obtained by recursively removing pairs of consecutive entries d_i, d_{i+1} with $d_{i+1} = -d_i$, until this is no longer possible.

We write $\widetilde{R}$ for the length of $\widetilde{\boldsymbol{d}}$. Thus, if $\boldsymbol{d} \in \mathbf{D}_R$, then $\widetilde{\boldsymbol{d}} \in \mathbf{D}_{\widetilde{R}}$.

Example 6.12. The above definition is best understood with an example: if

$$\boldsymbol{d} = (+5, -4, -1, +2, -2, +4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9)$$

then we may successively delete pairs of backtracking steps as follows:

$$
\begin{aligned}
&(+5, -4, -1, +2, -2, +4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9)\\
&(+5, -4, -1, \qquad\quad +4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9)\\
&(+5, -4, -1, \qquad\quad +4, \qquad\quad -4, -1, -9, -7, +7, +8, -8, +9)\\
&(+5, -4, -1, \qquad\qquad\qquad\qquad -1, -9, -7, +7, +8, -8, +9)\\
&(+5, -4, -1, \qquad\qquad\qquad\qquad -1, -9, \qquad\quad +8, -8, +9)\\
&(+5, -4, -1, \qquad\qquad\qquad\qquad -1, -9, \qquad\qquad\qquad +9)\\
&(+5, -4, -1, \qquad\qquad\qquad\qquad -1 \qquad\qquad\qquad\qquad\quad).
\end{aligned}
$$

Therefore, $\widetilde{\boldsymbol{d}} = (+5, -4, -1, -1)$.

Lemma 6.13. *Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Let $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$. Let $\boldsymbol{d}'$ be a vector obtained by recursively removing some pairs of backtracking steps from $\boldsymbol{d}$ (but not necessarily all).[3] Let R' be the length of $\boldsymbol{d}'$. There is a canonical injection*

$$\iota : [\![R']\!] \to [\![R]\!]$$

such that $d'_k = d_{\iota(k)}$ for all $k \in [\![R']\!]$.

---

[3] So $\boldsymbol{d}'$ could be the reduced walk $\widetilde{\boldsymbol{d}}$ or any vector obtained at an intermediate stage in the reduction process.

Define $\mathcal{S}'$ to be the set of single indices of $\boldsymbol{d}'$ (i.e. the set of pairs $(k,j) \in [\![R']\!] \times [\![J]\!]$ such that $d_{kj}'^2 = d_{\iota(k)j}^2$ does not divide $\rho_{\boldsymbol{d}'}$). We also define

$$\mathcal{L}' := \{(k,j) \in ([\![R']\!] \times [\![J]\!]) \setminus \mathcal{S}' : (\iota(k),j) \in \mathcal{L}\},$$
$$\mathcal{U}' := \{(k,j) \in ([\![R']\!] \times [\![J]\!]) \setminus \mathcal{S}' : (\iota(k),j) \in \mathcal{U}\}.$$

The following properties hold:

    (1) if $(i,j) \in \mathcal{S}$ then $i = \iota(k)$ for some $k \in [\![R']\!]$, and $(k,j) \in \mathcal{S}'$;

    (2) $[\![R']\!] \times [\![J]\!] = \mathcal{S}' \sqcup \mathcal{L}' \sqcup \mathcal{U}'$;

    (3) $|\mathcal{S}| \leqslant |\mathcal{S}'| \leqslant |\mathcal{S}| + \frac{1}{3}RJ$.

Proof.   (1) If $(i,j) \in \mathcal{S}$, then d_{ij} cannot appear in the backtracking part of $\boldsymbol{d}$, as otherwise d_{ij}^2 would divide $\rho_{\boldsymbol{d}}$. Thus, $d_{ij} \mid \rho_{\boldsymbol{d}'}$. Clearly, $d_{ij}^2 \nmid \rho_{\boldsymbol{d}'}$, as $\rho_{\boldsymbol{d}'} \mid \rho_{\boldsymbol{d}}$ and $(i,j) \in \mathcal{S}$. This means that single indices for $\boldsymbol{d}$ become single indices for $\boldsymbol{d}'$ (through ι^{-1}).

    (2) We have just seen that $\mathcal{S}' \supset \{(k,j) \in [\![R']\!] \times [\![J]\!] : (\iota(k),j) \in \mathcal{S}\}$. By definition of $\mathcal{L}'$ and $\mathcal{U}'$, this implies that $[\![R']\!] \times [\![J]\!] = \mathcal{S}' \sqcup \mathcal{L}' \sqcup \mathcal{U}'$.

    (3) We have $|\mathcal{S}'| = |\mathcal{S}| + t$, where t is the number of distinct primes p such that $p \mid \rho_{\boldsymbol{d}'}$, $p^2 \nmid \rho_{\boldsymbol{d}'}$ (p corresponds to a single index for $\boldsymbol{d}'$) and $p^2 \mid \rho_{\boldsymbol{d}}$ (p does not correspond to a single index for $\boldsymbol{d}$). Let p be a prime with these properties. Then p divides some d_i in the backtracking steps deleted in going from $\boldsymbol{d}$ to $\boldsymbol{d}'$, but since these d_i come in pairs we conclude that

$$p^2 \mid \frac{\rho_{\boldsymbol{d}}}{\rho_{\boldsymbol{d}'}}.$$

    Hence, $p^3 \mid \rho_{\boldsymbol{d}}$. Since $\rho_{\boldsymbol{d}}$ has RJ prime factors (with multiplicity), this shows that $3t \leqslant RJ$, which completes the proof.   □

Lemma 6.14. *Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Suppose that $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$ is such that*

$$(34) \qquad \qquad \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n} + b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n} + b_i \, \forall (i,j) \in \mathcal{U}}} \right] \neq 0.$$

Let $\widetilde{\boldsymbol{d}} \in \mathbf{D}_{\widetilde{R}}$ be the reduced walk, and let $\widetilde{\mathcal{S}}, \widetilde{\mathcal{L}}, \widetilde{\mathcal{U}}$ be the sets of single, lit and unlit indices associated to $\widetilde{\boldsymbol{d}}$ (as in Lemma 6.13).

    (3) For all $k_1 < k_2$ in $[\![R]\!]$ with $k_2 - k_1 < L$ and $[\![k_1, k_2]\!] \times [\![J]\!] \subset \widetilde{\mathcal{L}}$, neither $(\widetilde{d}_{k_1}, \widetilde{d}_{k_1+1}, \ldots, \widetilde{d}_{k_2})$ nor $(\widetilde{d}_{k_2}, \widetilde{d}_{k_2-1}, \ldots, \widetilde{d}_{k_1})$ are prohibited sequences.

See Definition 5.2 for the definition of prohibited sequences. Note that (3) is a property of the reduced walk $\widetilde{\boldsymbol{d}}$ only.

Proof. Let $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$ be such that (34) holds. In particular, there exists some $n \in \mathbb{Z}$ such that

    • $d_{ij} \mid n + b_i$ for all $(i,j) \in \mathcal{L}$,

    • and $n + b_i \in Y_L$ for all $i \in [\![R]\!]$.

Suppose that (3) fails. Thus, there are some $1 \leqslant k_1 < k_2 \leqslant \widetilde{R}$ with $k_2 - k_1 < L$, such that $[\![k_1, k_2]\!] \times [\![J]\!] \subset \widetilde{\mathcal{L}}$ and one of $(\widetilde{d}_{k_1}, \widetilde{d}_{k_1+1}, \ldots, \widetilde{d}_{k_2})$ or $(\widetilde{d}_{k_2}, \widetilde{d}_{k_2-1}, \ldots, \widetilde{d}_{k_1})$ is a prohibited sequence. Without loss of generality, we may assume that one of these two is a primitive prohibited sequence. Since $[\![k_1, k_2]\!] \times [\![J]\!] \subset \widetilde{\mathcal{L}}$ we know that $\widetilde{d}_k \mid n + b_{\iota(k)}$ for all $k \in [\![k_1, k_2]\!]$.

Let $k \in [\![k_1, k_2]\!]$. Note that

$$n + b_{\iota(k)} = (n + b_{\iota(k_1)}) + \sum_{\iota(k_1) \leqslant i < \iota(k)} d_i = (n + b_{\iota(k_1)}) + \sum_{k_1 \leqslant k' < k} \widetilde{d}_{k'}.$$

where the second equality follows from the definition of the reduced walk (the two sums differ by sums of pairs of backtracking steps, which cancel each other out). Therefore, if $(\widetilde{d}_{k_1}, \widetilde{d}_{k_1+1}, \ldots, \widetilde{d}_{k_2})$ is a primitive prohibited sequence, the fact that

$$\widetilde{d}_k \ \Big| \ (n + b_{\iota(k_1)}) + \sum_{k_1 \leqslant k' < k} \widetilde{d}_{k'}$$

for all $k \in [\![k_1, k_2]\!]$ implies that $n + b_{\iota(k_1)}$ belongs to the prohibited progression associated to $(\widetilde{d}_{k_1}, \widetilde{d}_{k_1+1}, \ldots, \widetilde{d}_{k_2})$. This contradicts the assumption that $n + b_{\iota(k_1)} \in Y_L$.

Similarly, if $(\widetilde{d}_{k_2}, \widetilde{d}_{k_2-1}, \ldots, \widetilde{d}_{k_1})$ is a primitive prohibited sequence, so is $(-\widetilde{d}_{k_2}, -\widetilde{d}_{k_2-1}, \ldots, -\widetilde{d}_{k_1})$, and the divisibility relations

$$-\widetilde{d}_k \ \Big| \ n + b_{\iota(k)} + \widetilde{d}_k = (n + b_{\iota(k_2)+1}) - \sum_{k < k' \leqslant k_2} \widetilde{d}_{k'}$$

imply that $n + b_{\iota(k_2)+1}$ belongs to the prohibited progression associated to $(-\widetilde{d}_{k_2}, -\widetilde{d}_{k_2-1}, \ldots, -\widetilde{d}_{k_1})$. Again, this contradicts the assumption that $n + b_{\iota(k_2)+1} \in Y_L$, and the proof is finished. $\square$

Definition 6.15. We denote by $\mathbf{D}_R^{\mathcal{S},\mathcal{L}}$ the set of all $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S}}$ satisfying conditions (1) and (2) of Lemma 6.10, and whose reduced walk $\widetilde{\boldsymbol{d}}$ satisfies condition (3) of Lemma 6.14.

The conclusion of this section is the following proposition.

Proposition 6.16. *We have*

$$\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant e^{O(KJ)} N \left(1 + \sup_{\substack{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!] \\ |\mathcal{U}| < K^{2\varepsilon_1}}} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \left| \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} | \mathbf{n}+b_i \ \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i \ \forall (i,j) \in \mathcal{U}}} \right] \right| \right).$$

Proof. By Lemmas 6.5 and 6.7, we have

$$\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant e^{O(KJ)} N \left(1 + \sup_{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!]} \sum_{\boldsymbol{d} \in \mathbf{D}_K} \left| \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} | \mathbf{n}+b_i \ \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i \ \forall (i,j) \in \mathcal{U}}} \right] \right| \right).$$

By Lemma 6.9, we may add the condition $|\mathcal{U}| < K^{2\varepsilon_1}$ in the supremum, at the cost of an error term which can be absorbed into the $e^{O(KJ)} N$ term. By Lemmas 6.10 and 6.14, we may restrict the sum to the elements of $\mathbf{D}_K^{\mathcal{S},\mathcal{L}}$ only. $\square$

7. Obtaining cancellation from single primes

We now implement the strategy of obtaining cancellation from the weights at single indices. As we mentioned in Section 6.4, the factors $W\big(\omega_{\mathcal{P}}(\mathbf{n} + b_i)\big)$ and $\mathbf{1}_{\mathbf{n}+b_i \in Y_L}$ prevent us from obtaining total cancellation. Instead, we will obtain a weaker amount of cancellation, that improves as the number of single indices increases. Namely, for every single index, we will save a factor $V^{-1/2}$ compared with the trivial bound. The main result of this section is the following.

Proposition 7.1. *We have*

$$\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant \sup_{\substack{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!] \\ |\mathcal{U}| < K^{2\varepsilon_1}}} e^{O(KJ)} N V^{-|\mathcal{S}|/2} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \prod_{p|\rho_{\boldsymbol{d}}} \frac{1}{p}.$$

7.1. Bad single indices.

There are special single indices for which we will not be able to obtain cancellation – we will call these indices 'bad'. Very roughly speaking, one can think of bad single indices as giving rise to certain undesired interactions between the values of $(\mathbf{1}_{d_{ij}|\mathbf{n}+b_i})_{(i,j)\in\mathcal{S}}$ and $(\omega_{\mathcal{P}}(\mathbf{n}+b_i))_{i\in[\![K]\!]}$. The definition of bad single indices may seem technical, but its relevance will become apparent in of the proof of Proposition 7.3.

Definition 7.2. Let $\mathcal{S} \subset [\![K]\!] \times [\![J]\!]$ and $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$. Define $\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})$ to be the set of $(i,j) \in \mathcal{S}$ such that either

(1) there exists $(i',j') \in \mathcal{S}$ with $b_i = b_{i'}$ and $i \neq i'$, or;

(2) there exists $(i',j') \in \mathcal{S}$ with $b_{i+1} = b_{i'+1}$ and $i \neq i'$, or;

(3) there exists $i' \in [\![K]\!]$ with $d_{ij} \mid b_{i'} - b_i$ and $b_{i'} \notin \{b_i, b_{i+1}\}$.

7.2. Cancellation over arithmetic progressions.

Assuming that the number of bad single indices is small, we can obtain some cancellation from the other single indices. To achieve this, we use a Laplace transform argument that replaces the smooth weights $W\big(\omega_{\mathcal{P}}(\mathbf{n}+b_i)\big)$ with expressions that can be directly analysed.

We also need to deal with the terms involving Y_L. Recall that Y_L is the complement of the union of all the prohibited progressions. By the inclusion-exclusion principle (in fact, a truncated version of it), it will be sufficient to bound a modified version of the expected value appearing in Proposition 6.16. In this simpler expected value, the terms $\mathbf{1}_{\mathbf{n}+b_i \in Y_L}$ are replaced with the indicator of a single arithmetic progression R.

Proposition 7.3. *Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$. Assume that*

$$\big|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})\big| \leqslant K^{1/2}.$$

Let R be an arithmetic progression whose modulus q_R is a square-free product of primes in $\mathcal{P}$. We assume that q_R is divisible by at most $K^{1-\varepsilon_1}$ primes $p \mid \rho_{\boldsymbol{d};\mathcal{S}}$.

Let

$$E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R) := \mathbb{E}\left[\mathbf{1}_{\mathbf{n} \in R} \prod_{i \in [\![K]\!]} W\big(\omega_{\mathcal{P}}(\mathbf{n}+b_i)\big) \prod_{(i,j) \in \mathcal{S}} \left(\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}\right) \mathbf{1}_{\substack{d_{ij}|\mathbf{n}+b_i \, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i \, \forall (i,j) \in \mathcal{U}}}\right].$$

Then

$$E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R) \ll e^{O(KJ)} V^{-|\mathcal{S}|/2} \prod_{p|q_R \rho_{\boldsymbol{d};\mathcal{S} \sqcup \mathcal{L}}} \frac{1}{p}.$$

If the prime d_{ij} associated to a single index (i,j) divides the modulus of R, the condition $\mathbf{n} \in R$ fixes the congruence class of $\mathbf{n}$ modulo d_{ij}, which prevents cancellation for that single index. This explains the extra assumption on the prime factors of q_R.

Proof. For $p \mid \rho_{\boldsymbol{d};\mathcal{S}}$, we write $b(p)$ for b_i, where (i,j) is the unique index with $d_{ij} = p$. Let $\alpha_1, \ldots, \alpha_{K_0}$ be the *distinct* integers appearing in the sequence $b_1, \ldots, b_K$, and let $m_1, \ldots, m_{K_0}$ be the corresponding multiplicities. Then

$$E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R) = \mathbb{E}\left[\mathbf{1}_{\mathbf{n}\in R} \prod_{1\leqslant k\leqslant K_0} W^{m_k}\big(\omega_{\mathcal{P}}(\mathbf{n}+\alpha_k)\big) \prod_{p\mid\rho_{\boldsymbol{d};\mathcal{S}}} \left(\mathbf{1}_{p\mid\mathbf{n}+b(p)} - \frac{1}{p}\right) \mathbf{1}_{\substack{d_{ij}\mid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{L} \\ d_{ij}\nmid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{U}}}\right].$$

In this proof, we will write $T := JV$ to lighten the notation. We introduce the Laplace transform $\widetilde{W}(z) := \int_0^\infty W(t)e^{-zt}\,dt$. Since W is compactly supported, $\widetilde{W}$ is entire. Moreover, for any $a \geqslant 1$, integration by parts yields

$$\widetilde{W}(z) = \frac{1}{z^a}\int_0^\infty W^{(a)}(t)e^{-zt}dt.$$

The same holds for W^m in place of W, for any power $m \geqslant 1$. Therefore, if $\mathrm{Re}(z) < 0$, by Lemma C.1 we have

$$(35) \quad |\widetilde{W^m}(z)| \leqslant \frac{1}{|z|^a}\cdot 2^m\left(\frac{Ca}{T}\right)^a\left(\int_0^{\frac{1}{2}T} e^{-\mathrm{Re}(z)t}dt + \int_{\frac{3}{2}T}^{2T} e^{-\mathrm{Re}(z)t}dt\right) \ll 2^m e^{2T\mathrm{Re}(-z)}\frac{(Ca)^a}{|z|^a T^{a-1}}$$

where C is an absolute constant. For any $\sigma \in \mathbb{R}$, the inverse Laplace transform formula says that

$$W^m(t) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \widetilde{W^m}(z)e^{zt}dz.$$

We use this formula for each term $W^{m_k}\big(\omega_{\mathcal{P}}(\mathbf{n}+\alpha_k)\big)$. Interchanging the integrals and the expected value, the expression $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)$ can thus be rewritten as a K_0-fold integral

$$\iint_{\substack{z_1,\ldots,z_{K_0} \\ \mathrm{Re}(z_k)=\sigma}} \mathbb{E}\left[\prod_{p\mid\rho_{\boldsymbol{d};\mathcal{S}}}\left(\mathbf{1}_{p\mid\mathbf{n}+b(p)} - \frac{1}{p}\right)\prod_{k\leqslant K_0}\exp\left(z_k\sum_{p\in\mathcal{P}}\mathbf{1}_{p\mid\mathbf{n}+\alpha_k}\right)\mathbf{1}_{\substack{\mathbf{n}\in R \\ d_{ij}\mid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{L} \\ d_{ij}\nmid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{U}}}\right]\prod_{k\leqslant K_0}\frac{\widetilde{W^{m_k}}(z_k)}{2\pi i}\,dz_k.$$

By independence of the variables $\mathbf{n}$ (mod p) for different primes p, we can rewrite $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)$ as

$$(36) \quad \iint_{\substack{z_1,\ldots,z_{K_0} \\ \mathrm{Re}(z_k)=\sigma}} \prod_{\substack{p\mid\rho_{\boldsymbol{d};\mathcal{S}} \\ p\nmid q_R}}\mathbb{E}\left[\left(\mathbf{1}_{p\mid\mathbf{n}+b(p)} - \frac{1}{p}\right)\exp\left(\sum_{k\leqslant K_0}z_k\mathbf{1}_{p\mid\mathbf{n}+\alpha_k}\right)\right]\cdot Z\cdot\prod_{k\leqslant K_0}\frac{\widetilde{W^{m_k}}(z_k)}{2\pi i}\,dz_k,$$

where

$$Z = \mathbb{E}\left[\mathbf{1}_{\substack{\mathbf{n}\in R \\ d_{ij}\mid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{L} \\ d_{ij}\nmid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{U}}}\prod_{\substack{p\mid\rho_{\boldsymbol{d};\mathcal{S}} \\ p\mid q_R}}\left(\mathbf{1}_{p\mid\mathbf{n}+b(p)} - \frac{1}{p}\right)\prod_{\substack{p\in\mathcal{P} \\ p\nmid\frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}},q_R)}}}\exp\left(\sum_{k\leqslant K_0}z_k\mathbf{1}_{p\mid\mathbf{n}+\alpha_k}\right)\right].$$

We choose $\sigma = -1/T$; as this is negative we can bound Z trivially by

$$(37) \quad |Z| \leqslant \mathbb{E}\left[\mathbf{1}_{\substack{\mathbf{n}\in R \\ d_{ij}\mid\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{L}}}\right] \leqslant \prod_{p\mid q_R\rho_{\boldsymbol{d};\mathcal{L}}}\frac{1}{p}.$$

We now estimate

$$(38) \quad \mathbb{E}\left[\left(\mathbf{1}_{p\mid\mathbf{n}+b(p)} - \frac{1}{p}\right)\exp\left(\sum_{k\leqslant K_0}z_k\mathbf{1}_{p\mid\mathbf{n}+\alpha_k}\right)\right]$$

for each $p \mid \frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}}, q_R)}$ (this simply means that $p \mid \rho_{\boldsymbol{d};\mathcal{S}}$ and $p \nmid q_R$ as $\rho_{\boldsymbol{d};\mathcal{S}}$ and q_R are square-free). For such a prime p, define

$$(39) \qquad M_p := \{k \leqslant K_0 : p \mid \alpha_k - b(p)\}.$$

We can directly compute that

$$\mathbb{E}\left[\mathbf{1}_{\mathbf{n} \equiv -b(p) \ (\mathrm{mod}\ p)} \left(\mathbf{1}_{p|\mathbf{n}+b(p)} - \frac{1}{p}\right) \exp\left(\sum_{k \leqslant K_0} z_k \mathbf{1}_{p|\mathbf{n}+\alpha_k}\right)\right] = \frac{1}{p}\left(1 - \frac{1}{p}\right)\exp\left(\sum_{k \in M_p} z_k\right),$$

and

$$\mathbb{E}\left[\mathbf{1}_{\mathbf{n} \not\equiv -\alpha_1,\dots,-\alpha_{K_0} \ (\mathrm{mod}\ p)} \left(\mathbf{1}_{p|\mathbf{n}+b(p)} - \frac{1}{p}\right) \exp\left(\sum_{k \leqslant K_0} z_k \mathbf{1}_{p|\mathbf{n}+\alpha_k}\right)\right] = \left(1 - \frac{O(K_0)}{p}\right)\frac{-1}{p}.$$

Finally, the contribution for when $\mathbf{n} \equiv -\alpha_k \ (\mathrm{mod}\ p)$ for some $\alpha_k \not\equiv b(p) \ (\mathrm{mod}\ p)$ is $O(K_0/p^2)$. We conclude that (38) is

$$\frac{1}{p}\left(1 - \frac{1}{p}\right)\exp\left(\sum_{k \in M_p} z_k\right) - \frac{1}{p} + O\left(\frac{K_0}{p^2}\right) = \frac{1}{p}\left(\exp\left(\sum_{k \in M_p} z_k\right) - 1\right) + O\left(\frac{K_0}{p^2}\right).$$

Observe that $\left|\exp\left(\sum_{k \in M_p} z_k\right) - 1\right| \geqslant \left|\exp\left(-\sum_{k \in M_p} \frac{1}{T}\right) - 1\right| \gg T^{-1}$. This is $\geqslant K_0/p$ by our choices of parameters, since $p \geqslant H_0$ and $T = JV$. Therefore,

$$(40) \qquad \mathbb{E}\left[\left(\mathbf{1}_{p|\mathbf{n}+b(p)} - \frac{1}{p}\right) \exp\left(\sum_{k \leqslant K_0} z_k \mathbf{1}_{p|\mathbf{n}+\alpha_k}\right)\right] \ll \frac{1}{p}\left|\exp\left(\sum_{k \in M_p} z_k\right) - 1\right|.$$

Substituting (37) and (40) into (36), we obtain that $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}; R)$ is, in absolute value, at most

$$(41) \qquad e^{O(|\mathcal{S}|)}\left(\prod_{p|q_R \rho_{\boldsymbol{d};\mathcal{S} \sqcup \mathcal{L}}} \frac{1}{p}\right) \iint_{\substack{z_1,\dots,z_{K_0} \\ \mathrm{Re}(z_k)=-1/T}} \prod_{\substack{p|\rho_{\boldsymbol{d};\mathcal{S}} \\ p \nmid q_R}} \left|\exp\left(\sum_{k \in M_p} z_k\right) - 1\right| \prod_{k \leqslant K_0} \left|\widetilde{W^{m_k}}(z_k)\right| |dz_k|.$$

To bound the expression $|\exp(\sum_{k \in M_p} z_k) - 1|$ non-trivially, we decompose the ranges of integration to be able to tell when each z_k is small or large. The multiple integral in (41) is thus

$$(42) \qquad \sum_{X \subset [\![K_0]\!]} \iint_{\substack{z_1,\dots,z_{K_0} \\ \mathrm{Re}(z_k)=-1/T \\ |z_k| \leqslant V^{-1/2}\ \forall k \in X \\ |z_k| > V^{-1/2}\ \forall k \notin X}} \prod_{\substack{p|\rho_{\boldsymbol{d};\mathcal{S}} \\ p \nmid q_R}} \left|\exp\left(\sum_{k \in M_p} z_k\right) - 1\right| \prod_{k \leqslant K_0} |\widetilde{W^{m_k}}(z_k)| |dz_k|.$$

Thus X is the set of all $k \in [\![K_0]\!]$ such that $|z_k| \leqslant V^{-1/2}$. By Taylor expansion, we have

$$\left|\exp\left(\sum_{k \in M_p} z_k\right) - 1\right| \ll \begin{cases} \sum_{k \in M_p} |z_k| \ll V^{-1/2} & \text{if } M_p \subset X \text{ and } |M_p| \leqslant 2, \\ 1 & \text{otherwise.} \end{cases}$$

Thus, (42) is bounded by

$$(43) \qquad e^{O(|\mathcal{S}|)} \sum_{X \subset [\![K_0]\!]} \left[\prod_{\substack{p|\frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}}, q_R)} \\ M_p \subset X,\ |M_p| \leqslant 2}} V^{-1/2}\right] \prod_{k \in X}\left(\int_{I_{\leqslant}} |\widetilde{W^{m_k}}(z)||dz|\right) \prod_{k \notin X}\left(\int_{I_{>}} |\widetilde{W^{m_k}}(z)||dz|\right),$$

where $I_{\leqslant} = \{z : \mathrm{Re}(z) = -1/T,\ |z| \leqslant V^{-1/2}\}$ and $I_{>} = \{z : \mathrm{Re}(z) = -1/T,\ |z| > V^{-1/2}\}$.

By (35) with $a = 2$, we have

$$\int_{I_\leqslant} |\widetilde{W^{m_k}}(z)||dz| \ll \frac{2^{m_k}}{T} \int_{\mathrm{Re}(z)=-1/T} |z|^{-2}|dz| \ll 2^{m_k}.$$

For the integral over $I_>$ we have, for any $a \geqslant 2$, using (35),

$$\int_{I_>} |\widetilde{W^{m_k}}(z)||dz| \ll 2^{m_k} \frac{(Ca)^a}{T^{a-1}} \int_{I_>} |z|^{-a}|dz| \ll 2^{m_k} \left(\frac{CaV^{1/2}}{T}\right)^{a-1}.$$

Choosing $a = \lfloor T/(eCV^{1/2}) \rfloor$, we obtain the bound

$$\int_{I_>} |\widetilde{W^{m_k}}(z)||dz| \ll 2^{m_k} e^{-a} \leqslant 2^{m_k} V^{-J},$$

where the last inequality holds provided that V is larger than some absolute constant, which is the case if ε_1 is sufficiently small, by Lemma 2.4.

Notice that $\prod_{k \leqslant K_0} 2^{m_k} = 2^K$. Putting everything together, we deduce that (43) is at most

$$(44) \qquad e^{O(KJ)} \sum_{X \subset \llbracket K_0 \rrbracket} \left[\prod_{\substack{p | \frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}},q_R)} \\ M_p \subset X, \, |M_p| \leqslant 2}} V^{-1/2} \right] (V^{-J})^{K_0 - |X|}.$$

We claim that

$$(45) \qquad \left| \left\{ p \mid \frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}},q_R)} : M_p \subset X, \, |M_p| \leqslant 2 \right\} \right| \geqslant |\mathcal{S}| - 2J(K_0 - |X|) - O(K^{1-\varepsilon_1}).$$

Assuming (45), we conclude that (44) is bounded by

$$e^{O(KJ)} \sum_{X \subset \llbracket K_0 \rrbracket} V^{-|\mathcal{S}|/2} V^{J(K_0-|X|)} V^{O(K^{1-\varepsilon_1})} (V^{-J})^{K_0-|X|} \ll e^{O(KJ)} V^{-|\mathcal{S}|/2},$$

which implies the desired bound on $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)$.

It remains to prove (45). This is where we will use our assumptions on q_R and on the number of bad single indices. Since q_R has at most $K^{1-\varepsilon_1}$ prime factors $p \mid \rho_{\boldsymbol{d};\mathcal{S}}$, we have

$$\left| \left\{ p \mid \frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}},q_R)} : M_p \subset X, \, |M_p| \leqslant 2 \right\} \right| \geqslant |\{ p \mid \rho_{\boldsymbol{d};\mathcal{S}} : M_p \subset X, \, |M_p| \leqslant 2 \}| - K^{1-\varepsilon_1}$$

$$\geqslant |\mathcal{S}| - |\{ p \mid \rho_{\boldsymbol{d};\mathcal{S}} : M_p \not\subset X \}| - |\{ p \mid \rho_{\boldsymbol{d};\mathcal{S}} : |M_p| > 2 \}| - K^{1-\varepsilon_1}.$$

Observe that $\{ p \mid \rho_{\boldsymbol{d};\mathcal{S}} : |M_p| > 2 \} \subset \mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})$. Indeed, suppose that $(i,j) \in \mathcal{S}$ is such that $|M_{d_{ij}}| > 2$. This implies that there are elements $i_1, i_2, i_3 \in \llbracket K \rrbracket$ with $b_{i_1}, b_{i_2}, b_{i_3}$ pairwise distinct such that

$$b_{i_1} - b_i \equiv b_{i_2} - b_i \equiv b_{i_3} - b_i \equiv 0 \pmod{d_{ij}}.$$

Since $b_{i_1}, b_{i_2}, b_{i_3}$ are distinct, one of them is not in $\{b_i, b_{i+1}\}$. By case (3) of Definition 7.2, this is only possible if $(i,j) \in \mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})$. Recall that $|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| \leqslant K^{1/2}$ by assumption. Therefore,

$$\left| \left\{ p \mid \frac{\rho_{\boldsymbol{d};\mathcal{S}}}{(\rho_{\boldsymbol{d};\mathcal{S}},q_R)} : M_p \subset X, \, |M_p| \leqslant 2 \right\} \right| \geqslant |\mathcal{S}| - |\{ p \mid \rho_{\boldsymbol{d};\mathcal{S}} : M_p \not\subset X \}| - O(K^{1-\varepsilon_1}).$$

Hence, to prove (45), it suffices to show that, for all $j \in \llbracket J \rrbracket$ and $k \in \llbracket K_0 \rrbracket \setminus X$,

$$\left| \{ i \in \llbracket K \rrbracket : (i,j) \in \mathcal{S} \setminus \mathcal{S}_{\mathrm{bad}}(\boldsymbol{d}), \, M_{d_{ij}} \ni k \} \right| \leqslant 2.$$

Suppose otherwise. Then, there are distinct $i_1, i_2, i_3 \in [\![K]\!]$ with $(i_1, j), (i_2, j), (i_3, j) \in \mathcal{S} \setminus \mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})$, and moreover $d_{i_1 j} \mid \alpha_k - b_{i_1}$, $d_{i_2 j} \mid \alpha_k - b_{i_2}$ and $d_{i_3 j} \mid \alpha_k - b_{i_3}$. By case (3) of Definition 7.2, these divisibilities imply that

$$\alpha_k \in \{b_{i_1}, b_{i_1+1}\} \cap \{b_{i_2}, b_{i_2+1}\} \cap \{b_{i_3}, b_{i_3+1}\}.$$

However, this intersection is empty by cases (1) and (2) of Definition 7.2. This is a contradiction. This finishes the proof of (45) and hence that of Proposition 7.3. $\qquad\square$

Proposition 7.3 dealt with the case where there are few bad single indices. The following lemma states that the contribution of the remaining walks, with many bad indices, is negligible. We will prove it in Section 11, along with other results of the same type. The idea behind the proof is that, by Definition 7.2, bad single indices force equality or divisibility constraints, and there can only be few walks $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$ for which a large number of such constraints are satisfied.

Lemma 7.4. *Let $\mathcal{S} \subset [\![K]\!] \times [\![J]\!]$. We have*

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}} \\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

7.3. **Cancellation over Y_L.** In this section, we use Proposition 7.3 to give a bound for the expected value in Proposition 6.16 that incorporates a saving of $V^{-1/2}$ for every single index.

Recall that $\mathcal{Y}$ is the set of all prohibited progressions, and Y_L is the complement of the union of these prohibited progressions. We need to use a suitable version of the inclusion-exclusion principle to express $\mathbf{1}_{Y_L}$ as a linear combination of indicators of intersections of prohibited progressions. By linearity of expectation, we will obtain a collection of expected values that can be treated by Proposition 7.3.

The exact inclusion-exclusion formula

$$(46) \quad \mathbf{1}_{n \in Y_L} = \mathbf{1}_{n \notin P \ \forall P \in \mathcal{Y}} = 1 - \sum_{P_1 \in \mathcal{Y}} \mathbf{1}_{n \in P_1} + \sum_{\substack{P_1, P_2 \in \mathcal{Y} \\ \text{distinct}}} \mathbf{1}_{n \in P_1 \cap P_2} - \sum_{\substack{P_1, P_2, P_3 \in \mathcal{Y} \\ \text{distinct}}} \mathbf{1}_{n \in P_1 \cap P_2 \cap P_3} + \cdots.$$

has too many terms to be useful. We require a truncated version, also known as a combinatorial sieve. The combinatorial sieve we will use was developed by Helfgott and Radziwiłł [5], using ideas from the theory of the Möbius function of partially ordered sets. Its two main features are the following.

- Because the progressions $P \in \mathcal{Y}$ have composite (square-free) moduli, several intersections of progressions in $\mathcal{Y}$ can yield the same result. For example,

$$5\mathbb{Z} \cap 6\mathbb{Z} \cap 7\mathbb{Z} = 14\mathbb{Z} \cap 30\mathbb{Z} = 2\mathbb{Z} \cap 6\mathbb{Z} \cap 15\mathbb{Z} \cap 21\mathbb{Z}.$$

  Let R be a progression. In the right-hand side of (46), all of the terms $\pm \mathbf{1}_{n \in P_1 \cap \ldots \cap P_i}$ with $i \geqslant 1$ and $P_1 \cap \ldots \cap P_i = R$ can be combined, and simplify to $c_R \mathbf{1}_{n \in R}$ for some integer coefficient c_R. However, if the modulus q_R of R has k prime factors, there can be close to 2^{2^k} ways of expressions R as an intersection of distinct arithmetic progressions. This means that the most naive bound would give $|c_R| \leqslant 2^{2^k}$. This is much larger than what we can allow. Fortunately, the combinatorial interpretation[4] of this coefficient c_R means that there

---

[4]In combinatorial language, c_R is a value of the Möbius function of the partially ordered set consisting of all possible intersections of prohibited progressions.

is an exceptional amount of cancellation from the ± 1 signs, and the much more reasonable bound $|c_R| \leqslant 2^k$ holds.[5]

- A classical way to approximate the inclusion-exclusion formula is by means of the *Bonferroni inequalities*. These imply that, for any $r \geqslant 1$,

$$\mathbf{1}_{n \in Y_L} = \sum_{i=0}^{r-1} \sum_{\substack{P_1,\ldots,P_i \in \mathcal{Y} \\ \text{distinct}}} (-1)^i \mathbf{1}_{n \in P_1 \cap \ldots \cap P_i} + O\left(\sum_{\substack{P_1,\ldots,P_r \in \mathcal{Y} \\ \text{distinct}}} \mathbf{1}_{n \in P_1 \cap \ldots \cap P_r} \right).$$

In this simple version, the terms $(-1)^i \mathbf{1}_{n \in P_1 \cap \ldots \cap P_i}$ with $i < r$ are kept in the main term, and those with $i > r$ can be discarded. We require a more flexible truncation method, not just based on the number i of sets in the intersection, but on specific properties of the progressions $P_1 \cap \ldots \cap P_i$. For Helfgott and Radziwiłł [5], this cut-off was determined by the number of prime factors of the moduli of the intersections $P_1 \cap \ldots \cap P_i$. In this paper, the truncation and its analysis are significantly more technical.

The combinatorial sieve of Helfgott and Radziwiłł is stated in Proposition A.3 for a general cut-off. We provide a self-contained proof of it in Appendix A (a shortened version of that in [5]). We now apply it to rewrite the term $\mathbf{1}_{\forall i,\, n+b_i \in Y_L}$ as a suitable combination of arithmetic progressions.

Notation 7.5. Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$ and let $\boldsymbol{b}$ be the associated vector of partial sums. We write

$$\mathcal{Y} - \boldsymbol{b} := \{P - b_i : P \in \mathcal{Y},\, i \in [\![K]\!]\}.$$

We also define

$$(\mathcal{Y} - \boldsymbol{b})^{\cap} := \left\{ \bigcap_{P \in X} P : X \subset \mathcal{Y} - \boldsymbol{b} \right\},$$

the set of all possible intersections of such shifted progressions (with the convention $\bigcap_{P \in \emptyset} P := \mathbb{Z}$).

The next lemma captures our application of the combinatorial sieve. It is rather technical, and we defer its proof to Section 11.4. The statement of Lemma 7.6 can be understood as follows. In (3), the approximate inclusion-exclusion formula is given, with a main term and a remainder term. The main term is a sum over all progressions with small *rank*. The rank of a progression can be thought as a measure of its complexity. It is a quantity depending on $\boldsymbol{d}$, but its precise definition is not immediately needed and hence will only be given later, in Definition 11.3. Two simple properties of the rank are given in (1) and (2). Finally, (4) and (5) contain important bounds to control the main and remainder terms, respectively.

Lemma 7.6. *Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. For every $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$, there exists a function*

$$\mathrm{rank}_{\boldsymbol{d}} : (\mathcal{Y} - \boldsymbol{b})^{\cap} \to \mathbb{Z}^{\geqslant 0} \cup \{+\infty\}$$

satisfying the following properties.

Define the arithmetic progression $A_{\boldsymbol{d}} := \{n \in \mathbb{Z} : \forall (i,j) \in \mathcal{L},\, d_{ij} \mid n + b_i\}$.

Let $X_{\boldsymbol{d}}$ be the set of all $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ such that $\mathrm{rank}_{\boldsymbol{d}}(R) < K^{5\varepsilon_1}$. Let $\partial X_{\boldsymbol{d}}$ be the set of all $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap} \setminus X_{\boldsymbol{d}}$ of the form $R = R' \cap P$ for some $R' \in X_{\boldsymbol{d}}$ and $P \in \mathcal{Y} - \boldsymbol{b}$.

(1) *(Primes dividing the modulus) For every $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$,*

$$\omega(q_R) \leqslant LJ \,\mathrm{rank}_{\boldsymbol{d}}(R) + KJ.$$

(2) *(Primes $p \mid \rho_{\boldsymbol{d};\mathcal{S}}$ dividing the modulus) For every $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$,*

$$|\{p : p \mid q_R,\, p \mid \rho_{\boldsymbol{d};\mathcal{S}}\}| \leqslant LJ \,\mathrm{rank}_{\boldsymbol{d}}(R).$$

---

[5]Optimal bounds for c_R are due to Sagan, Yeh and Ziegler (see [12, after Corollary 2.5]). Helfgott and Radziwiłł [5] gave a one-line proof of the slightly weaker bound $|c_R| \leqslant 2^k$ (see Lemma A.2).

(3) *(Combinatorial sieve)* Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$. For all $n \in \mathbb{Z}$, we have

$$\mathbf{1}_{\forall i,\, n+b_i \in Y_L \text{ and } n \in A_{\boldsymbol{d}}} = \sum_{R \in X_{\boldsymbol{d}}} c_{R,\boldsymbol{d}} \mathbf{1}_{n \in R \cap A_{\boldsymbol{d}}} + O\left(3^{3KJ} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \mathbf{1}_{n \in R \cap A_{\boldsymbol{d}}} \right),$$

where the coefficients $c_{R,\boldsymbol{d}}$ are independent of n and satisfy $|c_{R,\boldsymbol{d}}| \leqslant 2^{2KJ}$.

(4) *(Main term bound)* We have

$$\sum_{R \in X_{\boldsymbol{d}}} \prod_{\substack{p \mid q_R \\ p \nmid \rho_{\boldsymbol{d}}}} \frac{1}{p} \ll e^{O(KJ)}.$$

(5) *(Remainder term bound)* Suppose $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$. Then

$$\sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \prod_{p \mid q_R \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

We now have all the ingredients to prove Proposition 7.1.

Proof of Proposition 7.1, assuming Lemma 7.6. By Proposition 6.16, we have

$$(47) \qquad \mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant e^{O(KJ)} N \left(1 + \sup_{\substack{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!] \\ |\mathcal{U}| < K^{2\varepsilon_1}}} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \left| \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{U}}} \right] \right| \right).$$

We can ignore those $\boldsymbol{d}$ for which $|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}$ as, by the triangle inequality and Lemma 7.4,

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}} \\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}}} \left| \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{U}}} \right] \right| \leqslant \sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}} \\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

Thus, (47) becomes

$$(48) \quad \mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant e^{O(KJ)} N \left(1 + \sup_{\substack{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!] \\ |\mathcal{U}| < K^{2\varepsilon_1}}} \sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}} \\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| \leqslant K^{1/2}}} \left| \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{U}}} \right] \right| \right).$$

Fix $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}$ with $|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| \leqslant K^{1/2}$. By definition of $w_{\boldsymbol{d}}(\mathbf{n})$ we have

$$\left| \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{U}}} \right] \right| \leqslant \left(\prod_{p \mid \rho_{\boldsymbol{d};\mathcal{U}}} \frac{1}{p} \right) E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}),$$

where $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d})$ is defined by

$$E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}) := \mathbb{E}\left[\prod_{i \in [\![K]\!]} \mathbf{1}_{\mathbf{n}+b_i \in Y_L} W\big(\omega_{\mathcal{P}}(\mathbf{n}+b_i)\big) \prod_{(i,j) \in \mathcal{S}} \left(\mathbf{1}_{d_{ij} \mid \mathbf{n}+b_i} - \frac{1}{d_{ij}} \right) \mathbf{1}_{\substack{d_{ij} \mid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{L} \\ d_{ij} \nmid \mathbf{n}+b_i\, \forall (i,j) \in \mathcal{U}}} \right].$$

By part (3) of Lemma 7.6, we can write

$$(49) \qquad E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}) = \sum_{R \in X_{\boldsymbol{d}}} c_{R,\boldsymbol{d}}\, E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R) + \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} O\Big(3^{3KJ}\, E_{\mathcal{S},\mathcal{L},\mathcal{U}}^{|\cdot|}(\boldsymbol{d};R) \Big),$$

with $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)$ as defined in Proposition 7.3 and

$$E^{|\cdot|}_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R) := \mathbb{E}\left[\prod_{(i,j)\in\mathcal{S}}\left|\mathbf{1}_{d_{ij}|\mathbf{n}+b_i} - \frac{1}{d_{ij}}\right|\mathbf{1}_{\substack{\mathbf{n}\in R\\ d_{ij}|\mathbf{n}+b_i\,\forall(i,j)\in\mathcal{L}}}\right] \leqslant \prod_{p|q_R\rho_{\boldsymbol{d};\mathcal{S}\sqcup\mathcal{L}}}\frac{1}{p}.$$

Inserting (49) into (48) shows that $\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big)$ is bounded by the sum of a main term

$$(50)\qquad e^{O(KJ)}N \sup_{\substack{\mathcal{S}\sqcup\mathcal{L}\sqcup\mathcal{U}=[\![K]\!]\times[\![J]\!]\\ |\mathcal{U}|<K^{2\varepsilon_1}}} \sum_{\substack{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}\\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})|\leqslant K^{1/2}}}\left(\prod_{p|\rho_{\boldsymbol{d};\mathcal{U}}}\frac{1}{p}\right)\sum_{R\in X_{\boldsymbol{d}}}|E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)|$$

and a remainder term which is $\ll e^{O(KJ)}N$ since, by part (5) of Lemma 7.6,

$$\sum_{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}}\left(\prod_{p|\rho_{\boldsymbol{d};\mathcal{U}}}\frac{1}{p}\right)\sum_{\substack{R\in\partial X_{\boldsymbol{d}}\\ R\cap A_{\boldsymbol{d}}\neq\emptyset}}\prod_{p|q_R\rho_{\boldsymbol{d};\mathcal{S}\sqcup\mathcal{L}}}\frac{1}{p} \leqslant \sum_{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}}\sum_{\substack{R\in\partial X_{\boldsymbol{d}}\\ R\cap A_{\boldsymbol{d}}\neq\emptyset}}\prod_{p|q_R\rho_{\boldsymbol{d}}}\frac{1}{p} \ll 1.$$

We now use Proposition 7.3 to bound the expression $E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)$ in (50). Note that the main condition on the modulus of R in Proposition 7.3 is satisfied. Indeed, by part (2) of Lemma 7.6, we have, for $R\in X_{\boldsymbol{d}}$,

$$|\{p:p\mid q_R,\,p\mid\rho_{\boldsymbol{d};\mathcal{S}}\}|\leqslant LJK^{5\varepsilon_1}\leqslant K^{1-10\varepsilon_1}JK^{5\varepsilon_1}\leqslant K^{1-\varepsilon_1}.$$

We obtain

$$\sum_{R\in X_{\boldsymbol{d}}}|E_{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d};R)| \ll e^{O(KJ)}V^{-|\mathcal{S}|/2}\left(\prod_{p|\rho_{\boldsymbol{d};\mathcal{S}\sqcup\mathcal{L}}}\frac{1}{p}\right)\sum_{R\in X_{\boldsymbol{d}}}\prod_{\substack{p|q_R\\ p\nmid\rho_{\boldsymbol{d};\mathcal{S}\sqcup\mathcal{L}}}}\frac{1}{p}.$$

The sum on the right-hand side is $\ll e^{O(KJ)}$ by part (4) of Lemma 7.6. Therefore, (50) is at most

$$e^{O(KJ)}N \sup_{\substack{\mathcal{S}\sqcup\mathcal{L}\sqcup\mathcal{U}=[\![K]\!]\times[\![J]\!]\\ |\mathcal{U}|<K^{2\varepsilon_1}}} \sum_{\substack{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}\\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})|\leqslant K^{1/2}}}V^{-|\mathcal{S}|/2}\left(\prod_{p|\rho_{\boldsymbol{d};\mathcal{U}}}\frac{1}{p}\right)\left(\prod_{p|\rho_{\boldsymbol{d};\mathcal{S}\sqcup\mathcal{L}}}\frac{1}{p}\right).$$

We conclude that

$$\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big)\leqslant e^{O(KJ)}N\left(1 + \sup_{\substack{\mathcal{S}\sqcup\mathcal{L}\sqcup\mathcal{U}=[\![K]\!]\times[\![J]\!]\\ |\mathcal{U}|<K^{2\varepsilon_1}}}V^{-|\mathcal{S}|/2}\sum_{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}}\prod_{p|\rho_{\boldsymbol{d}}}\frac{1}{p}\right).$$

To finish the proof, note that the error term $e^{O(KJ)}N$ can be absorbed into the term with the supremum. To see why this is true, note that, for $\mathcal{S}=[\![K]\!]\times[\![J]\!]$ and $\mathcal{L}=\mathcal{U}=\emptyset$, we have

$$V^{-|\mathcal{S}|/2}\sum_{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}}\prod_{p|\rho_{\boldsymbol{d}}}\frac{1}{p} = V^{-KJ/2}\prod_{j\in[\![J]\!]}\left(\sum_{\substack{p_1,\ldots,p_K\in\mathcal{P}_j\\ \mathrm{distinct}}}\prod_{i\in[\![K]\!]}\frac{1}{p_i}\right) \gg V^{-KJ/2}V^{KJ}\gg 1. \qquad\square$$

8. Predictable walks

By Proposition 7.1, our task is reduced to giving a good bound for

$$(51)\qquad\qquad \sum_{\boldsymbol{d}\in\mathbf{D}_K^{\mathcal{S},\mathcal{L}}}\prod_{p|\rho_{\boldsymbol{d}}}\frac{1}{p}.$$

This means that we have to beat the naive bound given in Lemma 6.4 by leveraging the divisibility conditions of Lemmas 6.10 and 6.14 coming from the lit indices.

We first focus on the reduced, non-backtracking walks $\widetilde{\boldsymbol{d}}$.

The divisibility conditions arising from the lit indices may form a highly complicated system with lots of dependencies. Our strategy will be to consider only a subset of these conditions, in order to obtain a non-degenerate subsystem consisting of independent constraints. This strategy of extracting a simple subsystem will be implemented in Section 9.

However, there is a sparse set of very regular walks for which this strategy fails, because the original system of conditions can be highly degenerate. These walks, which we call *predictable* walks, need to be separated first. We will treat them in this section (see Proposition 8.8, the main result of this section). The remaining *unpredictable* walks will be dealt with in Section 9.

In Section 10, we will show how to pass from non-backtracking walks to general walks.

8.1. **Predictable words.** We found it convenient to express the combinatorial properties of walks in the language of words and letters. Ultimately, words will just be sequences of primes in $\mathcal{P}_j$ for some j, since we want to understand the repetition patterns of the primes appearing in walks.

Definition 8.1. Let $\mathcal{A}$ be a finite set (the alphabet). Let $\mathcal{W}_n$ be the set of all n-letter words on $\mathcal{A}$, where no two consecutive letters are the same. Let $\mathcal{W}_n^{\neq} \subset \mathcal{W}_n$ be the set of all n-letter words on $\mathcal{A}$ with distinct letters. Let $\mathcal{W} = \bigcup_{n \geqslant 1} \mathcal{W}_n$ and $\mathcal{W}^{\neq} = \bigcup_{n \geqslant 1} \mathcal{W}_n^{\neq}$.

For $w \in \mathcal{W}_n$ and $1 \leqslant k \leqslant n$, we write $w[k]$ for the k-th letter of w. We denote by $w[*]$ the set of all letters of w.

We denote the set of all positions of the letter $\mathtt{A}$ in w by $\mathrm{Pos}(\mathtt{A}; w) := \{k \in [\![n]\!] : w[k] = \mathtt{A}\}$. For $l \in [\![n]\!]$, we also write $\mathrm{Pos}(l; w) := \{k \in [\![n]\!] : w[k] = w[l]\}$ (instead of '$\mathrm{Pos}(w[l]; w)$').

The notation $v \sqsubset w$ means that v is a *substring* of w, i.e. a sequence of consecutive letters of w.

We write $\overline{w}$ for the word obtained by writing the letters of w in the reversed order.

The *concatenation* of two words w_1 and w_2 is the word obtained by appending the letters of w_2 at the end of w_1. We denote it by $w_1 w_2$.

We now introduce a measure of the amount of structure of a word. We will do so by counting the number of letters with *constant neighbours*. These are letters for which each occurrence is always surrounded by the same set of letters. If most of the letters of a word have constant neighbours, the repetition patterns of these letters can be jointly well understood.

Definition 8.2. Let $w \in \mathcal{W}$ and $\mathtt{A} \in \mathcal{A}$. If there are two occurrences of $\mathtt{A}$ in w such that the sets of letters immediately adjacent to $\mathtt{A}$ are not the same in both occurrences, then we say that $\mathtt{A}$ has *variable neighbours in w*. Otherwise we say that $\mathtt{A}$ has *constant neighbours in w*.

For example,

w	neighbours of every occurrence of $\mathtt{A}$ in w	neighbours of $\mathtt{A}$ in w
XAYZYAXAY	{X,Y}, {X,Y}, {X,Y}	constant
AXYXAXZY	{X}, {X}	constant
XAYZYAYZXAY	{X,Y}, {Y}, {X,Y}	variable
YAXYZAXA	{X,Y}, {X,Z}, {X}	variable

Definition 8.3. A word $w \in \mathcal{W}$ is said to be *t-predictable* if the following conditions both hold.

(1) Every letter appears $\leqslant t$ times in w.

(2) There are $\leqslant t$ letters with variable neighbours in w.

Otherwise w is called *t-unpredictable*.

8.2. Counting predictable words.
Bounding the contribution of predictable walks requires us to show that there are few predictable words (up to relabelling of the letters).

For this section, we could have used the language of partitions since our primary focus is on the positions of the letters, and not the letters themselves. However, we found it more convenient to use words for Section 9, so we will use them here as well.

Lemma 8.4. *Let $n \geqslant 2$ and let $w_1, w_2 \in \mathcal{W}_n$. For $i \in \{1, 2\}$, let $L_i \subset w_i[*]$ be the set of letters*

$$L_i = \{w_i[1], w_i[2]\} \cup \{A \in w_i[*] : A \text{ has variable neighbours in } w_i\}$$

$$\cup \{B \in w_i[*] : B \text{ appears in } w_i \text{ next to a letter } A \text{ having variable neighbours in } w_i\}.$$

Suppose that $\{\mathrm{Pos}(A; w_1) : A \in L_1\} = \{\mathrm{Pos}(A; w_2) : A \in L_2\}$. Then

$$\{\mathrm{Pos}(A; w_1) : A \in w_1[*]\} = \{\mathrm{Pos}(A; w_2) : A \in w_2[*]\}.$$

In other words, the sets of positions of the letters in L_1 uniquely determine the sets of positions of all the letters of w_1.

Proof. Suppose that the conclusion does not hold, and let $k \geqslant 1$ be minimal with the property that $\mathrm{Pos}(k; w_1) \neq \mathrm{Pos}(k; w_2)$. Hence, $w_1[k] \notin L_1$ and $w_2[k] \notin L_2$ by the assumption in the statement. In particular, $k \geqslant 3$ since $w_i[1], w_i[2] \in L_i$.

Note that $w_1[k] \neq w_1[k-2]$. Indeed, if $w_1[k] = w_1[k-2]$, we would have $k \in \mathrm{Pos}(k-2; w_1)$, but $\mathrm{Pos}(k-2; w_1) = \mathrm{Pos}(k-2; w_2)$ by minimality of k, so $k \in \mathrm{Pos}(k-2; w_2)$ and thus $\mathrm{Pos}(k; w_2) = \mathrm{Pos}(k; w_1)$ which is not the case, by assumption.

By definition of L_1, both $w_1[k]$ and $w_1[k-1]$ have constant neighbours in w_1. This means that every occurrence of the letter $w_1[k-1]$ in w_1 is surrounded by the letters $w_1[k-2]$ and $w_1[k]$ (in any order). In addition, every appearance of $w_1[k]$ is adjacent to an occurrence of $w_1[k-1]$. Thus, we may describe $\mathrm{Pos}(k; w_1)$ exactly as

$$(52) \qquad \mathrm{Pos}(k; w_1) = \{l \in [\![n]\!] : \{l-1, l+1\} \cap \mathrm{Pos}(k-1; w_1) \neq \emptyset, \ l \notin \mathrm{Pos}(k-2; w_1)\}.$$

The same reasoning with w_2 shows that

$$(53) \qquad \mathrm{Pos}(k; w_2) = \{l \in [\![n]\!] : \{l-1, l+1\} \cap \mathrm{Pos}(k-1; w_2) \neq \emptyset, \ l \notin \mathrm{Pos}(k-2; w_2)\}.$$

However, $\mathrm{Pos}(k-2; w_1) = \mathrm{Pos}(k-2; w_2)$ and $\mathrm{Pos}(k-1; w_1) = \mathrm{Pos}(k-1; w_2)$ by minimality of k, so (52) and (53) imply that $\mathrm{Pos}(k; w_1) = \mathrm{Pos}(k; w_2)$, a contradiction. $\square$

The next lemma states that there are $\leqslant n^{O(t^3)}$ words $w \in \mathcal{W}_n$ which are t-predictable, considering two words equivalent if one can be obtained from the other by relabelling its letters.

Lemma 8.5. *Let $n, t \geqslant 1$. There are $\leqslant n^{O(t^3)}$ partitions of $\{1, \ldots, n\}$ of the form*

$$\{\mathrm{Pos}(A; w) : A \in w[*]\}$$

for some t-predictable word $w \in \mathcal{W}_n$.

Proof. By Lemma 8.4, it suffices to bound the number of possibilities for the set

$$(54) \qquad \{\mathrm{Pos}(A; w) : A \in L_w\},$$

where $L_w \subset w[*]$ is the set defined in Lemma 8.4 (with w in place of w_i), and w ranges over the set of t-predictable words in $\mathcal{W}_n$.

If w is t-predictable, there are $\leqslant t$ letters with variable neighbours. Moreover, every letter appears $\leqslant t$ times, so for every letter $\mathtt{A}$ there are $\leqslant 2t$ letters adjacent to an occurrence of $\mathtt{A}$. Thus, the set L_w has size $\leqslant 2 + t + t \cdot 2t \leqslant 5t^2$. For every $\mathtt{A} \in L_w$, the set $\mathrm{Pos}(\mathtt{A}; w)$ of positions of $\mathtt{A}$ in w is a subset of $[\![n]\!]$ of size $\leqslant t$, and there are $\leqslant n^t$ such sets.

Hence, there are $\leqslant (n^t)^{5t^2}$ possibilities for the set in (54), which concludes the proof. $\qquad\square$

8.3. Contribution of predictable walks. Let us introduce some notation for non-backtracking walks.

Definition 8.6. Let $R \geqslant 1$. Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$.

We define $\widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ to be the the set of all $\boldsymbol{d} \in \mathbf{D}_R^{\mathcal{S},\mathcal{L}}$ such that $d_{i+1} \neq -d_i$ for all $i \in [\![R-1]\!]$, i.e. those which are non-backtracking. In particular, by Definition 6.15, every $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ enjoys the following properties:

  (1) Whenever two indices $(i,j), (i',j) \in \mathcal{L}$ are such that $d_{ij} = d_{i'j}$, we have
$$d_{ij} \mid b_{i'} - b_i.$$

  (2) For every $k \in [\![R]\!]$, there are at most $2JV$ distinct primes $p \mid \rho_{\boldsymbol{d}}$ for which there exists an index $(i,j) \in \mathcal{L}$ such that $p = d_{ij}$ and $p \mid b_i - b_k$.

  (3) For all $k_1 < k_2$ in $[\![R]\!]$ with $k_2 - k_1 < L$ and $[\![k_1, k_2]\!] \times [\![J]\!] \subset \mathcal{L}$, neither $(d_{k_1}, d_{k_1+1}, \ldots, d_{k_2})$ nor $(d_{k_2}, d_{k_2-1}, \ldots, d_{k_1})$ are prohibited sequences.

Here we kept the usual notation: for $(i,j) \in [\![R]\!] \times [\![J]\!]$, d_{ij} is the unique prime in $\mathcal{P}_j$ dividing d_i, we write $b_i := \sum_{k<i} d_k$ and $\rho_{\boldsymbol{d}} := \prod_{i \in [\![R]\!]} d_i$.

We can now define predictable and unpredictable walks.

Definition 8.7. Let $R \geqslant 1$. Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Let $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$.

For $j \in [\![J]\!]$, we define two words $v_{j,\boldsymbol{d}}$ and $w_{j,\boldsymbol{d}}$ on the alphabet $\mathcal{P}_j$ as follows. Let $v_{j,\boldsymbol{d}}$ be the word

(55)
$$d_{1j} d_{2j} \cdots d_{Rj}.$$

This word can have repeated consecutive letters, so we define $w_{j,\boldsymbol{d}}$ to be the *compression* of $v_{j,\boldsymbol{d}}$, meaning the word formed by replacing, in $v_{j,\boldsymbol{d}}$, any string of consecutive occurrences of a letter with a single instance of that letter. Thus, $w_{j,\boldsymbol{d}} \in \mathcal{W}_r$ for some $r \leqslant R$.

We write $\mathbf{P}_R$ for the set of $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ such that, for all $j \in [\![J]\!]$, the word $w_{j,\boldsymbol{d}}$ is $K^{1/4}$-predictable. Similarly, we define $\mathbf{U}_R$ to be the set of $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ such that $w_{j,\boldsymbol{d}}$ is $K^{1/4}$-unpredictable for some $j \in [\![J]\!]$.

The next proposition bounds the contribution of predictable walks.

Proposition 8.8. *Let $1 \leqslant R \leqslant K$. Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. We have*
$$\sum_{\boldsymbol{d} \in \mathbf{P}_R} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll e^{O(KJ)} V^{|\mathcal{S}| + (|\mathcal{L}| + |\mathcal{U}|)/2}.$$

The proof resembles that of Lemma 6.4. The main difference is that we are restricting ourselves to partitions coming from $K^{1/4}$-predictable words, which prevents combinatorial explosion.

Proof. Any $\boldsymbol{d} \in \mathbf{P}_R$ induces a partition $\Pi_{\boldsymbol{d}}$ of $[\![R]\!] \times [\![J]\!]$, where (i,j) and (i',j') are in the same class if and only if $d_{ij} = d_{i'j'}$. Let us count the number of possible partitions.

Fix some $j \in [\![J]\!]$. Let $r \leqslant R$ be the length of $w_{j,\boldsymbol{d}}$. We know that $w_{j,\boldsymbol{d}}$ is $K^{1/4}$-predictable. By Lemma 8.5, there are $\leqslant R \cdot R^{O(K^{3/4})} \ll e^K$ possibilities for r and for the partition of $[\![r]\!]$ given by

$$(56) \qquad \qquad \left\{ \mathrm{Pos}(\mathtt{A}; w_{j,\boldsymbol{d}}) : \mathtt{A} \in w_{j,\boldsymbol{d}}[*] \right\}.$$

Let $\Pi_{j,\boldsymbol{d}}$ be the partition of $[\![R]\!]$ where i and i' are in the same class if and only if $v_{j,\boldsymbol{d}}[i] = v_{j,\boldsymbol{d}}[i']$, i.e. $d_{ij} = d_{i'j}$. Since $w_{j,\boldsymbol{d}}$ is the compressed word of $v_{j,\boldsymbol{d}}$, the partition $\Pi_{j,\boldsymbol{d}}$ is uniquely determined by r, the partition (56) of $[\![r]\!]$, and a sequence $(c_1, c_2, \ldots, c_r)$ of positive integers summing to R (these c_i correspond to the number of consecutive occurrences of each letter in $v_{j,\boldsymbol{d}}$). There are $\leqslant e^{O(R)}$ vectors of positive integers summing to R. Therefore, there are $\ll e^K e^{O(R)} \leqslant e^{O(K)}$ possibilities for the partition $\Pi_{j,\boldsymbol{d}}$. Since the partitions $(\Pi_{j,\boldsymbol{d}})_{j \in [\![J]\!]}$ determine $\Pi_{\boldsymbol{d}}$, we conclude that there are $\leqslant e^{O(KJ)}$ possible partitions $\Pi_{\boldsymbol{d}}$ of $[\![R]\!] \times [\![J]\!]$.

Observe that any $\boldsymbol{d} \in \mathbf{P}_R$ is fully determined by the signs of its coordinates d_i, the partition $\Pi_{\boldsymbol{d}}$ and the assignment of a prime p to every class α of this partition, with $p \in \mathcal{P}_j$ when $\alpha \subset [\![R]\!] \times \{j\}$.

Fix a partition Π of $[\![R]\!] \times [\![J]\!]$ and a sequence of signs $\sigma \in \{\pm 1\}^R$. For any $\boldsymbol{d} \in \mathbf{P}_R$ with $\Pi_{\boldsymbol{d}} = \Pi$, the number of distinct primes dividing $\rho_{\boldsymbol{d}}$ is $\leqslant |\mathcal{S}| + \frac{1}{2}(|\mathcal{L}| + |\mathcal{U}|)$, as every d_{ij} with $(i, j) \notin \mathcal{S}$ appears at least twice. Thus, the contribution of $\prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p}$ of all $\boldsymbol{d}$ with partition $\Pi_{\boldsymbol{d}} = \Pi$ and signs $(\mathrm{sign}(d_i))_{i \in [\![R]\!]} = \sigma$ is bounded by $V^{|\mathcal{S}| + (|\mathcal{L}| + |\mathcal{U}|)/2}$ (since $\sum_{p \in \mathcal{P}_j} 1/p = V_j \leqslant V$ for every j).

Thus, we obtain

$$\sum_{\boldsymbol{d} \in \mathbf{P}_R} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \leqslant 2^K e^{O(KJ)} V^{|\mathcal{S}| + (|\mathcal{L}| + |\mathcal{U}|)/2}$$

as desired. $\qquad \square$

9. TRIANGULAR SYSTEMS AND UNPREDICTABLE WALKS

The goal of this section is to prove the following proposition, which states that the contribution of non-backtracking, unpredictable walks is negligible.

Proposition 9.1. *Let $1 \leqslant R \leqslant K$. Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ and $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$. We have*

$$\sum_{\boldsymbol{d} \in \mathbf{U}_R} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

Our strategy is as follows. Every time a prime is repeated at lit indices, we obtain a divisibility condition. These conditions restrict the possibilities for $\boldsymbol{d}$, and generically we might hope to win a factor of about H_0 from each such condition, which would be more than sufficient. Unfortunately, there are many dependencies between the conditions, so it is very difficult to rule out the possibility that the system is very degenerate. However, since H_0 is much larger than K, it is enough to win a moderate number of factors H_0 to beat the trivial bound of Lemma 6.4. To do so, we extract from the original system of lit conditions a trivially non-singular subset of the constraints. These simple subsystems will be called *triangular systems*. These are triangular in the sense that, for a suitable ordering of the variables, the n-th variable is essentially determined by the n-th condition and the first $n-1$ variables.

9.1. Constraints and triangular systems.
We will often need to count the number of vectors $\boldsymbol{d} = (d_1, \ldots, d_R)$, with coordinates $d_i \in \mathcal{D}$, satisfying certain divisibility relations. The specific shape of these divisibility relations will depend on the situation. In Definition 9.2, we describe a fairly general type of divisibility relations that encompasses all the cases that will need to cover.

Definition 9.2. Let $R \geqslant 1$ and let $\boldsymbol{d} \in (\pm\mathcal{D})^R$. For $(i,j) \in [\![R]\!] \times [\![J]\!]$, write d_{ij} for the unique prime in $\mathcal{P}_j$ dividing d_i. Thus $|d_i| = \prod_{j \in [\![J]\!]} d_{ij}$. As before, we set $\rho_{\boldsymbol{d}} := \prod_{i \in [\![R]\!]} d_i$.

We define a *constraint* on $\boldsymbol{d}$ to be any predicate of the form

$$(57) \qquad\qquad d_{i_0 j_0} \ \Big| \ \sum_{i \in I} d_i + \kappa$$

for some $I \subset [\![R]\!]$, $(i_0, j_0) \in [\![R]\!] \times [\![J]\!]$ and $\kappa \in \mathbb{Z}$. We denote this constraint by $C_{I, i_0, j_0, \kappa}(\boldsymbol{d})$.

This constraint (57) should be viewed as a polynomial divisibility condition on the primes d_{ij}. In most of our applications, κ will be zero.

We now define what it means for a prime to be *absent* from a constraint, and *involved* in a constraint.

Definition 9.3. A prime $p \in \mathcal{P}$ is *absent* from a constraint '$d_{i_0 j_0} \mid \sum_{i \in I} d_i + \kappa$' if $p \neq d_{i_0 j_0}$ and $p \nmid d_i$ for all $i \in I$.

The definition of a prime p being *involved* in a constraint is not just the negation of the property in Definition 9.3, because we want to make sure that the constraint is not 'degenerate' when viewed as a condition on p. For example, consider the constraint $d_{11} \mid d_2 + d_3 = d_{21} d_{22} + d_{31} d_{32}$ (with $J = 2$). If $d_{11} = d_{21} = d_{31}$, this constraint will be satisfied regardless of the exact values of the primes d_{ij}, so we would like to say that none of the d_{ij} are involved in this constraint.

Definition 9.4. A prime $p \in \mathcal{P}$ is said to be *involved* in a constraint '$d_{i_0 j_0} \mid \sum_{i \in I} d_i + z$' if (at least) one of the following holds:

(i) $z = 0$, $\displaystyle\sum_{i \in I} d_i = 0$ and $\displaystyle\sum_{\substack{i \in I \\ p \mid d_i}} d_i \neq 0$, or

(ii) $z = 0$, $p = d_{i_0 j_0}$ and $\displaystyle\sum_{\substack{i \in I \\ p \nmid d_i}} d_i \neq 0$, or

(iii) $p \neq d_{i_0 j_0}$ and $\displaystyle\sum_{\substack{i \in I \\ p \mid d_i}} d_i \not\equiv 0 \pmod{d_{i_0 j_0}}$.

If case (i) holds, we will say that p is (i)-involved in the corresponding constraint. We similarly define (ii)-involved and (iii)-involved primes.

Definition 9.4 is by no means the most natural or general possible, but it is well adapted to the cases we will encounter.

In our applications, R will be fixed, and we will want to give an upper bound for the number of vectors $\boldsymbol{d}$ satisfying certain systems of constraints. Since constraints are non-linear divisibility conditions to very large, possibly distinct moduli $d_{i_0 j_0}$, these systems of constraints can be quite complicated to handle. We will use the basic 'substitution method', which only really works for *triangular systems*.

Definition 9.5. A *triangular system of T constraints* on $\boldsymbol{d}$ is a sequence $C_1(\boldsymbol{d}), \ldots, C_T(\boldsymbol{d})$ of constraints on $\boldsymbol{d}$ such that, for each $t \in [\![T]\!]$, there is a prime p_t involved in $C_t(\boldsymbol{d})$ and absent from $C_1(\boldsymbol{d}), C_2(\boldsymbol{d}), \ldots, C_{t-1}(\boldsymbol{d})$.

We will say that a triangular system of constraints on $\boldsymbol{d}$ has *complexity* $(T; c, B)$ if it is of the form $\big(C_{I_t, i_t, j_t, \kappa}(\boldsymbol{d}) \big)_{t \in [\![T]\!]}$, where each I_t is a union of at most c discrete intervals, and $|\kappa| \leqslant B$ (in particular, this integer κ is the same for all constraints).

Lemma 9.6. *Let $1 \leqslant T \leqslant R \leqslant 2K$. Let $B \geqslant 1$. Let $\mathbf{T} \subset (\pm\mathcal{D})^R$ be a set such that each $\boldsymbol{d} \in \mathbf{T}$ satisfies a triangular system of complexity $(T; 3, B)$ (thus, the system may depend on $\boldsymbol{d}$). Then*

$$\sum_{\boldsymbol{d} \in \mathbf{T}} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \ll BK^{11RJ} H_0^{-T/2}.$$

This lemma will be proved in Section 11. The proof consists in simple iterated substitutions, but is quite heavy on the notational side. The key takeaway is that every constraint of a triangular system produces a saving of a factor $H_0^{-1/2}$.

9.2. **Structure of unpredictable words.** The goal of this section is to prove Proposition 9.15, which states that t-unpredictable words must contain some special patterns. These patterns will allow us to extract large triangular systems for those $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ not covered by Section 8.

Recall that $\mathcal{W}$ denotes the set of words on the alphabet $\mathcal{A}$ with no two consecutive equal letters, and $\mathcal{W}^{\neq} \subset \mathcal{W}$ is the set of words with distinct letters.

Definition 9.7. A word $w \in \mathcal{W}$ contains n *separated repetitions* if it has a substring of the form

$$\mathtt{A}_1 \ldots \mathtt{A}_1 \ldots \mathtt{A}_2 \ldots \mathtt{A}_2 \ldots \cdots \ldots \mathtt{A}_n \ldots \mathtt{A}_n,$$

for some non-necessarily distinct letters $\mathtt{A}_1, \ldots, \mathtt{A}_n$. The three dots $\ldots$ represent a string of letters of arbitrary length (possibly empty). In other words, there are $k_1 < l_1 < k_2 < \ldots < k_n < l_n$ such that $w[k_i] = w[l_i]$ for all i.

Lemma 9.8. *Let $m \geqslant 10$. Let $k_1 < k_2 < \ldots < k_m$ be positive integers. Let $w \in \mathcal{W}$ be a word of length $\geqslant k_m$. Then, either w contains $\gg m^{1/2}$ separated repetitions, or there are $i, j \in [\![m]\!]$ with $j - i \gg m^{1/2}$ such that the substring*

$$w[k_i]w[k_i + 1] \cdots w[k_j]$$

of w has distinct letters.

Proof. Let $n = \lfloor m^{1/2} \rfloor$. If the second conclusion does not hold, there must be a repeated letter in the substring $w[k_{rn+1}]w[k_{rn+1} + 1] \cdots w[k_{(r+1)n}]$, for each $0 \leqslant r \leqslant n - 1$. This implies that w contains n separated repetitions. $\square$

Lemma 9.9. *Let $\mathtt{A}, \mathtt{B}, \mathtt{C} \in \mathcal{A}$. Let $w_1, w_2 \in \mathcal{W}^{\neq}$ be two words of the form*

$$\mathtt{A} \ldots \mathtt{B} \ldots \mathtt{C}.$$

Suppose that $\mathtt{B}$ has variable neighbours in the concatenation $w_1 w_2$ (this just means that the two letters adjacent to $\mathtt{B}$ in w_1 are not the same as the two letters adjacent to $\mathtt{B}$ in w_2).

Then, there exist substrings $v_1 \sqsubset w_1$ and $v_2 \sqsubset w_2$, both of the form $\mathtt{A} \ldots \mathtt{Y}$ for some letter $\mathtt{Y}$ (possibly equal to $\mathtt{B}$ or $\mathtt{C}$), with distinct sets of letters (i.e. $v_1[*] \neq v_2[*]$).

Proof. If $w_1[*] \neq w_2[*]$, we can just take $v_1 := w_1$, $v_2 := w_2$ and $\mathtt{Y} := \mathtt{C}$.

Otherwise, w_1 and w_2 have the same sets of letters, and thus the same length as $w_1, w_2 \in \mathcal{W}^{\neq}$. Let $k \geqslant 2$ be minimal such that $w_1[k] \neq w_2[k]$. We know that k exists, since $w_1 \neq w_2$. We set $\mathtt{X} := w_1[k]$ and $\mathtt{Y} := w_2[k]$. The letter $\mathtt{Y}$ is present in w_1 as both words have the same letters. By minimality of k, we must have $\mathtt{Y} = w_1[l]$ for some $l > k$. Set $v_1 := w_1[1]w_1[2] \cdots w_1[l] = \mathtt{A} \ldots \mathtt{X} \ldots \mathtt{Y}$ and $v_2 := w_2[1]w_2[2] \cdots w_2[k] = \mathtt{A} \ldots \mathtt{Y}$. Then $v_1[*] \neq v_2[*]$ as $\mathtt{X} \in v_1[*] \setminus v_2[*]$. $\square$

Notation 9.10. Let $w \in \mathcal{W}^{\neq}$, and suppose that w is of the form $\mathtt{A} \ldots \mathtt{X} \ldots \mathtt{Y} \ldots \mathtt{B}$. We write $w|_{\mathtt{X} \ldots \mathtt{Y}}$ for the unique substring of w of the form $\mathtt{X} \ldots \mathtt{Y}$. This is well-defined as w has distinct letters.

Lemma 9.11. *Let $m \geqslant 1$. Let $\mathtt{A}_0, \ldots, \mathtt{A}_{2m} \in \mathcal{A}$. Let $w_1, w_2 \in \mathcal{W}^{\neq}$ be two words of the form*

$$\mathtt{A}_0 \ldots \mathtt{A}_1 \ldots \mathtt{A}_2 \ldots \cdots \ldots \mathtt{A}_{2m}$$

such that, for all $1 \leqslant i \leqslant 2m-1$, the letter $\mathtt{A}_i$ has variable neighbours in the concatenation $w_1 w_2$.

Then, there are substrings $v_1 \sqsubset w_1$ and $v_2 \sqsubset w_2$, both of the form

$$\mathtt{Y}_0 \ldots \mathtt{Y}_1 \ldots \mathtt{Y}_2 \ldots \cdots \ldots \mathtt{Y}_m,$$

for some letters $\mathtt{Y}_0, \ldots, \mathtt{Y}_m$ (possibly equal to some of the $\mathtt{A}_i$) such that, for all $j \in [\![m]\!]$, the sets of letters of $v_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ and $v_2|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ are distinct.

Proof. Set $\mathtt{Y}_0 := \mathtt{A}_0$. Let $w_1^1 := w_1|_{\mathtt{Y}_0 \ldots \mathtt{A}_1 \ldots \mathtt{A}_2}$ and $w_2^1 := w_2|_{\mathtt{Y}_0 \ldots \mathtt{A}_1 \ldots \mathtt{A}_2}$. Applying Lemma 9.9 to these words w_1^1 and w_2^1, we find two further substrings $v_1^1 \sqsubset w_1^1$ and $v_2^1 \sqsubset w_2^1$ of the form $\mathtt{Y}_0 \ldots \mathtt{Y}_1$ for some common ending letter $\mathtt{Y}_1$, such that $v_1^1[*] \neq v_2^1[*]$. Notice that w_1 and w_2 are of the form

$$\mathtt{Y}_0 \ldots \mathtt{Y}_1 \ldots \mathtt{A}_3 \ldots \mathtt{A}_4 \ldots \cdots \ldots \mathtt{A}_{2m}.$$

We may thus define the substrings $w_1^2 := w_1|_{\mathtt{Y}_1 \ldots \mathtt{A}_3 \ldots \mathtt{A}_4}$ and $w_2^2 := w_2|_{\mathtt{Y}_1 \ldots \mathtt{A}_3 \ldots \mathtt{A}_4}$. Applying Lemma 9.9 again with w_1^2 and w_2^2, we obtain two substrings $v_1^2 \sqsubset w_1^2$ and $v_2^2 \sqsubset w_2^2$ of the form $\mathtt{Y}_1 \ldots \mathtt{Y}_2$, with $v_1^2[*] \neq v_2^2[*]$. In particular, w_1 and w_2 can now be written as

$$\mathtt{Y}_0 \ldots \mathtt{Y}_1 \ldots \mathtt{Y}_2 \ldots \mathtt{A}_5 \ldots \mathtt{A}_6 \ldots \cdots \ldots \mathtt{A}_{2m}.$$

We can repeat this process; after m applications of Lemma 9.9, we obtain substrings of w_1 and w_2 of the form $\mathtt{Y}_0 \ldots \mathtt{Y}_1 \ldots \mathtt{Y}_2 \ldots \cdots \ldots \mathtt{Y}_m$ with the required properties. $\square$

Lemma 9.12. *Let $m \geqslant 1$. Let $\mathtt{Y}_0, \ldots, \mathtt{Y}_{4m} \in \mathcal{A}$. Let $w_1, w_2 \in \mathcal{W}^{\neq}$ be two words of the form*

$$\mathtt{Y}_0 \ldots \mathtt{Y}_1 \ldots \mathtt{Y}_2 \ldots \cdots \ldots \mathtt{Y}_{4m}.$$

Suppose that, for all $j \in [\![4m]\!]$, the words $w_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ and $w_2|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ have distinct sets of letters.

Then, there is a pair of words $(w_1', w_2') \in \{(w_1, w_2), (w_2, w_1), (\overline{w_1}, \overline{w_2}), (\overline{w_2}, \overline{w_1})\}$ with the following properties.

There are letters $\mathtt{X}_1, \ldots, \mathtt{X}_m, \mathtt{Z}_0, \mathtt{Z}_1, \ldots, \mathtt{Z}_m, \mathtt{Z}_{m+1}$ (possibly equal to some of the $\mathtt{Y}_j$) such that w_1' is of the form

$$\mathtt{Z}_0 \ldots \mathtt{X}_1 \ldots \mathtt{Z}_1 \ldots \mathtt{X}_2 \ldots \mathtt{Z}_2 \ldots \cdots \ldots \mathtt{Z}_{m-1} \ldots \mathtt{X}_m \ldots \mathtt{Z}_m \ldots \mathtt{Z}_{m+1},$$

w_2' is of the form

$$\mathtt{Z}_0 \ldots \mathtt{Z}_1 \ldots \mathtt{Z}_2 \ldots \cdots \ldots \mathtt{Z}_m \ldots \mathtt{Z}_{m+1},$$

and, for all $j \in [\![m]\!]$, the letter $\mathtt{X}_j$ does not appear in $w_2'|_{\mathtt{Z}_0 \ldots \mathtt{Z}_j}$.

Proof. Let J_1 be the set of all $j \in [\![4m]\!]$ such that $w_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ contains a letter not appearing in $w_2|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$. Similarly, let J_2 be the set of all $j \in [\![4m]\!]$ such that $w_2|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ has a letter that is not present in $w_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$. By assumption, $J_1 \cup J_2 = [\![4m]\!]$, so one of J_1 and J_2 has size $\geqslant 2m$. Without loss of generality, assume that $|J_1| \geqslant 2m$, swapping w_1 and w_2 if necessary.

Let $j \in J_1$, and let $\mathtt{X}$ be a letter present in $w_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ but not in $w_2|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$. The letter $\mathtt{X}$ could possibly appear in $w_2|_{\mathtt{Y}_0 \ldots \mathtt{Y}_{j-1}}$ or in $w_2|_{\mathtt{Y}_j \ldots \mathtt{Y}_{4m}}$, but not in both as $w_2 \in \mathcal{W}^{\neq}$.

We define $J_1^{<}$ to be the set of all $j \in J_1$ for which there exists a letter $\mathtt{X}$ present in $w_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ but not in $w_2|_{\mathtt{Y}_0 \ldots \mathtt{Y}_j}$. Similarly, we define $J_1^{>}$ to be the set of all $j \in J_1$ for which there exists a letter $\mathtt{X}$ of $w_1|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_j}$ not appearing in $w_2|_{\mathtt{Y}_{j-1} \ldots \mathtt{Y}_{4m}}$. By the previous observation, we have $J_1^{<} \cup J_1^{>} = J_1$, so one of $J_1^{<}$ and $J_1^{>}$ has size $\geqslant m$. Considering the reversed words if necessary, we may assume without loss of generality that $|J_1^{<}| \geqslant m$.

Let $j_1 < j_2 < \ldots < j_m$ be elements of $J_1^<$. For $i \in [\![m]\!]$, let $\mathtt{X}_i$ be a letter of $w_1|_{\mathtt{Y}_{j_i}-1\ldots\mathtt{Y}_{j_i}}$ not appearing in the substring $w_2|_{\mathtt{Y}_0\ldots\mathtt{Y}_{j_i}}$ of w_2. Then w_1 is of the form

$$\mathtt{Y}_0\ldots\mathtt{X}_1\ldots\mathtt{Y}_{j_1}\ldots\mathtt{X}_2\ldots\mathtt{Y}_{j_2}\ldots\cdots\ldots\mathtt{Y}_{j_{m-1}}\ldots\mathtt{X}_m\ldots\mathtt{Y}_{j_m}\ldots\mathtt{Y}_{4m}.$$

The lemma follows, defining $\mathtt{Z}_0 := \mathtt{Y}_0$, $\mathtt{Z}_{m+1} := \mathtt{Y}_{4m}$ and $\mathtt{Z}_i := \mathtt{Y}_{j_i}$ for all $i \in [\![m]\!]$. $\qquad\square$

Combining Lemma 9.11 and Lemma 9.12, we immediately obtain the following.

Lemma 9.13. *Let $m \geqslant 1$. Let $\mathtt{A}_0, \ldots, \mathtt{A}_{8m} \in \mathcal{A}$. Let $w_1, w_2 \in \mathcal{W}^{\neq}$ be two words of the form*

$$\mathtt{A}_0\ldots\mathtt{A}_1\ldots\mathtt{A}_2\ldots\cdots\ldots\mathtt{A}_{8m}.$$

Suppose that, for $1 \leqslant i \leqslant 8m - 1$, the letter $\mathtt{A}_i$ has variable neighbours in the concatenation w_1w_2.

Then, after possibly replacing (w_1, w_2) with an element of $\{(w_1, w_2), (w_2, w_1), (\overline{w_1}, \overline{w_2}), (\overline{w_2}, \overline{w_1})\}$, the following applies.

For some letters $\mathtt{X}_1, \ldots, \mathtt{X}_m, \mathtt{Z}_0, \mathtt{Z}_1, \ldots, \mathtt{Z}_m$ (possibly equal to some of the $\mathtt{A}_j$), there are words $v_1 \sqsubset w_1$ and $v_2 \sqsubset w_2$, with v_1 of the form

$$\mathtt{Z}_0\ldots\mathtt{X}_1\ldots\mathtt{Z}_1\ldots\mathtt{X}_2\ldots\mathtt{Z}_2\ldots\cdots\ldots\mathtt{Z}_{m-1}\ldots\mathtt{X}_m\ldots\mathtt{Z}_m$$

and v_2 of the form

$$\mathtt{Z}_0\ldots\mathtt{Z}_1\ldots\mathtt{Z}_2\ldots\cdots\ldots\mathtt{Z}_m,$$

such that, for all $j \in [\![m]\!]$, the letter $\mathtt{X}_j$ does not appear in the substring $v_2|_{\mathtt{Z}_0\ldots\mathtt{Z}_j}$.

It is a well-known combinatorial fact that from any sequence of n distinct real numbers one can always extract an increasing or decreasing subsequence of length $\gg \sqrt{n}$. We will use a similar result about *pairs* of real numbers.

Lemma 9.14. *Let S be a set of n pairs of real numbers, such that*

- *if $(a, b) \in S$ then $a < b$, and*
- *if $(a, b), (c, d) \in S$ are two distinct pairs, then $\{a, b\} \cap \{c, d\} = \emptyset$.*

There exists $S' \subset S$ of size $n' \geqslant n^{1/4}$ such that one of the following holds.[6]

  (i) $S' = \{(a_1, b_1), (a_2, b_2), \ldots, (a_{n'}, b_{n'})\}$ *for some $a_1 < b_1 < a_2 < b_2 < \cdots < b_{n'}$.*

  (ii) $S' = \{(a_1, b_1), (a_2, b_2), \ldots, (a_{n'}, b_{n'})\}$ *for some $a_1 < a_2 < \cdots < a_{n'} < b_1 < b_2 < \cdots < b_{n'}$.*

  (iii) $S' = \{(a_1, b_1), (a_2, b_2), \ldots, (a_{n'}, b_{n'})\}$ *for some $a_1 < a_2 < \cdots < a_{n'} < b_{n'} < b_{n'-1} < \cdots < b_1$.*

Proof. Define a strict partial order $\prec^1$ on S by setting $(a, b) \prec^1 (c, d)$ iff $b < c$. A well-known consequence of Dilworth's theorem states that any partially ordered set on n elements contains a chain or an antichain[7] of size $\geqslant n^{1/2}$ (see [13, Proposition 2.5.9]). If S contains a chain of size $\geqslant n^{1/2}$ for $\prec^1$, we are in case (i). Suppose that S contains an antichain A of size $\geqslant n^{1/2}$. We introduce another partial order $\prec^2$ on A by defining $(a, b) \prec^2 (c, d)$ iff $a < c < d < b$. By the same combinatorial fact, either A contains a chain for $\prec^2$ of size $\geqslant n^{1/4}$, and case (iii) applies, or A contains an antichain A' for $\prec^2$ of size $\geqslant n^{1/4}$. Suppose that the latter possibility occurs. Let $(a_1, b_1), \ldots, (a_{n'}, b_{n'})$ be the elements of A', with $a_1 < a_2 < \cdots < a_{n'}$. Since A' is an antichain for $\prec^1$, all the b_i are greater than $a_{n'}$. Since A' is also an antichain for $\prec^2$, we deduce that $b_1 < b_2 < \cdots < b_{n'}$, and we are in case (ii). $\qquad\square$

We will combine the previous lemmas to extract useful substructures in unpredictable words.

---

[6] The bound $n' \geqslant n^{1/4}$ can be improved, but that is not relevant for us.

[7] Recall that a *chain* is a totally ordered subset of a partially ordered set, and an *antichain* is a subset in which no two elements are comparable.

Proposition 9.15. *There is an absolute constant $c_1 > 0$ such that the following holds.*

Let $n, t \geqslant 10$. Let $w \in \mathcal{W}_n$ be a t-unpredictable word. Then, for some $m \geqslant t^{c_1}$, at least one of the properties below is satisfied.

(1) w *has* m *separated repetitions.*

(2) *There are words* v_1, v_2 *with all of the following properties:*

    (i) $v_1 \in \mathcal{W}^{\neq}$ *and* $v_2 \in \mathcal{W}^{\neq}$;

    (ii) $v_1 \sqsubset w$ *or* $\overline{v_1} \sqsubset w$;

    (iii) $v_2 \sqsubset w$ *or* $\overline{v_2} \sqsubset w$;

    (iv) *there are letters* $\mathtt{X}_1, \ldots, \mathtt{X}_m, \mathtt{Z}_0, \ldots, \mathtt{Z}_m$ *such that* v_1 *is of the form*

$$\mathtt{Z}_0 \ldots \mathtt{X}_1 \ldots \mathtt{Z}_1 \ldots \mathtt{X}_2 \ldots \mathtt{Z}_2 \ldots \cdots \ldots \mathtt{Z}_{m-1} \ldots \mathtt{X}_m \ldots \mathtt{Z}_m$$

    and v_2 *is of the form*

$$\mathtt{Z}_0 \ldots \mathtt{Z}_1 \ldots \mathtt{Z}_2 \ldots \cdots \ldots \mathtt{Z}_m.$$

    Moreover, for all $j \in [\![m]\!]$, *the letter* $\mathtt{X}_j$ *does not appear in* $v_2|_{\mathtt{Z}_0 \ldots \mathtt{Z}_j}$.

Proof. By definition of unpredictability, either w contains a letter repeated $> t$ times, or it has $> t$ letters with variable neighbours. In the first case, we immediately see that w has $\lfloor t/2 \rfloor$ repetitions. This is $\geqslant t^{c_1}$ if c_1 is sufficiently small.

Suppose now that there are $> t$ letters with variable neighbours in w. Let E be the set of all these letters, with the possible exception of the first and last letters of w which are discarded (to simplify the notation below). Thus, $|E| \geqslant t - 2$. For every letter $\mathtt{A} \in E$, there are two positions $1 < k_\mathtt{A} < l_\mathtt{A} < n$ such that $w[k_\mathtt{A}] = w[l_\mathtt{A}] = \mathtt{A}$, and the sets of letters adjacent to these two occurrences of $\mathtt{A}$ are different, i.e. $\{w[k_\mathtt{A} - 1], w[k_\mathtt{A} + 1]\} \neq \{w[l_\mathtt{A} - 1], w[l_\mathtt{A} + 1]\}$.

We apply Lemma 9.14 to the set $S = \{(k_\mathtt{A}, l_\mathtt{A}) : \mathtt{A} \in E\}$. If case (i) occurs, we can immediately conclude that w has $\gg t^{1/4}$ separated repetitions and we are done.

Suppose that case (ii) of Lemma 9.14 applies. This implies that, for some $c \gg 1$, there exists a subset

$$F = \{\mathtt{A}_1, \ldots, \mathtt{A}_{|F|}\} \subset E$$

of size $|F| \geqslant t^c$ such that

$$1 < k_{\mathtt{A}_1} < k_{\mathtt{A}_2} < \cdots < k_{\mathtt{A}_{|F|}} < l_{\mathtt{A}_1} < l_{\mathtt{A}_2} < \cdots < l_{\mathtt{A}_{|F|}} < n.$$

By Lemma 9.8, either w has $\gg t^{c/2}$ separated repetitions, and the first conclusion holds, or we can find a 'large' substring of $w[k_{\mathtt{A}_1}]w[k_{\mathtt{A}_1} + 1] \cdots w[k_{\mathtt{A}_{|F|}}]$ with distinct letters. Without loss of generality (by replacing F with a smaller subset, c with a smaller absolute constant and relabelling the letters), we may thus assume that the word $w[k_{\mathtt{A}_1}]w[k_{\mathtt{A}_1} + 1] \cdots w[k_{\mathtt{A}_{|F|}}]$ itself has distinct letters. By a further application of Lemma 9.8, we may also assume that the word $w[l_{\mathtt{A}_1}]w[l_{\mathtt{A}_1} + 1] \cdots w[l_{\mathtt{A}_{|F|}}]$ has distinct letters.

We apply Lemma 9.13 with $w_1 := w[k_{\mathtt{A}_1}]w[k_{\mathtt{A}_1} + 1] \cdots w[k_{\mathtt{A}_{|F|}}]$ and $w_2 := w[l_{\mathtt{A}_1}]w[l_{\mathtt{A}_1} + 1] \cdots w[l_{\mathtt{A}_{|F|}}]$. These are two words in $\mathcal{W}^{\neq}$ of the form

$$\mathtt{A}_1 \ldots \mathtt{A}_2 \ldots \mathtt{A}_3 \ldots \cdots \ldots \mathtt{A}_{|F|},$$

so the assumptions of Lemma 9.13 are satisfied (of course, we may assume that $|F| \equiv 1 \pmod 8$ without loss of generality). The conclusion of Lemma 9.13 provides us with two words v_1 and v_2 precisely satisfying the second conclusion of Proposition 9.15.

The treatment of case (iii) of Lemma 9.14 is similar. For some $c \gg 1$, there exists a subset

$$F = \{\mathtt{A}_1, \ldots, \mathtt{A}_{|F|}\} \subset E$$

of size $|F| \geqslant t^c$ such that

$$1 < k_{\mathtt{A}_1} < k_{\mathtt{A}_2} < \cdots < k_{\mathtt{A}_{|F|}} < l_{\mathtt{A}_{|F|}} < l_{\mathtt{A}_{|F|-1}} < \cdots < l_{\mathtt{A}_1} < n.$$

By two successive applications Lemma 9.8, we may assume, without loss of generality, that the substrings $w_1 := w[k_{\mathtt{A}_1}]w[k_{\mathtt{A}_1}+1]\cdots w[k_{\mathtt{A}_{|F|}}]$ and $w_2 := w[l_{\mathtt{A}_{|F|}}]w[l_{\mathtt{A}_{|F|}}-1]\cdots w[l_{\mathtt{A}_1}]$ each have distinct letters. Then, applying Lemma 9.13 with these two substrings w_1 and w_2 produces two words v_1 and v_2 with the required properties. $\qquad\square$

9.3. Contribution of non-backtracking, unpredictable walks.

We now use our combinatorial work from the previous section to prove Proposition 9.1.

For the rest of this section, we fix some $1 \leqslant R \leqslant K$ and a decomposition $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ with $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$.

Definition 9.16. Let $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$. Let $1 \leqslant x < y \leqslant R$ and $p \in \mathcal{P}$. We will say that (x,y,p) is a *divisibility triple* if $p \mid d_x$, $p \mid d_y$ and there is at least one $x < i < y$ such that $p \nmid d_i$. In particular, $y \geqslant x + 2$.

We shall say that the triple (x,y,p) is *minimal* if there is no divisibility triple (x',y',p') with $x \leqslant x' < y' \leqslant y$ and $|y'-x'| < |y-x|$.

Lemma 9.17. *If (x,y,p) is a minimal divisibility triple for $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$, then for every $q \in \mathcal{P}$, the sets $\{x \leqslant i < y : q \mid d_i\}$ and $\{x < i \leqslant y : q \mid d_i\}$ are discrete intervals.*

Proof. This is an immediate consequence of Definition 9.16. $\qquad\square$

Lemma 9.18. *Let $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$ and let (x,y,p) be a minimal divisibility triple. There is some $q \in \mathcal{P}$ such that*

$$\sum_{\substack{x<i<y \\ q \mid d_i}} d_i \not\equiv 0 \pmod{p}.$$

Proof. First, note that $p \nmid d_i$ for all $x < i < y$ by minimality of (x,y,p). For $q \in \mathcal{P}$, define

$$I(q) = \{x < i < y : q \mid d_i\}.$$

Observe that $I(q)$ is a discrete interval by Lemma 9.17 and minimality of (x,y,p).

Consider the collection $\mathcal{I}$ of all sets $I(q)$, where q ranges over the prime divisors of d_{y-1}. This is a partially ordered set (where the partial order is set inclusion). Choose a prime $q_0 \mid d_{y-1}$ such that $I(q_0)$ is minimal in $\mathcal{I}$ for inclusion. This implies that $d_{ij} = d_{(y-1)j}$ for all $i \in I(q_0)$ and all $j \in [\![J]\!]$, and thus $|d_i| = |d_{y-1}|$ for all $i \in I(q_0)$. Since $\boldsymbol{d}$ is non-backtracking, we actually have $d_i = d_{y-1}$ for all $i \in I(q_0)$. Therefore,

$$\sum_{\substack{x<i<y \\ q_0 \mid d_i}} d_i = \sum_{i \in I(q_0)} d_i = |I(q_0)| \, d_{y-1}.$$

This is not divisible by p since $p \nmid d_{y-1}$ and $0 < |I(q_0)| \leqslant R \leqslant K < H_0 \leqslant p$. $\qquad\square$

Lemma 9.19. *Let $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$, let $j \in [\![J]\!]$, and suppose that the word $w_{j,\boldsymbol{d}}$ (see Definition 8.7) has $\geqslant 2m$ separated repetitions, for some $m \geqslant K^{2\varepsilon_1}$.*

Then there are $1 \leqslant x_1 < y_1 < x_2 < y_2 < \ldots < x_m < y_m \leqslant R$ *and primes* $p_1, \ldots, p_m \in \mathcal{P}$ *such that, for all* $i \in [\![m]\!]$, (x_i, y_i, p_i) *is a minimal divisibility triple, and moreover*

$$\big([\![x_i, y_i]\!] \times [\![J]\!] \big) \cap \mathcal{U} = \emptyset.$$

Proof. The assumption that $w_{j,\boldsymbol{d}}$ has $\geqslant 2m$ separated repetitions immediately tells us that there are $1 \leqslant x_1 < y_1 < x_2 < y_2 < \ldots < x_{2m} < y_{2m} \leqslant R$ and primes $p_1, \ldots, p_{2m} \in \mathcal{P}$ such that, for all $i \in [\![2m]\!]$, (x_i, y_i, p_i) is a divisibility triple. Without loss of generality, we may assume that, for every i, the triple (x_i, y_i, p_i) is minimal, as otherwise we may replace it with a divisibility triple having a smaller value of $|y_i - x_i|$, and this process eventually stops.

To get the second property, just note that there are at most $|\mathcal{U}| \leqslant K^{2\varepsilon_1} \leqslant m$ values of i for which $\big([\![x_i, y_i]\!] \times [\![J]\!] \big) \cap \mathcal{U} \neq \emptyset$, so we may simply discard the corresponding triples. $\qquad\square$

In the following lemmas, Lemmas 9.20 and 9.21, we extract a triangular system of suitable complexity for unpredictable walks. The two lemmas correspond to the two cases in the conclusion of Proposition 9.15. They are the only places in the paper where we use of condition (3) of Lemma 6.14 on prohibited sequences, which is essential to make the combinatorial analysis work.

Lemma 9.20. *Let* $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S},\mathcal{L}}$, *let* $j \in [\![J]\!]$, *and suppose that the word* $w_{j,\boldsymbol{d}}$ *(see Definition 8.7) has* $\geqslant 2m$ *separated repetitions, for some* $m \geqslant 8K^{10\varepsilon_1}$. *Then* $\boldsymbol{d}$ *satisfies a triangular system of constraints of complexity* $(\lceil m/4 \rceil; 1, 0)$.

Proof. By Lemma 9.19, there are $1 \leqslant x_1 < y_1 < x_2 < y_2 < \ldots < x_m < y_m \leqslant R$ and $p_1, \ldots, p_m \in \mathcal{P}$ such that, for all $i \in [\![m]\!]$, (x_i, y_i, p_i) is a minimal divisibility triple, and

$$(58) \qquad \big([\![x_i, y_i]\!] \times [\![J]\!] \big) \cap \mathcal{U} = \emptyset.$$

By definition of divisibility triple, and by part (1) of Definition 8.6, for every $i \in [\![m]\!]$, we have

$$(59) \qquad p_i \; \Big| \; \sum_{x_i < z < y_i} d_z.$$

By Lemma 9.18, there is, for each n, a prime $q_i \in \mathcal{P}$ such that

$$(60) \qquad \sum_{\substack{x_i < z < y_i \\ q_i | d_z}} d_z \not\equiv 0 \pmod{p_i}.$$

Let $I^<$ be the set of all $i \in [\![m]\!]$ such that q_i does not divide $\prod_{k<i} \prod_{z \in [\![x_k, y_k]\!]} d_z$.

Suppose that $|I^<| \geqslant m/4$. Observe that (59) is a constraint C_i on $\boldsymbol{d}$ in which q_i is (iii)-involved by (60). If $i \in I^<$, we know that q_i is absent from the constraints C_k with $k < i$. Therefore, the constraints $(C_i)_{i \in I^<}$ form a triangular system of complexity $(\lceil m/4 \rceil; 1, 0)$ and we are done. Henceforth, we assume that $|I^<| < m/4$.

Now, let I_1 be the set of all $i \in [\![m]\!]$ such that $[\![x_i + 1, y_i - 1]\!] \times [\![J]\!]$ contains an index $(s_i, t_i) \in \mathcal{S}$. Suppose that $|I_1| \geqslant m/4$. We will use the previous constraints C_i, but with the $d_{s_i t_i}$ as the involved primes, in place of q_i. For $i \in I_1$, notice that $d_{s_i t_i}$ is (iii)-involved in the constraint (59), because

$$(61) \qquad \sum_{\substack{x_i < z < y_i \\ d_{s_i t_i} | d_z}} d_z = d_{s_i} \not\equiv 0 \pmod{p_i}.$$

Here we used that $(s_i, t_i) \in \mathcal{S}$ for the first equality and the minimality of (x_i, y_i, p_i) to say that $p_i \nmid d_{s_i}$. In addition, $d_{s_i t_i}$ is absent from the other constraints C_k, $k \neq i$, as $(s_i, t_i) \in \mathcal{S}$. Thus, $(C_i)_{i \in I_1}$ is a triangular system of complexity $(\lceil m/4 \rceil; 1, 0)$ satisfied by $\boldsymbol{d}$, as desired. We now assume that $|I_1| < m/4$.

Let I_2 be the set of all $i \in [\![m]\!]$ such that $\{x_i\} \times [\![J]\!]$ contains an index $(s_i, t_i) \in \mathcal{S}$ (thus $s_i = x_i$). Suppose that $|I_2| \geqslant m/2$. Then $|I_2 \setminus I^<| \geqslant m/4$. This time, we will use a different sequence of constraints. Let $i \in I_2 \setminus I^<$. By definition of $I^<$, we know that there exists

$$z_i \in \bigcup_{k<i} [\![x_k, y_k]\!]$$

such that $q_i \mid d_{z_i}$. By (58), we have $(\{z_i\} \times [\![J]\!]) \cap \mathcal{U} = \emptyset$. We also know that $q_i \mid d_{u_i}$ for some $x_i < u_i < y_i$ by (60). By (58) again, we have $(\{u_i\} \times [\![J]\!]) \cap \mathcal{U} = \emptyset$. Hence, by part (1) of Definition 8.6, we obtain the constraint

$$q_i \, \Big| \sum_{z_i < l < u_i} d_l,$$

that we call C_i'. Since

(62) $$z_i < x_i = s_i < u_i < y_i$$

and $(s_i, t_i) \in \mathcal{S}$, we have $d_{s_i t_i} \neq q_i$ and

$$\sum_{\substack{z_i < l < u_i \\ d_{s_i t_i} \mid d_l}} d_l = d_{s_i} \not\equiv 0 \pmod{q_i};$$

therefore $d_{s_i t_i}$ is (iii)-involved in C_i'. Moreover, for $k, i \in I_2 \setminus I^<$ with $k < i$, the same inequalities (62) and the fact that $(s_i, t_i) \in \mathcal{S}$ show that $d_{s_i t_i}$ is absent from C_k'. Thus, $\boldsymbol{d}$ satisfies a triangular system of complexity $(\lceil m/4 \rceil; 1, 0)$. We may assume henceforth that $|I_2| < m/2$.

We have reached the final case of the proof. We will show that this case is impossible using the prohibited sequences condition. Let $I_3 = [\![m]\!] \setminus (I_1 \cup I_2)$, so that $|I_3| \geqslant m/4$. For $i \in I_3$, by definition of I_1 and I_2, the set $[\![x_i, y_i - 1]\!] \times [\![J]\!]$ has empty intersection with $\mathcal{S}$. By (58), this implies that $[\![x_i, y_i - 1]\!] \times [\![J]\!] \subset \mathcal{L}$.

Let $i \in I_3$ and suppose for a moment that $|y_i - x_i| \leqslant L$. We claim that $(d_{x_i}, d_{x_i+1}, \ldots, d_{y_i-1})$ is a prohibited sequence (see Definition 5.2). This vector is non-backtracking as $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S}, \mathcal{L}}$; it satisfies the consecutiveness assumption by Lemma 9.17 and minimality of (x_i, y_i, p_i); and it satisfies the prohibited pattern (59). Therefore, $(d_{x_i}, d_{x_i+1}, \ldots, d_{y_i-1})$ is indeed a prohibited sequence, but this cannot happen by part (3) of Definition 8.6.

We deduce that $y_i - x_i > L$ for all $i \in I_3$. This implies

$$|I_3| L \leqslant \sum_{i \in I_3} (y_i - x_i) \leqslant R \leqslant K,$$

but that is impossible as $|I_3| \geqslant m/4 \geqslant 2K^{10\varepsilon_1}$ and $L = K^{1-10\varepsilon_1}$. This concludes the proof. $\qquad\square$

The previous lemma dealt with the first case of Proposition 9.15, when $w_{j,\boldsymbol{d}}$ has many separated repetitions. Let us now consider the second case.

Lemma 9.21. *Let $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S}, \mathcal{L}}$ and $j \in [\![J]\!]$. Suppose that the word $w_{j,\boldsymbol{d}}$ satisfies the second conclusion of Proposition 9.15 for some $m \geqslant 200 K^{10\varepsilon_1}$. Then, the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$ satisfies a triangular system of constraints of complexity $(\lceil m/(200 K^{10\varepsilon_1}) \rceil; 2, KH).$*[8]

Proof. Let $w := w_{j,\boldsymbol{d}}$. Consider the second conclusion of Proposition 9.15. There are eight possibilities:

---

[8]We work with the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$ to allow for negative signs in the constraints. The reason for this will be apparent in the proof.

- $v_1 \sqsubset w$, $v_2 \sqsubset w$ and v_1 appears before v_2 in w;[9]
- $v_1 \sqsubset w$, $v_2 \sqsubset w$ and v_1 appears after v_2 in w;
- $v_1 \sqsubset w$, $\overline{v_2} \sqsubset w$ and v_1 appears before $\overline{v_2}$ in w;
  $\vdots$
- $\overline{v_1} \sqsubset w$, $\overline{v_2} \sqsubset w$ and $\overline{v_1}$ appears after $\overline{v_2}$ in w.

We will only consider the case where $v_1 \sqsubset w$, $v_2 \sqsubset w$ and v_1 appears before v_2 in w. The proofs of the seven other cases are completely analogous and left to the reader.

In this case, the second conclusion of Proposition 9.15 tells us that there are integers

$$(63) \qquad k_0 < x_1 < k_1 < x_2 < k_2 < \cdots < x_m < k_m \leqslant l_0 < l_1 < l_2 < \cdots < l_m$$

in $[\![R]\!]$ such that $d_{k_i j} = d_{l_i j}$ for all $i \in [\![0, m]\!]$. Moreover, for all $i \in [\![m]\!]$, the prime $d_{x_i j}$ does not divide $\prod_{l_0 \leqslant z \leqslant l_i} d_z$. Furthermore, the fact that v_1 has distinct letters implies in particular that, for all $i \in [\![m-1]\!]$, $d_{k_{i+1} j} \nmid \prod_{z \in [\![k_{i-1}, k_i]\!]} d_z$, and for all $i \in [\![m]\!]$, $d_{x_i j} \nmid \prod_{k_0 \leqslant z \leqslant k_{i-1}} d_z$. These observations will be useful later.

Call an integer $i \in [\![m]\!]$ *unsuitable* if one of the following holds:

- there exists an unlit index in $[\![k_{i-1}, k_i]\!] \times [\![J]\!]$ or in $[\![l_{i-1}, l_i]\!] \times [\![J]\!]$;
- $|k_i - k_{i-1}| \geqslant L$ or $|l_i - l_{i-1}| \geqslant L$;
- there exists a divisibility triple (x, y, p) with $k_{i-1} < x < y \leqslant k_i$.

Otherwise, we shall say that i is *suitable*.

Since $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$ and $L = K^{1-10\varepsilon_1}$, the first two scenarios can only happen for $\leqslant 3K^{10\varepsilon_1}$ values of $i \in [\![m]\!]$. Moreover, if there are $\geqslant 8K^{10\varepsilon_1}$ values of i for which the third scenario occurs, then w has $\geqslant 8K^{10\varepsilon_1}$ repetitions, and we are done by Lemma 9.20. Therefore, there are at most $11K^{10\varepsilon_1}$ unsuitable integers $i \in [\![m]\!]$.

Let $[\![m_1, m_2]\!]$ be a subinterval of $[\![m]\!]$ of maximal length that does not contain any unsuitable integer. Then $m_2 - m_1 \geqslant m/(33K^{10\varepsilon_1})$.

Let $i \in [\![m_1, m_2]\!]$. Note that $(k_i, j), (l_i, j) \notin \mathcal{U}$ since i is suitable, and $(k_i, j), (l_i, j) \notin \mathcal{S}$ as $d_{k_i j} = d_{l_i j}$. This means that $(k_i, j), (l_i, j) \in \mathcal{L}$. Hence, by part (1) of Definition 8.6, we have the following constraint on $\boldsymbol{d}$:

$$d_{k_i j} \; \Big| \sum_{k_i < z < l_i} d_z.$$

We rewrite this as

$$(64) \qquad d_{k_i j} \; \Big| \; \kappa + \sum_{k_0 \leqslant z < k_i} -d_z + \sum_{l_0 < z < l_i} d_z,$$

with $\kappa := \sum_{k_0 \leqslant z \leqslant l_0} d_z$. We call this constraint C_i; it is a constraint on the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$. We will show that an appropriate subset of these constraints forms a triangular system.

Note that $|\kappa| \leqslant KH$, and κ is the same for all C_i.

We define I to be the set of all $i \in [\![m_1, m_2]\!]$ such that $\big([\![k_{i-1}, k_i]\!] \times [\![J]\!]\big) \cap \mathcal{S} \neq \emptyset$.

Suppose first that $|I| \geqslant (m_2 - m_1)/2$. Then one of the sets $I_0 := \{i \in I : i \equiv 0 \pmod{2}\}$ and $I_1 := \{i \in I : i \equiv 1 \pmod{2}\}$ has size $\geqslant (m_2 - m_1)/4$. Without loss of generality, suppose that we are in the case $|I_0| \geqslant (m_2 - m_1)/4$. For each element $i \in I_0$, there is some $(s_i, t_i) \in \mathcal{S}$ such that

---

[9]Technically speaking, we should say that there is *an occurrence of* v_1 before/after *an occurrence of* v_2 in w, as v_1 and v_2 could appear several times in w.

$\{s_i\} \in [\![k_{i-1}, k_i]\!]$. We claim that the constraints $(C_{i+1})_{i \in I_0}$ form a triangular system. Indeed, the prime $d_{s_i t_i}$ is (iii)-involved in C_{i+1} as

$$\sum_{\substack{k_0 \leqslant z < k_{i+1} \\ d_{s_i t_i} | d_z}} -d_z + \sum_{\substack{l_0 < z < l_{i+1} \\ d_{s_i t_i} | d_z}} d_z = -d_{s_i} \not\equiv 0 \pmod{d_{k_{i+1}j}},$$

using that (s_i, t_i) is a single index. The last step $d_{s_i} \not\equiv 0 \pmod{d_{k_{i+1}j}}$ follows from the above-mentioned fact that $d_{k_{i+1}j}$ does not divide $\prod_{z \in [\![k_{i-1}, k_i]\!]} d_z$. Furthermore, $d_{s_i t_i}$ is absent from C_{r+1} for all $r \in I_0$ with $r < i$, as for such r we have $s_i \notin [\![k_0, k_{r+1}-1]\!] \cup [\![l_0+1, l_{r+1}-1]\!]$. Thus, the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$ satisfies a triangular system of constraints of complexity $(\lceil (m_2 - m_1)/4 \rceil; 2, KH)$, as required.

Thus, we may assume that $|I| < (m_2 - m_1)/2$. Let $i \in [\![a, b]\!] \setminus I$. We will finally make use of the integers x_i introduced in (63). We claim that the prime $d_{x_i j}$ is (iii)-involved in the constraint C_i defined by (64).

Suppose for contradiction that $d_{x_i j}$ is not (iii)-involved in C_i. Recalling that $d_{x_i j}$ does not divide $\prod_{l_0 \leqslant z \leqslant l_i} d_z$, this means that

$$(65) \qquad \sum_{\substack{k_{i-1} < z < k_i \\ d_{x_i j} | d_z}} d_z \equiv 0 \pmod{d_{k_i j}}.$$

Since i is suitable, the set $\{k_{i-1} < z < k_i : d_{x_i j} \mid d_z\}$ is a discrete interval, by Lemma 9.17, say $\{k_{i-1} < z < k_i : d_{x_i j} \mid d_z\} = [\![r, t]\!]$. Observe that the vector $(d_{k_i}, d_{k_i-1}, \dots, d_r)$ is a prohibited sequence. Indeed, it satisfies all the assumptions of Definition 5.2: it has length $\leqslant |k_i - k_{i-1}| < L$ as i is suitable; it is non-backtracking as $\boldsymbol{d} \in \widetilde{\mathbf{D}}_R^{\mathcal{S}, \mathcal{L}}$; it meets the consecutiveness assumption by Lemma 9.17 (using that i is suitable); and finally, by (65), it satisfies the prohibited pattern

$$d_{k_i j} \; \Big| \sum_{\substack{k_{i-1} < z < k_i \\ d_{x_i j} | d_z}} d_z = \sum_{r \leqslant z \leqslant t} d_z.$$

However, using that $i \notin I$ and that i is suitable, we see that $[\![k_{i-1}, k_i]\!] \times [\![J]\!] \subset \mathcal{L}$. This contradicts part (3) of Definition 8.6. We deduce that $d_{x_i j}$ is (iii)-involved in the constraint C_i.

We know, by definition of x_i, that $d_{x_i j}$ does not divide $\prod_{l_0 \leqslant z \leqslant l_i} d_z$ or $\prod_{k_0 \leqslant z \leqslant k_{i-1}} d_z$. This implies that $d_{x_i j}$ is absent from C_k for all $k < i$. Therefore, the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$ satisfies the triangular system of constraints $(C_i)_{i \in [\![a, b]\!] \setminus I}$, which has complexity $(\lceil (m_2 - m_1)/2 \rceil; 2, KH)$. This concludes the proof. □

We recall Proposition 9.1, which was our goal for this section.

Proposition 9.1. *Let $1 \leqslant R \leqslant K$. Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ and $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$. We have*

$$\sum_{\boldsymbol{d} \in \mathbf{U}_R} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

Proof of Proposition 9.1, assuming Lemma 9.6. Let $\boldsymbol{d} \in \mathbf{U}_R$. By Definition 8.7, there is some $j \in [\![J]\!]$ such that $w_{j, \boldsymbol{d}}$ is $K^{1/4}$-unpredictable.

Let $c_1 > 0$ be the constant in the statement of Proposition 9.15. We can safely assume that $K^{c_1/8} \geqslant 400 K^{10\varepsilon_1}$ since c_1 is a fixed absolute constant (that could in principle be computed) and ε_1 is assumed to be sufficiently small.

By Proposition 9.15, one the following holds.

- The first possibility is that $w_{j,\boldsymbol{d}}$ has $\geqslant K^{c_1/4}$ separated repetitions. By Lemma 9.20, $\boldsymbol{d}$ satisfies a triangular system of complexity $(\lceil K^{c_1/4}/100 \rceil; 1, 0)$.
- Otherwise, the second conclusion of Proposition 9.15 holds with $m \geqslant K^{c_1/4}$, which means that the hypotheses of Lemma 9.21 are satisfied, and hence the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$ satisfies a triangular system of constraints of complexity $(\lceil K^{c_1/4}/(200 K^{10\varepsilon_1}) \rceil; 2, KH)$.

In either case, the concatenation of $\boldsymbol{d}$ and $-\boldsymbol{d}$ satisfies a triangular system of constraints of complexity $(\lceil 2K^{c_1/8} \rceil; 2, KH)$.

By Lemma 9.6, we obtain the bound

$$\sum_{\boldsymbol{d} \in \mathbf{U}_R} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \ll KHK^{22RJ} H_0^{-K^{c_1/8}}.$$

Note that $KHK^{22RJ} \ll K^{O(KJ)}$ as $H \leqslant e^K$ and $R \leqslant K$. Moreover, since $\log H_0 \gg K^{1-\varepsilon_1}$ and $c_1 \geqslant 80\varepsilon_1$ we have $H_0^{-K^{c_1/8}} \ll \exp\big(-K^{1+c_1/16}\big)$. Recalling that $J \leqslant \log K$, we get

$$\sum_{\boldsymbol{d} \in \mathbf{U}_R} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \ll K^{O(KJ)} e^{-K^{1+c_1/16}} \ll 1$$

as desired.   □

10. BACKTRACKING WALKS AND PROOF OF THE HIGH TRACE BOUND

In this section, we pass from non-backtracking walks to general walks. We start by bounding the number of possibilities when adding one pair of backtracking steps. We will then iterate this procedure to obtain a general bound for the backtracking part of a walk (see Proposition 10.8). At the end of this section, we will combine results from the current and previous sections to prove Proposition 3.5.

10.1. **Adding one pair of backtracking steps.**

Definition 10.1. Let $R \geqslant 2$. Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Let $\boldsymbol{d}' \in \mathbf{D}_{R-2}$ and $\boldsymbol{d} \in \mathbf{D}_R$. We say that $\boldsymbol{d}$ is an *extension of* $\boldsymbol{d}'$ if

$$\boldsymbol{d} = (d_1', d_2', \ldots, d_{R-2}', x, -x)$$

for some $x \in \pm\mathcal{D}$. The *type* of this extension is defined to be the triple $(J_{\mathcal{N}}, J_{\mathcal{L}}, J_{\mathcal{U}})$, where

(1) $J_{\mathcal{N}}$ is the set of all $j \in [\![J]\!]$ such that $d_{Rj} \nmid \rho_{\boldsymbol{d}'}$;
(2) $J_{\mathcal{L}}$ is the set of all $j \in [\![J]\!]$ such that $(R, j) \in \mathcal{L}$ and there exists $i \in [\![R-2]\!]$ with $d_{Rj} = d_{ij}$ and $(i, j) \in \mathcal{L}$;
(3) $J_{\mathcal{U}}$ is the set of all remaining $j \in [\![J]\!]$, i.e. $J_{\mathcal{U}}$ is defined by $[\![J]\!] = J_{\mathcal{N}} \sqcup J_{\mathcal{L}} \sqcup J_{\mathcal{U}}$.

Lemma 10.2. *Keeping the notations of Definition 10.1, $J_{\mathcal{U}}$ is exactly the set of $j \in [\![J]\!]$ such that*

- *either $(R, j) \in \mathcal{U}$,*
- *or $(R, j) \notin \mathcal{U}$, and the set $\{i \in [\![R-2]\!] : d_{Rj} = d_{ij}\}$ is non-empty and contained in $\mathcal{U}$.*

Proof. This is immediate by Definition 10.1.   □

Lemma 10.3. *Let $R \geqslant 2$. Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![R]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Let $J_{\mathcal{N}}, J_{\mathcal{L}}, J_{\mathcal{U}}$ be any sets such that $[\![J]\!] = J_{\mathcal{N}} \sqcup J_{\mathcal{L}} \sqcup J_{\mathcal{U}}$. Let $\boldsymbol{d}' \in \mathbf{D}_{R-2}$.*

Write $\mathrm{Ext}_{J_\mathcal{N},J_\mathcal{L},J_\mathcal{U}}^{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}')$ for the set of all extensions $\boldsymbol{d}$ of $\boldsymbol{d}'$ of type $(J_\mathcal{N}, J_\mathcal{L}, J_\mathcal{U})$ satisfying properties (1) and (2) of Lemma 6.10. Then

$$\sum_{\boldsymbol{d}\in\mathrm{Ext}_{J_\mathcal{N},J_\mathcal{L},J_\mathcal{U}}^{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}')} \prod_{\substack{p|\rho_{\boldsymbol{d}}\\ p\nmid\rho_{\boldsymbol{d}'}}} \frac{1}{p} \leqslant e^{O(J)}V^J R^{|J_\mathcal{U}|}.$$

Proof. Let us write $\mathrm{Ext}(\boldsymbol{d}')$ instead of $\mathrm{Ext}_{J_\mathcal{N},J_\mathcal{L},J_\mathcal{U}}^{\mathcal{S},\mathcal{L},\mathcal{U}}(\boldsymbol{d}')$ to shorten notation. By definition, the elements $\boldsymbol{d} \in \mathrm{Ext}(\boldsymbol{d}')$ are uniquely determined by the R-th coordinate $d_R \in \pm\mathcal{D}$. Just as any element of $\pm\mathcal{D}$, d_R is of the form

$$(66) \qquad\qquad d_R = \sigma \prod_{j\in[\![J]\!]} d_{Rj}$$

for some $\sigma \in \{\pm 1\}$ and $d_{Rj} \in \mathcal{P}_j$. Thus,

$$(67) \qquad \sum_{\boldsymbol{d}\in\mathrm{Ext}(\boldsymbol{d}')} \prod_{\substack{p|\rho_{\boldsymbol{d}}\\ p\nmid\rho_{\boldsymbol{d}'}}} \frac{1}{p} = \sum_{\sigma} \sum_{(d_{Rj})_{j\in J_\mathcal{N}}} \sum_{(d_{Rj})_{j\in J_\mathcal{L}}} \sum_{(d_{Rj})_{j\in J_\mathcal{U}}} \prod_{j\in J_\mathcal{N}} \frac{1}{d_{Rj}},$$

where the quadruple sum is over all choices of σ and $(d_{Rj})_{j\in[\![J]\!]}$ such that, defining d_R by (66) and letting $\boldsymbol{d} := (d_1', d_2', \ldots, d_{R-2}', -d_R, d_R)$, we have $\boldsymbol{d} \in \mathrm{Ext}(\boldsymbol{d}')$.

We treat the elements of $J_\mathcal{N}$, $J_\mathcal{L}$ and $J_\mathcal{U}$ separately.

For every $j \in J_\mathcal{N}$, we have

$$(68) \qquad\qquad \sum_{d_{Rj}\in\mathcal{P}_j} \frac{1}{d_{Rj}} = V_j \leqslant V.$$

Let $j \in J_\mathcal{L}$. By definition of $J_\mathcal{L}$, we know that $(R, j) \in \mathcal{L}$. We need to count the number of possibilities for d_{Rj}, given that it should be of the form $d_{Rj} = d_{ij}$ for some $i \in [\![R-2]\!]$ with $(i, j) \in \mathcal{L}$. Since $\boldsymbol{d}$ has to satisfy property (1) of Lemma 6.10, we know that d_{Rj} must be an element of the set

$$A_j := \{d_{ij} \, : \, i \in [\![R-2]\!], \, (i,j) \in \mathcal{L}, \, d_{ij} \mid b_{R-1} - b_i\}$$

(recalling that $d_{Rj} = d_{(R-1)j}$, and thus $b_{R-1} = \sum_{k<R-1} d_k \equiv b_R \pmod{d_{Rj}}$). Note that this set A_j depends only on $\boldsymbol{d}'$ and $\mathcal{L}$, which are fixed.

For $j \in J_\mathcal{U}$, we know that $d_{Rj} \mid \rho_{\boldsymbol{d}'}$, so d_{Rj} must be chosen in the set $\{d_{1j}, d_{2j}, \ldots, d_{(R-2)j}\}$. Thus, there are $\leqslant R$ possibilities for d_{Rj} when $j \in J_\mathcal{U}$.

Putting everything together, we obtain that

$$(69) \qquad \sum_{\boldsymbol{d}\in\mathrm{Ext}(\boldsymbol{d}')} \prod_{\substack{p|\rho_{\boldsymbol{d}}\\ p\nmid\rho_{\boldsymbol{d}'}}} \frac{1}{p} \leqslant 2V^{|J_\mathcal{N}|}R^{|J_\mathcal{U}|} \prod_{j\in J_\mathcal{L}} |A_j|.$$

By the AM-GM inequality, we have

$$\left(\prod_{j\in J_\mathcal{L}} |A_j|\right)^{1/|J_\mathcal{L}|} \leqslant \frac{1}{|J_\mathcal{L}|} \sum_{j\in J_\mathcal{L}} |A_j| = \frac{1}{|J_\mathcal{L}|}\left|\bigsqcup_{j\in J_\mathcal{L}} A_j\right|.$$

This is a disjoint union as $A_j \subset \mathcal{P}_j$ for all $j \in J_\mathcal{L}$, and the sets $\mathcal{P}_j$ are disjoint. Clearly, $\bigsqcup_{j\in J_\mathcal{L}} A_j$ is contained in the set of all $p \mid \rho_{\boldsymbol{d}'}$ for which there is an index $(i, j) \in \mathcal{L}$, with $i \leqslant R - 2$, such that $p = d_{ij}$ and $p \mid b_i - b_{R-1}$. If $\left|\bigsqcup_{j\in J_\mathcal{L}} A_j\right| > 2JV$, no extension $\boldsymbol{d}$ of $\boldsymbol{d}'$ can be in $\mathrm{Ext}(\boldsymbol{d}')$ as such

a $\boldsymbol{d}$ will not satisfy property (2) of Lemma 6.10. Thus, in this case, $\text{Ext}(\boldsymbol{d}')$ is empty and there is nothing to prove. Otherwise, we have

$$\left| \bigsqcup_{j \in J_{\mathcal{L}}} A_j \right| \leqslant 2JV.$$

Hence, (69) becomes

$$\sum_{\boldsymbol{d} \in \text{Ext}(\boldsymbol{d}')} \prod_{\substack{p \mid \rho_{\boldsymbol{d}} \\ p \nmid \rho_{\boldsymbol{d}'}}} \frac{1}{p} \leqslant 2V^{|J_{\mathcal{N}}|} R^{|J_{\mathcal{U}}|} \left(\frac{2JV}{|J_{\mathcal{L}}|} \right)^{|J_{\mathcal{L}}|} \leqslant e^{O(J)} V^{|J_{\mathcal{N}}| + |J_{\mathcal{L}}|} R^{|J_{\mathcal{U}}|}.$$

This concludes the proof as $|J_{\mathcal{N}}| + |J_{\mathcal{L}}| \leqslant J$. $\qquad\square$

Note that the proof of Lemma 10.3 is the only place in the paper where we have made essential use of part (2) of Lemma 6.10.

10.2. **Reconstructing a walk from its non-backtracking part.** It remains to iterate Lemma 10.3 to generate multiple pairs of backtracking steps.

For notational convenience, we have defined extensions as vectors with a pair of backtracking steps in the last two coordinates. Of course, backtracking steps can be present anywhere in a walk, not just at the end, so we need to allow for cyclic permutations if we are to use Lemma 10.3 repeatedly. This is merely a technical formality that does not affect the proof other than in terms of notation.

Definition 10.4. Let $0 \leqslant h \leqslant R$ and let $\boldsymbol{d} \in \mathbf{D}_R$. We denote by $\tau_h \boldsymbol{d}$ the vector obtained by cyclically permuting the entries of $\boldsymbol{d}$:

$$\tau_h \boldsymbol{d} := (d_{R-h+1}, d_{R-h+2}, \ldots, d_R, d_1, d_2, \ldots, d_{R-h}).$$

Example 10.5. Let $\boldsymbol{d}$ and $\widetilde{\boldsymbol{d}}$ be the vectors from Example 6.12. Observe that $\boldsymbol{d}$ may be recovered from $\widetilde{\boldsymbol{d}}$ by successive cyclic permutations and extensions:

Initial vector $\boldsymbol{d}^{(0)} := \widetilde{\boldsymbol{d}}$:	$(+5, -4, -1, -1)$
Apply permutation τ_0:	$(+5, -4, -1, -1)$
Extension $\boldsymbol{d}^{(1)}$ (by -9):	$(+5, -4, -1, -1, -9, +9)$
Apply permutation τ_1:	$(+9, +5, -4, -1, -1, -9)$
Extension $\boldsymbol{d}^{(2)}$ (by $+8$):	$(+9, +5, -4, -1, -1, -9, +8, -8)$
Apply permutation τ_2:	$(+8, -8, +9, +5, -4, -1, -1, -9)$
Extension $\boldsymbol{d}^{(3)}$ (by -7):	$(+8, -8, +9, +5, -4, -1, -1, -9, -7, +7)$
Apply permutation τ_4:	$(-1, -9, -7, +7, +8, -8, +9, +5, -4, -1)$
Extension $\boldsymbol{d}^{(4)}$ (by $+4$):	$(-1, -9, -7, +7, +8, -8, +9, +5, -4, -1, +4, -4)$
Apply permutation τ_1:	$(-4, -1, -9, -7, +7, +8, -8, +9, +5, -4, -1, +4)$
Extension $\boldsymbol{d}^{(5)}$ (by $+5$):	$(-4, -1, -9, -7, +7, +8, -8, +9, +5, -4, -1, +4, +5, -5)$
Apply permutation τ_3:	$(+4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9, +5, -4, -1)$
Extension $\boldsymbol{d}^{(6)}$ (by $+2$):	$(+4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9, +5, -4, -1, +2, -2)$
Apply permutation τ_5:	$(+5, -4, -1, +2, -2, +4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9) = \boldsymbol{d}$.

Note that in total, over the whole procedure, the first coordinate of $\widetilde{\boldsymbol{d}}$ (i.e. $+5$) has been shifted by $0 + 1 + 2 + 4 + 1 + 3 + 5 = 16$, which also corresponds to the length of $\boldsymbol{d}$.

We formalise this observation in the following lemma.

Lemma 10.6. *Let $\boldsymbol{d} \in \mathbf{D}_K$ and let $\widetilde{\boldsymbol{d}} \in \mathbf{D}_{\widetilde{K}}$ be the reduced vector. Let $M = (K - \widetilde{K})/2$. There is a canonical choice of non-negative integers $h_0, \ldots, h_M$ with $\sum_i h_i = K$ and vectors $\boldsymbol{d}^{(0)}, \ldots, \boldsymbol{d}^{(M)}$ such that*

- $\boldsymbol{d}^{(0)} = \widetilde{\boldsymbol{d}}$,
- $\boldsymbol{d}^{(i+1)}$ *is an extension of $\tau_{h_i}\boldsymbol{d}^{(i)}$ for all $i \in [\![0, M-1]\!]$,*
- $\tau_{h_M}\boldsymbol{d}^{(M)} = \boldsymbol{d}$.

Lemma 10.6 should be intuitively clear, but we provide a formal proof for completeness.

Proof. We associate to $\boldsymbol{d}$ a string s consisting of spaces, left and right parentheses, with a pair of matching parentheses for the backtracking steps and a blank space for the non-backtracking steps. For example, to the vector

$$\boldsymbol{d} = (+5, -4, -1, +2, -2, +4, +5, -5, -4, -1, -9, -7, +7, +8, -8, +9)$$

of Example 6.12 we attach the string

$$s = \underline{}\,\underline{}\,\underline{}\,\underline{}\,(\,\underline{}\,)\,(\,\underline{}\,(\,\underline{}\,)\,\underline{}\,)\,\underline{}\,\underline{}\,(\,\underline{}\,(\,\underline{}\,)\,\underline{}\,(\,\underline{}\,)\,\underline{}\,).$$

Let $e_1 > \ldots > e_M$ be the positions of the right parentheses, in decreasing order. In our example, these would be $16, 15, 13, 9, 8$ and 5. Let x_i be the e_i-th coordinate of $\boldsymbol{d}$, for $i \in [\![M]\!]$. We also set $e_0 := K$ and $e_{M+1} := 0$. For $i \in [\![0, M]\!]$, we define $h_i = e_{i+1} - e_i$.

Let $\boldsymbol{d}^{(0)} := \widetilde{\boldsymbol{d}}$. For $i \in [\![0, M-1]\!]$, let $\boldsymbol{d}^{(i+1)}$ be the extension of $\tau_{h_i}\boldsymbol{d}^{(i)}$ obtained by appending $-x_i$ and x_i at the end of $\tau_{h_i}\boldsymbol{d}^{(i)}$. Note that this is exactly reproducing the steps in Example 10.5 for a general $\boldsymbol{d}$. It is straightforward to check that $\tau_{h_M}\boldsymbol{d}^{(M)} = \boldsymbol{d}$, by construction. $\square$

To be able to apply Lemma 10.3, we need some control on the sets $J_{\mathcal{U}}$ appearing at each stage of the iterated extension procedure.

Lemma 10.7. *Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}$ with reduced vector $\widetilde{\boldsymbol{d}} \in \mathbf{D}_{\widetilde{K}}$. Let $M = (K - \widetilde{K})/2$.*

Let $h_0, \ldots, h_M$ and $\boldsymbol{d}^{(0)}, \ldots, \boldsymbol{d}^{(M)}$ be as in Lemma 10.6.

For $m \in [\![0, M]\!]$, let K_m be the length of $\boldsymbol{d}^{(m)}$. There is a canonical injection $\iota_m : [\![K_m]\!] \to [\![K]\!]$ such that $d_k^{(m)} = d_{\iota_m(k)}$ for all $k \in [\![K_m]\!]$.[10] *Let $\mathcal{S}_m$, $\mathcal{L}_m$ and $\mathcal{U}_m$ be the sets associated to $\boldsymbol{d}^{(m)}$ as in Lemma 6.13.*

The following holds.

   (i) *For all $m \in [\![0, M]\!]$, the vector $\boldsymbol{d}^{(m)}$ satisfies properties* (1) *and* (2) *of Lemma 6.10.*[11]

   (ii) *For $m \in [\![0, M-1]\!]$, let $(J_{\mathcal{N},m}, J_{\mathcal{L},m}, J_{\mathcal{U},m})$ be the type of the extension $\boldsymbol{d}^{(m+1)}$ of $\tau_{h_m}\boldsymbol{d}^{(m)}$. Then*

$$\sum_{m \in [\![0, M-1]\!]} |J_{\mathcal{U},m}| \leqslant 2\,|\mathcal{U}|.$$

Proof. Property (i). Suppose that there are indices $(k, j), (k', j) \in \mathcal{L}_m$ such that $d_{kj}^{(m)} = d_{k'j}^{(m)}$, where $d_{kj}^{(m)}$ is the unique prime in $\mathcal{P}_j$ dividing $d_k^{(m)}$. Then $(\iota(k), j), (\iota(k'), j) \in \mathcal{L}$ and

$$d_{\iota(k)j} = d_{kj}^{(m)} = d_{k'j}^{(m)} = d_{\iota(k')j}.$$

---

[10]Note that this map ι_m may not be increasing, due to the cyclic permutations.

[11]Of course, with $K_m, \mathcal{S}_m, \mathcal{L}_m, \mathcal{U}_m$ in place of $K, \mathcal{S}, \mathcal{L}, \mathcal{U}$, respectively.

By property (1) of Lemma 6.10 applied to $\boldsymbol{d}$, we have $d_{\iota(k)j} \mid b_{\iota(k')} - b_{\iota(k)}$, and thus

$$d_{kj}^{(m)} \mid \sum_{l<k'} d_l^{(m)} - \sum_{l<k} d_l^{(m)},$$

since the expression on the right only differs from $b_{\iota(k')} - b_{\iota(k)}$ by pairs of backtracking steps, which cancel each other out. This proves that $\boldsymbol{d}^{(m)}$ satisfies property (1) of Lemma 6.10. The proof that property (2) of Lemma 6.10 passes down from $\boldsymbol{d}$ to $\boldsymbol{d}^{(m)}$ is analogous and shall be omitted.

Property (ii). Let us make a preliminary observation. For $0 \leqslant m_1 < m_2 \leqslant M$, the composition

$$\iota_{m_2}^{-1} \circ \iota_{m_1} : [\![K_{m_1}]\!] \to [\![K_{m_2}]\!]$$

is well-defined, injective, and its image is contained in $[\![K_{m_2} - 2]\!]$ since the last two entries of $\boldsymbol{d}^{(m_2)}$ correspond to a new backtracking pair.

For $m \in [\![0, M-1]\!]$, by Lemma 10.2, we can write $J_{\mathcal{U},m} = J_{\mathcal{U},m}^- \sqcup J_{\mathcal{U},m}^+$, where

- $J_{\mathcal{U},m}^- = \{j \in [\![J]\!] : (K_{m+1}, j) \in \mathcal{U}_{m+1}\}$, and
- $J_{\mathcal{U},m}^+$ is the set of all $j \in [\![J]\!]$ such that $(K_{m+1}, j) \notin \mathcal{U}_{m+1}$ and

(70) $$\left\{k \in [\![K_{m+1} - 2]\!] : d_{K_{m+1}j}^{(m+1)} = d_{kj}^{(m+1)}\right\} \times \{j\}$$

  is a non-empty set contained in $\mathcal{U}_{m+1}$.

Define a map $F^- : \bigsqcup_{m \in [\![0, M-1]\!]} \left(\{m\} \times J_{\mathcal{U},m}^-\right) \to \mathcal{U}$ as follows. For $m \in [\![0, M-1]\!]$ and $j \in J_{\mathcal{U},m}$, let $F^-(m, j) := (\iota_{m+1}(K_{m+1}), j)$. We know that $(K_{m+1}, j) \in \mathcal{U}_{m+1}$, so $(\iota_{m+1}(K_{m+1}), j)$ is indeed in $\mathcal{U}$. Note that $\iota_{m+1}(K_{m+1})$ uniquely determines m. To check this, note that there cannot exist $1 \leqslant m_1 < m_2 \leqslant M$ such that $\iota_{m_1}(K_{m_1}) = \iota_{m_2}(K_{m_2})$ by our preliminary observation. Hence, F^- is injective, and thus

$$\sum_{m \in [\![0, M-1]\!]} |J_{\mathcal{U},m}^-| \leqslant |\mathcal{U}|.$$

Define a map $F^+ : \bigsqcup_{m \in [\![0, M-1]\!]} \left(\{m\} \times J_{\mathcal{U},m}^+\right) \to \mathcal{U}$ as follows. For $m \in [\![0, M-1]\!]$ and $j \in J_{\mathcal{U},m}$, let (k_{m+1}, j) be any element in the (non-empty) set (70), and define $F^+(m, j) := (\iota_{m+1}(k_{m+1}), j)$. Since $(k_{m+1}, j) \in \mathcal{U}_{m+1}$, we know that $(\iota_{m+1}(k_{m+1}), j) \in \mathcal{U}$, so F^+ is well-defined.

We shall prove that F^+ is an injective map. Suppose that $(\iota_{m_1}(k_{m_1}), j) = (\iota_{m_2}(k_{m_2}), j)$ for some $1 \leqslant m_1 < m_2 \leqslant M$ and $j \in J_{\mathcal{U},m_1-1}^+ \cap J_{\mathcal{U},m_2-1}^+$. Let $\iota := \iota_{m_2}^{-1} \circ \iota_{m_1} : [\![K_{m_1}]\!] \to [\![K_{m_2}]\!]$. By definition of ι, ι_{m_1}, ι_{m_2}, k_{m_1} and k_{m_2}, and using the equality $\iota_{m_1}(k_{m_1}) = \iota_{m_2}(k_{m_2})$, we have

$$d_{\iota(K_{m_1})j}^{(m_2)} = d_{K_{m_1}j}^{(m_1)} = d_{k_{m_1}j}^{(m_1)} = d_{\iota_{m_1}(k_{m_1})j} = d_{\iota_{m_2}(k_{m_2})j} = d_{k_{m_2}j}^{(m_2)} = d_{K_{m_2}j}^{(m_2)}.$$

Recall, moreover, that $\iota(K_{m_1}) \in [\![K_{m_2} - 2]\!]$ by our preliminary observation. Therefore, the pair $(\iota(K_{m_1}), j)$ in the set (70), with m_2 in place of $m+1$. Since $j \in J_{\mathcal{U},m_2-1}^+$, that set is contained in $\mathcal{U}_{m_2}$, so $(\iota(K_{m_1}), j) \in \mathcal{U}_{m_2}$.

On the other hand, since $j \in J_{\mathcal{U},m_1-1}^+$, we have $(K_{m_1}, j) \notin \mathcal{U}_{m_1}$, by definition of $J_{\mathcal{U},m_1-1}^+$. In addition, $(K_{m_1}, j) \notin \mathcal{S}_{m_1}$ as $d_{(K_{m_1}-1)j} = d_{K_{m_1}j}$. Thus, (K_{m_1}, j) has to be $\mathcal{L}_{m_1}$. Hence, $(\iota(K_{m_1}), j) \in \mathcal{L}_{m_2}$, which contradicts the fact that $(\iota(K_{m_1}), j) \in \mathcal{U}_{m_2}$.

Therefore, F^+ is injective, so

$$\sum_{m \in [\![0, M-1]\!]} |J_{\mathcal{U},m}^+| \leqslant |\mathcal{U}|$$

which is what we had to show.   □

We can now achieve our stated goal for this section.

Proposition 10.8. *Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ and $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$. Let $0 \leqslant R \leqslant K$ and let $\boldsymbol{d}' \in (\pm \mathcal{D})^R$.*

Then

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}} \\ \widetilde{\boldsymbol{d}} = \boldsymbol{d}'}} \prod_{\substack{p | \rho_{\boldsymbol{d}} \\ p \nmid \rho_{\boldsymbol{d}'}}} \frac{1}{p} \ll e^{O(KJ)} V^{(K-R)J/2},$$

where $\widetilde{\boldsymbol{d}}$ is the reduced vector associated to $\boldsymbol{d}$ (see Definition 6.11).

Proof. Let $M := (K - R)/2$. By Lemma 10.6, any $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}$ with $\widetilde{\boldsymbol{d}} = \boldsymbol{d}'$ is obtained from $\boldsymbol{d}'$ by a succession of cyclic permutations and extensions. We sum over all possibilities for the integers $h_0, \ldots, h_M$ characterising the cyclic permutations, and for the types $(J_{\mathcal{N},m}, J_{\mathcal{L},m}, J_{\mathcal{U},m})$ of these extensions. There are $e^{O(K)}$ tuples of non-negative integers $(h_0, \ldots, h_M)$ with sum $\leqslant K$. For every $m \in [\![0, M-1]\!]$, there are $e^{O(J)}$ decompositions of $[\![J]\!]$ into three sets $[\![J]\!] = J_{\mathcal{N},m} \sqcup J_{\mathcal{L},m} \sqcup J_{\mathcal{U},m}$. Thus, there are $e^{O(KJ)}$ possibilities for $h_0, \ldots, h_M$ and $(J_{\mathcal{N},m}, J_{\mathcal{L},m}, J_{\mathcal{U},m})_{m \in [\![0,M-1]\!]}$.

Fix some $h_0, \ldots, h_M$ and $(J_{\mathcal{N},m}, J_{\mathcal{L},m}, J_{\mathcal{U},m})_{m \in [\![0,M-1]\!]}$. By part (ii) of Lemma 10.7, we may assume that

(71)
$$\sum_{m \in [\![0,M-1]\!]} |J_{\mathcal{U},m}| \leqslant 2 |\mathcal{U}|.$$

The remaining task is to show that the sum in the statement, restricted to those $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}$ generated from $\boldsymbol{d}'$ via the cyclic permutations $\tau_{h_0}, \ldots, \tau_{h_M}$ and extensions of types $(J_{\mathcal{N},m}, J_{\mathcal{L},m}, J_{\mathcal{U},m})$, is at most $e^{O(KJ)} V^{MJ}$. We do so by repeatedly applying Lemma 10.3 to obtain the bound

$$\prod_{m \in [\![0,M-1]\!]} e^{O(J)} V^J K^{|J_{\mathcal{U},m}|} \leqslant e^{O(KJ)} V^{MJ} K^{\sum_{m \in [\![0,M-1]\!]} |J_{\mathcal{U},m}|}.$$

Note that we have used part (i) of Lemma 10.7 to be able to apply Lemma 10.3. By (71), and since $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$, the term $K^{\sum_{m \in [\![0,M-1]\!]} |J_{\mathcal{U},m}|}$ is $e^{O(K)}$. This concludes the proof. $\qquad\square$

10.3. **Proof of the high trace bound.**
Combining our work in several of the previous sections, we can finally prove the high trace bound for G.

Proof of Proposition 3.5. The weighted graph G introduced in Definition 5.4 satisfies the first two properties of Proposition 3.5: the first one by Lemma 5.6, and the second by construction.

For the trace bound, we have, by Proposition 7.1, that

$$\mathrm{Tr}\big((\mathrm{Ad}_G)^K\big) \leqslant \sup_{\substack{\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!] \\ |\mathcal{U}| < K^{2\varepsilon_1}}} e^{O(KJ)} N V^{-|\mathcal{S}|/2} \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p}.$$

Fix some sets $\mathcal{S}$, $\mathcal{L}$ and $\mathcal{U}$ with $\mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U} = [\![K]\!] \times [\![J]\!]$ and $|\mathcal{U}| < K^{2\varepsilon_1}$. It remains to show that

(72)
$$\sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \prod_{p | \rho_{\boldsymbol{d}}} \frac{1}{p} \leqslant e^{O(KJ)} V^{|\mathcal{S}|/2} V^{2KJ/3}.$$

To do this, we sum over the backtracking and non-backtracking parts separately. We first sum over all possibilities for the length R of the reduced walk, and the sets $\mathcal{S}'$, $\mathcal{L}'$ and $\mathcal{U}'$ associated to the

reduced walk (see Lemma 6.13). We then sum over all possibilities $\boldsymbol{d}'$ for the reduced walk given this data, and finally over all $\boldsymbol{d}$ with reduced walk $\widetilde{\boldsymbol{d}} = \boldsymbol{d}'$. This gives

$$(73) \qquad \sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \prod_{p|\rho_{\boldsymbol{d}}} \frac{1}{p} = \sum_{0 \leqslant R \leqslant K} \sum_{[\![R]\!] \times [\![J]\!] = \mathcal{S}' \sqcup \mathcal{L}' \sqcup \mathcal{U}'} \sum_{\boldsymbol{d}' \in \widetilde{\mathbf{D}}_R^{\mathcal{S}',\mathcal{L}'}} \prod_{p|\rho_{\boldsymbol{d}'}} \frac{1}{p} \sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}} \\ \widetilde{\boldsymbol{d}} = \boldsymbol{d}'}} \prod_{\substack{p|\rho_{\boldsymbol{d}} \\ p \nmid \rho_{\boldsymbol{d}'}}} \frac{1}{p}.$$

By Lemma 6.13, we may add the constraints $|\mathcal{U}'| \leqslant K^{2\varepsilon_1}$ and $|\mathcal{S}| \leqslant |\mathcal{S}'| \leqslant |\mathcal{S}| + \frac{1}{3}KJ$ to the second sum.

By Proposition 10.8, the innermost sum in (73) satisfies

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}} \\ \widetilde{\boldsymbol{d}} = \boldsymbol{d}'}} \prod_{\substack{p|\rho_{\boldsymbol{d}} \\ p \nmid \rho_{\boldsymbol{d}'}}} \frac{1}{p} \ll e^{O(KJ)} V^{(K-R)J/2}.$$

We can split the sum over non-backtracking walks $\boldsymbol{d}'$ as a sum over predictable walks, and a sum over unpredictable walks:

$$\sum_{\boldsymbol{d}' \in \widetilde{\mathbf{D}}_R^{\mathcal{S}',\mathcal{L}'}} \prod_{p|\rho_{\boldsymbol{d}'}} \frac{1}{p} = \sum_{\boldsymbol{d}' \in \mathbf{P}_R} \prod_{p|\rho_{\boldsymbol{d}'}} \frac{1}{p} + \sum_{\boldsymbol{d}' \in \mathbf{U}_R} \prod_{p|\rho_{\boldsymbol{d}'}} \frac{1}{p}.$$

The first and second sums on the right-hand side are $\ll e^{O(KJ)} V^{|\mathcal{S}'| + (|\mathcal{L}'| + |\mathcal{U}'|)/2}$ and $\ll 1$ respectively, by Proposition 8.8 and Proposition 9.1.

Putting everything together, we obtain that (73) is

$$\leqslant e^{O(KJ)} \sum_{0 \leqslant R \leqslant K} \sum_{\substack{[\![R]\!] \times [\![J]\!] = \mathcal{S}' \sqcup \mathcal{L}' \sqcup \mathcal{U}' \\ |\mathcal{U}'| \leqslant K^{2\varepsilon_1} \\ |\mathcal{S}| \leqslant |\mathcal{S}'| \leqslant |\mathcal{S}| + KJ/3}} V^{(K-R)J/2} V^{|\mathcal{S}'| + (|\mathcal{L}'| + |\mathcal{U}'|)/2}.$$

Note that

$$\begin{aligned}
\tfrac{1}{2}(K-R)J + |\mathcal{S}'| + \tfrac{1}{2}(|\mathcal{L}'| + |\mathcal{U}'|) &= \tfrac{1}{2}(K-R)J + \tfrac{1}{2}|\mathcal{S}'| + \tfrac{1}{2}RJ \\
&= \tfrac{1}{2}KJ + \tfrac{1}{2}|\mathcal{S}'| \\
&\leqslant \tfrac{2}{3}KJ + \tfrac{1}{2}|\mathcal{S}|,
\end{aligned}$$

using $|\mathcal{S}'| \leqslant |\mathcal{S}| + KJ/3$ for the last inequality. Since there are $\leqslant e^{O(KJ)}$ choices for R, $\mathcal{S}'$, $\mathcal{L}'$ and $\mathcal{U}'$, we exactly get (72). This finishes the proof of Proposition 3.5. $\qquad \square$

11. Walks with many divisibility conditions

In this section, we prove Lemma 9.6 on systems of triangular constraints, Lemma 6.9 on bad unlit indices, Lemma 5.5 on the size of $I_N \backslash Y_L$, and Lemma 7.6 on the cut-off function for the combinatorial sieve. All of these were stated without proof in the previous sections.

11.1. Proof of the triangular system bound.

We start this section by proving the bound on the weighted number of solutions to triangular systems of constraints, which we restate here for convenience.

Lemma 9.6. *Let $1 \leqslant T \leqslant R \leqslant 2K$. Let $B \geqslant 1$. Let $\mathbf{T} \subset (\pm\mathcal{D})^R$ be a set such that each $\boldsymbol{d} \in \mathbf{T}$ satisfies a triangular system of complexity $(T; 3, B)$ (thus, the system may depend on $\boldsymbol{d}$). Then*

$$\sum_{\boldsymbol{d} \in \mathbf{T}} \prod_{p|\rho_{\boldsymbol{d}}} \frac{1}{p} \ll B K^{11RJ} H_0^{-T/2}.$$

The proof is very heavy in notations, but the idea is just to fix the shape of the system and use the fact that it is triangular to take advantage of the constraints one by one.

Proof of Lemma 9.6. Let $\sigma \in \{\pm 1\}^R$ be a sequence of signs and let Π be a partition of $[\![R]\!] \times [\![J]\!]$. For $t \in [\![T]\!]$, let $I_t \subset [\![R]\!]$ be a union of at most three discrete intervals and let $(i_t, j_t) \in [\![R]\!] \times [\![J]\!]$. Let $\kappa \in [\![-B, B]\!]$. Let $f : [\![T]\!] \to \Pi$. We define $\mathbf{T}_{(\sigma,\Pi,(I_t),(i_t),(j_t),\kappa,f)}$ to be the set of all $\boldsymbol{d} \in \mathbf{T}$ such that

- $\mathrm{sign}(d_i) = \sigma_i$ for $i \in [\![R]\!]$;
- for all $(i,j),(i',j') \in [\![R]\!] \times [\![J]\!]$, $d_{ij} = d_{i'j'}$ iff (i,j) and (i',j') are in the same class in Π;[12]
- the constraints $(C_t(\boldsymbol{d}))_{t \in [\![T]\!]}$ are satisfied by $\boldsymbol{d}$, where $C_t(\boldsymbol{d})$ is short for $C_{I_t,i_t,j_t,\kappa}(\boldsymbol{d})$;
- for $t \in [\![T]\!]$, the prime[13] $d_{f(t)}$ is involved in $C_t(\boldsymbol{d})$ but absent from $C_s(\boldsymbol{d})$ for $s < t$.

We will show that, for each such choice of $\sigma, \Pi, (I_t), (i_t), (j_t), \kappa, f$, we have

$$(74) \qquad \sum_{\boldsymbol{d} \in \mathbf{T}_{(\sigma,\Pi,(I_t),(i_t),(j_t),\kappa,f)}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \leqslant V^{RJ} H_0^{-T/2}.$$

This is enough to prove Lemma 9.6. Indeed, $\mathbf{T}$ is contained in the union of $\mathbf{T}_{(\sigma,\Pi,(I_t),(i_t),(j_t),\kappa,f)}$ over all possible choices of $\sigma, \Pi, (I_t), (i_t), (j_t), \kappa$ and f. Hence, to bound the sum over $\boldsymbol{d} \in \mathbf{T}$, it suffices to multiply the right-hand side of (74) by the number of possibilities for these parameters. There are 2^R choices for σ. The number of partitions of $[\![R]\!] \times [\![J]\!]$ is $\leqslant (RJ)^{RJ}$. For $t \in [\![T]\!]$, since I_t is a union of at most three discrete intervals, it is uniquely determined by six elements of $[\![R]\!]$. Thus, the number of choices for $(I_t, i_t, j_t)_{t \in [\![T]\!]}$ is $\leqslant (R^6 RJ)^T$. There are $\leqslant 2B + 1$ choices for $\kappa \in [\![-B, B]\!]$. Any function $f : [\![T]\!] \to \Pi$ induces a function $[\![T]\!] \to [\![R]\!] \times [\![J]\!]$ which uniquely determines f, so there are $\leqslant (RJ)^T$ possibilities for f. Therefore, assuming (74), we have

$$\sum_{\boldsymbol{d} \in \mathbf{T}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \leqslant 3B\, 2^R (RJ)^{RJ+8T} V^{RJ} H_0^{-T/2}.$$

By property (b) of Lemma 2.4 and the inequality $(a/n)^n \leqslant e^a$, we have $V^J \leqslant K$. Using $T \leqslant R \leqslant 2K$ and $J \ll \log \log H$, we can simplify the above to obtain

$$\sum_{\boldsymbol{d} \in \mathbf{T}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll B K^{11RJ} H_0^{-T/2}$$

as desired.

It remains to prove (74). Let $\sigma, \Pi, (I_t), (i_t), (j_t), \kappa$ and f be such that the set $\mathbf{T}_{(\sigma,\Pi,(I_t),(i_t),(j_t),\kappa,f)}$ (which will henceforth be denoted by $\mathbf{T}_*$) is non-empty. Note that every class α of Π is contained in $[\![R]\!] \times \{j(\alpha)\}$ for some $j(\alpha) \in [\![J]\!]$, which is the unique integer such that $d_\alpha \in \mathcal{P}_{j(\alpha)}$ for all $\boldsymbol{d} \in \mathbf{T}_*$.

Any $\boldsymbol{d} \in \mathbf{T}_*$ is uniquely determined by the sequence of primes $(d_\alpha)_{\alpha \in \Pi}$.

Let $\Pi_0 := \Pi \setminus \{f(t) : t \in [\![T]\!]\}$ and, for $t \in [\![T]\!]$, let $\Pi_t := \Pi_{t-1} \cup \{f(t)\}$.

Let W_0 be the set of all sequences $(p_\alpha)_{\alpha \in \Pi_0}$ with $p_\alpha \in \mathcal{P}_{j(\alpha)}$ for all $\alpha \in \Pi_0$. For any $t \in [\![T]\!]$ and any sequence of primes $(p_\alpha)_{\alpha \in \Pi_{t-1}}$, we define $W_t\big[(p_\alpha)_{\alpha \in \Pi_{t-1}}\big]$ to be the set of all primes $p \in \mathcal{P}_{j(f(t))}$ for which there is some $\boldsymbol{d} \in \mathbf{T}_*$ such that $d_\alpha = p_\alpha$ for all $\alpha \in \Pi_{t-1}$ and $d_{f(t)} = p$.

Then, we have

$$(75) \qquad \sum_{\boldsymbol{d} \in \mathbf{T}_*} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} = \sum_{(p_\alpha) \in W_0} \left(\prod_{\alpha \in \Pi_0} \frac{1}{p_\alpha} \right) \sum_{p_{f(1)} \in W_1} \frac{1}{p_{f(1)}} \sum_{p_{f(2)} \in W_2} \frac{1}{p_{f(2)}} \cdots \sum_{p_{f(T)} \in W_T} \frac{1}{p_{f(T)}},$$

---

[12]Recall from Definition 9.2, that for $(i,j) \in [\![R]\!] \times [\![J]\!]$, write d_{ij} for the unique prime in $\mathcal{P}_j$ dividing d_i.

[13]For $\alpha \in \Pi$, we write d_α for the prime d_{ij}, where (i,j) is any element of α; this is well-defined by construction.

writing W_t instead of $W_t\big[(p_\alpha)_{\alpha\in\Pi_{t-1}}\big]$ to shorten notation.

Fix some $t \in [\![T]\!]$ and some sequence $(p_\alpha)_{\alpha\in\Pi_{t-1}}$. We claim that

$$(76) \qquad\qquad \sum_{p\in W_t} \frac{1}{p} \leqslant H_0^{-1/2}.$$

Recall that, for any $p \in W_t$, there is some $\boldsymbol{d} \in \mathbf{T}_*$ with $d_\alpha = p_\alpha$ for all $\alpha \in \Pi_{t-1}$ and $d_{f(t)} = p$. In particular, p is involved in $C_t(\boldsymbol{d})$. By Definition 9.4, this means that p is (i)-involved, (ii)-involved or (iii)-involved in $C_t(\boldsymbol{d})$.

If p is (i)-involved in $C_t(\boldsymbol{d})$, then by definition $\sum_{i\in I_t} d_i = 0$ and $\sum_{i\in I_t,\, p|d_i} d_i \neq 0$. This means that p satisfies the linear equation $Ap + B = 0$ where

$$A := \frac{1}{p}\sum_{\substack{i\in I_t \\ p|d_i}} d_i \qquad \text{and} \qquad B := \sum_{\substack{i\in I_t \\ p\nmid d_i}} d_i.$$

Observe that A and B are explicit expressions of the primes $(p_\alpha)_{\alpha\in\Pi_{t-1}}$. Indeed, $f(t)$ is of the form $f(t) = Z_t \times \{j(f(t))\}$ for some $Z_t \subset [\![R]\!]$, and we may rewrite

$$A = \sum_{i\in I_t\cap Z_t}\prod_{j\in[\![J]\!]\setminus j(f(t))} d_{ij} \qquad \text{and} \qquad B = \sum_{i\in I_t\setminus Z_t}\prod_{j\in[\![J]\!]} d_{ij}.$$

By definition of $f(t)$, the prime $p = d_{f(t)}$ does not appear in A or B. By construction, the primes $d_{f(t+1)}, .., d_{f(T)}$ are absent from $C_t(\boldsymbol{d})$, which means that $d_{f(t+1)}, ..., d_{f(T)}$ cannot be any of the primes d_{ij} occurring in A or B either. Hence, A and B are fully determined by the primes $(p_\alpha)_{\alpha\in\Pi_{t-1}}$. Since $A \neq 0$ by assumption, the equation $Ap + B = 0$ has at most one solution p in $\mathcal{P}_{j(f(t))}$.

If p is (ii)-involved in $C_t(\boldsymbol{d})$, we know that p must be a prime divisor of

$$A := \sum_{\substack{i\in I_t \\ p\nmid d_i}} d_i,$$

and that $A \neq 0$. As before, A is can be explicitly computed from the primes $(p_\alpha)_{\alpha\in\Pi_{t-1}}$. Note that A is non-zero by assumption, and $|A| \leqslant RH$, so A has at most $\log_2(RK)$ prime factors.

Finally, if p is (iii)-involved in $C_t(\boldsymbol{d})$, we have $d_{i_t j_t} \mid Ap + B + \kappa$ with A and B as in case (i), but this time we assume that A is not divisible by $d_{i_t j_t}$. Once again, $d_{i_t j_t}$, A and B only depend on the primes $(p_\alpha)_{\alpha\in\Pi_{t-1}}$, and κ is fixed. Thus, this divisibility condition uniquely determines the congruence class of p modulo the prime $d_{i_t j_t}$. Using that $\mathcal{P} \subset (H_0, H)$, we have, for any x,

$$\sum_{\substack{p\in\mathcal{P} \\ p\equiv x\;(\mathrm{mod}\;d_{i_t j_t})}} \frac{1}{p} \leqslant \sum_{\substack{1\leqslant n\leqslant H \\ n\equiv 1\;(\mathrm{mod}\;H_0)}} \frac{1}{n} \leqslant \frac{10\log H}{H_0}.$$

Gathering the three cases, we conclude that

$$\sum_{p\in W_t} \frac{1}{p} \leqslant \frac{1}{H_0} + \frac{\log_2(RH)}{H_0} + \frac{10\log H}{H_0} \leqslant H_0^{-1/2},$$

so (76) is proved. Using this fact in (75) successively for $t = T, T-1, \ldots, 1$ yields

$$\sum_{\boldsymbol{d}\in\mathbf{T}_*}\prod_{p|\rho_{\boldsymbol{d}}}\frac{1}{p} \leqslant H_0^{-T/2}\sum_{(p_\alpha)\in W_0}\prod_{\alpha\in\Pi_0}\frac{1}{p_\alpha} = H_0^{-T/2}\prod_{\alpha\in\Pi_0}V_{j(\alpha)} \leqslant H_0^{-T/2}V^{|\Pi_0|}.$$

Equation (74) follows, which concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

11.2. **Bad single indices.** In this section, we prove Lemma 7.4 by extracting a large triangular system from the bad single indices conditions.

Lemma 11.1. *Let $\mathcal{S} \subset [\![K]\!] \times [\![J]\!]$ and let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$ be such that $|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}$. Then $\boldsymbol{d}$ satisfies a triangular system of complexity $(\lfloor \frac{1}{6}K^{1/2}/J \rfloor; 1, 0)$.*

Proof. We split the proof according to which case of Definition 7.2 occurs most often.

Suppose first that there are $\geqslant \frac{1}{3}K^{1/2}$ indices $(i,j) \in \mathcal{S}$ for which there exists $(i',j') \in \mathcal{S}$ with $b_i = b_{i'}$ and $i \neq i'$. By symmetry, there are $\geqslant \frac{1}{6}K^{1/2}$ indices $(i,j) \in \mathcal{S}$ for which there exists $(i',j') \in \mathcal{S}$ with $b_i = b_{i'}$ and $i < i'$. We use the pigeonhole principle on the second coordinate j. We see that, for some $j_0 \in [\![J]\!]$, there is a set $\mathcal{I}$ of $\geqslant \frac{1}{6}K^{1/2}/J$ elements $i \in [\![K]\!]$ with the above properties, i.e. $(i,j_0) \in \mathcal{S}$ and there exists $(i',j') \in \mathcal{S}$ with $b_i = b_{i'}$ and $i < i'$. In particular, for any $i \in \mathcal{I}$, there is some $i' > i$ such that

$$d_{K1} \ \Big| \ 0 = b_{i'} - b_i = \sum_{i \leqslant k < i'} d_k,$$

meaning that $\boldsymbol{d}$ satisfies the constraint $C_{[\![i,i'-1]\!],K,1,0}(\boldsymbol{d})$. For any $i \in \mathcal{I}$, we choose such an i' (arbitrarily) and denote by C_i the resulting constraint $C_{[\![i,i'-1]\!],K,1,0}(\boldsymbol{d})$. Note that the prime d_{ij_0} is (i)-involved in C_i, since we have $\sum_{i \leqslant k < i', d_{ij_0} | d_k} d_k = d_i \neq 0$ as $(i,j_0) \in \mathcal{S}$. Moreover, for any $i_1, i_2 \in \mathcal{I}$ with $i_1 < i_2$, the prime $d_{i_1 j_0}$ is absent from C_{i_2}. Therefore, the sequence $(C_i)_{i \in \mathcal{I}}$ (in decreasing order of $i \in \mathcal{I}$) forms a triangular system of constraints satisfied by $\boldsymbol{d}$, of complexity $(\lfloor \frac{1}{6}K^{1/2}/J \rfloor; 1, 0)$.

Case (2) of Definition 7.2 is treated in an analogous way. Suppose there are $\geqslant \frac{1}{3}K^{1/2}$ indices $(i,j) \in \mathcal{S}$ for which there exists $(i',j') \in \mathcal{S}$ with $b_{i+1} = b_{i'+1}$ and $i \neq i'$. As before, we can find some $j_0 \in [\![J]\!]$ and some set $\mathcal{I}$ of size $\geqslant \frac{1}{6}K/J$ such that, for all $i \in \mathcal{I}$, $(i,j_0) \in \mathcal{S}$ and there exists $(i',j') \in \mathcal{S}$ with $b_{i+1} = b_{i'+1}$ and $i' < i$. For $i \in \mathcal{I}$, define C_i to be the constraint $C_{[\![i'+1,i]\!],1,1,0}(\boldsymbol{d})$, for some $(i',j') \in \mathcal{S}$ with these properties. Then, for all $i \in \mathcal{I}$, d_{ij_0} is (i)-involved in C_i. In addition, for all $i_1, i_2 \in \mathcal{I}$ with $i_1 < i_2$, the prime $d_{i_2 j_0}$ is absent from C_{i_1}. Thus $(C_i)_{i \in \mathcal{I}}$ (in increasing order of $i \in \mathcal{I}$) forms a triangular system of constraints satisfied by $\boldsymbol{d}$, of complexity $(\lfloor \frac{1}{6}K^{1/2}/J \rfloor; 1, 0)$.

Finally, we split case (3) of Definition 7.2 into two sub-cases, according to whether $i' < i$ or $i' > i$. Suppose that there are $\geqslant \frac{1}{6}K^{1/2}$ indices $(i,j) \in \mathcal{S}$ for which there exists $1 \leqslant i' < i$ such that $d_{ij} \mid b_{i'} - b_i$ and $b_{i'} \notin \{b_i, b_{i+1}\}$. By the pigeonhole principle, there is some $j_0 \in [\![J]\!]$ and some $\mathcal{I} \subset [\![K]\!]$ of size $\geqslant \frac{1}{6}K^{1/2}/J$ with the following properties. For all $i \in \mathcal{I}$, we have $(i,j_0) \in \mathcal{S}$ and there exists $1 \leqslant i' < i$ with $d_{ij_0} \mid b_{i'} - b_i$ and $b_{i'} \notin \{b_i, b_{i+1}\}$. Thus, for every $i \in \mathcal{I}$, there is some $i' < i$ such that $\boldsymbol{d}$ satisfies the constraint

$$d_{ij_0} \ \Big| \ b_i - b_{i'} = \sum_{i' \leqslant k < i} d_k \neq 0.$$

For every $i \in \mathcal{I}$, we choose an appropriate i' and denote by C_i the constraint $C_{[\![i',i-1]\!],i,j_0,0}(\boldsymbol{d})$. The prime d_{ij_0} is (ii)-involved in C_i, as $(i,j_0) \in \mathcal{S}$. For $i_1, i_2 \in \mathcal{I}$ with $i_1 < i_2$, observe that $d_{i_2 j_0}$ is absent from C_{i_1} (this again follows from the fact that $(i_2, j_0) \in \mathcal{S}$). Thus $(C_i)_{i \in \mathcal{I}}$ (in increasing order of $i \in \mathcal{I}$) forms a triangular system of constraints satisfied by $\boldsymbol{d}$, of complexity $(\lfloor \frac{1}{6}K^{1/2}/J \rfloor; 1, 0)$.

The remaining sub-case is when there are $\geqslant \frac{1}{6}K^{1/2}$ indices $(i,j) \in \mathcal{S}$ for which there exists $i < i' \leqslant K$ such that $d_{ij} \mid b_{i'} - b_i$ and $b_{i'} \notin \{b_i, b_{i+1}\}$. The proof is identical to the previous paragraph.

Since $|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}$, at least one of the previous cases must occur, and in each of them the conclusion of the lemma holds. $\qquad\square$

We now restate and prove Lemma 7.4.

Lemma 7.4. *Let $\mathcal{S} \subset [\![K]\!] \times [\![J]\!]$. We have*

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}} \\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

Proof of Lemma 7.4. Let $T := \left\lfloor \frac{1}{6} K^{1/2}/J \right\rfloor$. By Lemma 11.1, we know that every $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$ with $|\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}$ satisfies a triangular system of complexity $(T; 1, 0)$. By Lemma 9.6, we deduce that

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}} \\ |\mathcal{S}_{\mathrm{bad}}(\boldsymbol{d})| > K^{1/2}}} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll K^{11KJ} H_0^{-T/2},$$

which is $\ll 1$ since $J \leqslant \log K$, $\log H_0 \gg K^{1-\varepsilon_1}$ and $T \gg K^{1/3}$. $\qquad\square$

11.3. **Primitive prohibited sequences.**

In this section, we prove a technical lemma that allows us to find constraints and involved primes in primitive prohibited sequences. This will allow us to immediately deduce Lemma 5.5, and will be useful for the proof of Lemma 7.6.

The divisibility condition in the definition of prohibited sequences (see Definition 5.2) only brings up a subset of the prime factors of the d_i. Even the primes that do appear in that constraint might not be involved in the sense of Definition 9.4. Lemma 11.2 is a useful tool to circumvent this problem: it allows us to pass from an arbitrary prime to a (possibly different) involved prime.

Lemma 11.2. *Let $2 < \ell \leqslant L$ and let $\boldsymbol{d} = (d_1, \ldots, d_\ell)$ be a primitive prohibited sequence.*

Let Γ be the set of all constraints C satisfied by $\boldsymbol{d}$, that are of the form $C = C_{I, i_0, j_0, 0}(\boldsymbol{d})$ for some discrete interval $I \subset [\![\ell]\!]$ and some $(i_0, j_0) \in [\![\ell]\!] \times [\![J]\!]$.

For every prime $p \mid \rho_{\boldsymbol{d}}$,

(1) either there is a constraint $C \in \Gamma$ in which p is involved,

(2) or there is another prime q involved in a constraint of Γ, such that $q \mid d_{i'}$ for some $i' \in [\![\ell]\!]$ and

$$\sum_{\substack{1 \leqslant i < i' \\ p \mid d_i}} d_i \not\equiv 0 \pmod{q}.$$

Proof. For $p \in \mathcal{P}$, let

$$I(p) := \{i \in [\![\ell]\!] : p \mid d_i\},$$

it is a discrete interval by definition of a prohibited sequence (Definition 5.2).

By Definition 5.2, there are some $1 < \ell_0 < \ell$ and $j_0 \in [\![J]\!]$ such that $d_{1 j_0} \nmid d_\ell$ and

$$(77) \qquad d_{1 j_0} \mid \sum_{\ell_0 \leqslant i \leqslant \ell} d_i.$$

In particular, the constraint $C_1 := C_{[\![\ell_0, \ell]\!], 1, j_0, 0}(\boldsymbol{d})$ is satisfied by $\boldsymbol{d}$, so $C_1 \in \Gamma$.

Among all the primes dividing d_ℓ, choose some prime p_1 such that $I(p_1)$ is minimal for inclusion. We claim that p_1 is (iii)-involved in C_1. Clearly $d_{1 j_0} \neq p_1$ since $d_{1 j_0} \nmid d_\ell$. Moreover, it is easy to see that $d_i = d_\ell$ for all $i \in I(p_1)$, using the fact that $I(p_1)$ is minimal for inclusion and the first two assumptions of Definition 5.2, as in the proof of Lemma 9.18. Hence,

$$(78) \qquad \sum_{\substack{i \in [\![\ell_0, \ell]\!] \\ p_1 \mid d_i}} d_i = |I(p_1) \cap [\![\ell_0, \ell]\!]| \, d_\ell.$$

Since $|I(p_1) \cap [\![\ell_0, \ell]\!]| \leqslant L < H_0 < d_{1j_0}$ and $d_{1j_0} \nmid d_\ell$, the expression (78) is not divisible by d_{1j_0}, which means that p_1 is (iii)-involved in C_1 as claimed.

We are now ready to start the proof of Lemma 11.2 in earnest. Let $p \mid \rho_{\boldsymbol{d}}$ be a prime.

If $I(p) \subset I(p_1)$, then $I(p) = I(p_1)$ by minimality of $I(p_1)$. Repeating the previous paragraph with p in place of p_1, we conclude that p is involved in C_1, so we are in case (1). We henceforth assume that $I(p) \setminus I(p_1)$ is non-empty. Note that $I(p) \setminus I(p_1)$ is discrete interval; we denote it by $[\![a_1, a_2]\!]$.

Assume that $\sum_{i \in [\![\ell]\!], \, p \mid d_i} d_i \equiv 0 \pmod{p_1}$, as otherwise we are in case (2) with $q = p_1$. This can be rewritten as

$$(79) \qquad p_1 \ \Big| \ \sum_{i \in [\![a_1, a_2]\!]} d_i.$$

However, this implies that $\left(d_\ell, d_{\ell-1}, \ldots, d_{a_1+1}, d_{a_1}\right)$ is a prohibited sequence. Since $\boldsymbol{d}$ is a primitive prohibited sequence, this is only possible if $a_1 = 1$. Hence, $I(p) \setminus I(p_1) = [\![1, a_2]\!]$. Note that $p_1 \nmid d_1$ and thus $a_2 \geqslant 2$ by (79).

We will now exhibit another prime p_2 for which the case (2) of the lemma holds with $q = p_2$.

Let p_2 be the prime d_{2j_0}. Note that $p_2 \nmid d_1$, or else we would have $p_2 = d_{1j_0}$, and thus $p_2 \mid \sum_{\ell_0 \leqslant i \leqslant \ell} d_i$ by (77). This would imply that $(d_2, d_3, \ldots, d_\ell)$ is a prohibited sequence, which is impossible since $\boldsymbol{d}$ is primitive.

Next, observe that (79) is exactly saying that $\boldsymbol{d}$ satisfies the constraint $C_2 := C_{[\![1, a_2]\!], \ell, j_1, 0}(\boldsymbol{d})$, where j_1 is the unique integer such that $p_1 = d_{\ell j_1}$. Let us show that p_2 is (iii)-involved in C_2. Recall that $[\![1, a_2]\!] = I(p) \setminus I(p_1) \supset \{1, 2\}$, so $p_1 \nmid d_2$ and hence $p_2 \neq p_1$. Suppose for contradiction that p_2 is not (iii)-involved in C_2. Then, we would have

$$(80) \qquad p_1 \ \Big| \ \sum_{\substack{i \in [\![1, a_2]\!] \\ p_2 \mid d_i}} d_i = \sum_{\substack{i \in [\![2, a_2]\!] \\ p_2 \mid d_i}} d_i,$$

using that $p_2 \nmid d_1$. Note that $\{i \in [\![2, a_2]\!] : p_2 \mid d_i\}$ is a discrete interval containing 2 and not containing ℓ. Thus, (80) implies that $\left(d_\ell, d_{\ell-1}, \ldots, d_2\right)$ is a prohibited sequence, contradicting that $\boldsymbol{d}$ is primitive. Hence, p_2 is (iii)-involved in C_2.

To summarise, we have shown that the prime p_2 is involved in $C_2 \in \Gamma$. Since $p_2 \mid d_2$, we can easily check that case (2) applies with $q = p_2$ and $i' = 2$: $p_2 \mid d_2$ and

$$\sum_{\substack{1 \leqslant i < 2 \\ p \mid d_i}} d_i = d_1 \not\equiv 0 \pmod{p_2}.$$

This concludes the proof. $\qquad \square$

We can use the previous lemma (in fact, a much weaker version would suffice) to prove Lemma 5.5.

Lemma 5.5. $|I_N \setminus Y_L| \ll H_0^{-1/3} N$.

Proof of Lemma 5.5. Recall that $\mathbb{Z} \setminus Y_L$ is the union of all prohibited progressions $P \in \mathcal{Y}$. By the union bound, we have

$$|I_N \setminus Y_L| \leqslant \sum_{P \in \mathcal{Y}} |I_N \cap P| \ll N \sum_{P \in \mathcal{Y}} \frac{1}{q_P}.$$

For any $P \in \mathcal{Y}$, there is a primitive prohibited sequence $\boldsymbol{d}$ of length $\leqslant L$ such that P is the prohibited progression associated to $\boldsymbol{d}$. By Lemma 11.2, there is a constraint C satisfied by $\boldsymbol{d}$ which involves at least one prime. This constraint alone can be viewed as a triangular system of complexity $(1; 1, 0)$.

We apply Lemma 9.6 with $\mathbf{T} = \mathbf{T}_\ell$ being the set of $\boldsymbol{d} \in (\pm\mathcal{D})^\ell$ satisfying a triangular system of complexity $(1;1,0)$, for $2 < \ell \leqslant L$. This gives

$$\sum_{P \in \mathcal{Y}} \frac{1}{q_P} \leqslant \sum_{2 < \ell \leqslant L} \sum_{\boldsymbol{d} \in \mathbf{T}_\ell} \prod_{p \mid \rho_{\boldsymbol{d}}} \frac{1}{p} \ll LK^{11LJ} H_0^{-1/2}.$$

This is $\ll H_0^{-1/3}$ since $J \leqslant \log K$, $\log H_0 \gg K^{1-\varepsilon_1}$ and $L \ll K^{1-10\varepsilon_1}$. $\square$

11.4. **Cut-off function for the combinatorial sieve.** We finally turn to the proof of Lemma 7.6. Recall that $L := K^{1-10\varepsilon_1}$, $\mathcal{Y}$ is the set of all prohibited arithmetic progressions (see Definition 5.3) and $(\mathcal{Y} - \boldsymbol{b})^\cap$ is the set defined in Notation 7.5.

In the next definition, we introduce the function $\text{rank}_{\boldsymbol{d}} : (\mathcal{Y} - \boldsymbol{b})^\cap \to \mathbb{Z}^{\geqslant 0} \cup \{+\infty\}$ which is used as a cut-off for the combinatorial sieve (or rather, a family of such functions, one for every $\boldsymbol{d}$).

Definition 11.3. Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ and let $\boldsymbol{d} \in \mathbf{D}_K^\mathcal{S}$. We define $A_{\boldsymbol{d}} \subset \mathbb{Z}$ to be the arithmetic progression

$$A_{\boldsymbol{d}} := \{n \in \mathbb{Z} : \forall (i,j) \in \mathcal{L}, \ d_{ij} \mid n + b_i\}.$$

Let $R \in (\mathcal{Y} - \boldsymbol{b})^\cap$. If $R \cap A_{\boldsymbol{d}} = \emptyset$, we set $\text{rank}_{\boldsymbol{d}}(R) := +\infty$. Otherwise, we define $\text{rank}_{\boldsymbol{d}}(R)$ to be the largest integer $T \geqslant 0$ for which there exist progressions $Q_1, \ldots, Q_T \in \mathcal{Y} - \boldsymbol{b}$ containing R such that, for each $t \in [\![T]\!]$, the modulus q_{Q_t} does not divide $q_{A_{\boldsymbol{d}}} \prod_{s \in [\![T]\!] \setminus \{t\}} q_{Q_s}$.

We need to show that these rank functions satisfy the five properties of Lemma 7.6. We will be able to quickly derive the first few properties from the following simple fact.

Lemma 11.4. *Let $R \in (\mathcal{Y} - \boldsymbol{b})^\cap$ be such that $\text{rank}_{\boldsymbol{d}}(R) = T < +\infty$. Then, there are progressions $Q_1, \ldots, Q_T \in \mathcal{Y} - \boldsymbol{b}$ such that*

$$\emptyset \neq R \cap A_{\boldsymbol{d}} = \bigcap_{t \in [\![T]\!]} Q_t \cap A_{\boldsymbol{d}}.$$

Proof. By definition of $(\mathcal{Y} - \boldsymbol{b})^\cap$, we may write $R = \bigcap_{i \in I} Q_i$ for some finite set I and some $Q_i \in \mathcal{Y} - \boldsymbol{b}$. Let I_0 be a minimal subset of I such that

$$(81) \qquad\qquad\qquad\qquad R \cap A_{\boldsymbol{d}} = \bigcap_{i \in I_0} Q_i \cap A_{\boldsymbol{d}}.$$

Note that the modulus of a non-empty intersection of arithmetic progressions is the least common multiple of the moduli of these progressions. There is no $i_0 \in I_0$ such that $q_{Q_{i_0}}$ divides $q_{A_{\boldsymbol{d}}} \prod_{i \in I_0 \setminus \{i_0\}} q_{Q_i}$, for otherwise $\bigcap_{i \in I_0 \setminus \{i_0\}} Q_i \cap A_{\boldsymbol{d}}$ and $\bigcap_{i \in I_0} Q_i \cap A_{\boldsymbol{d}}$ would have the same modulus, so these progressions would be equal, contradicting the minimality of I_0. This shows that $|I_0| \leqslant \text{rank}_{\boldsymbol{d}}(R)$. Thus, (81) means that we have been able to write $R \cap A_{\boldsymbol{d}}$ as an intersection of at most $\text{rank}_{\boldsymbol{d}}(R)$ progressions $Q_i \cap A_{\boldsymbol{d}}$. Repeating some Q_i if necessary, we can make it an intersection of exactly $\text{rank}_{\boldsymbol{d}}(R)$ sets. $\square$

We reproduce Lemma 7.6 here for convenience.

Lemma 7.6. *Let $\mathcal{S}, \mathcal{L}, \mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$. For every $\boldsymbol{d} \in \mathbf{D}_K^\mathcal{S}$, there exists a function*

$$\text{rank}_{\boldsymbol{d}} : (\mathcal{Y} - \boldsymbol{b})^\cap \to \mathbb{Z}^{\geqslant 0} \cup \{+\infty\}$$

satisfying the following properties.

Define the arithmetic progression $A_{\boldsymbol{d}} := \{n \in \mathbb{Z} : \forall (i,j) \in \mathcal{L}, \ d_{ij} \mid n + b_i\}$.

Let $X_{\boldsymbol{d}}$ be the set of all $R \in (\mathcal{Y} - \boldsymbol{b})^\cap$ such that $\text{rank}_{\boldsymbol{d}}(R) < K^{5\varepsilon_1}$. Let $\partial X_{\boldsymbol{d}}$ be the set of all $R \in (\mathcal{Y} - \boldsymbol{b})^\cap \setminus X_{\boldsymbol{d}}$ of the form $R = R' \cap P$ for some $R' \in X_{\boldsymbol{d}}$ and $P \in \mathcal{Y} - \boldsymbol{b}$.

(1) *(Primes dividing the modulus)* For every $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$,
$$\omega(q_R) \leqslant LJ \operatorname{rank}_{\boldsymbol{d}}(R) + KJ.$$

(2) *(Primes $p \mid \rho_{\boldsymbol{d};\mathcal{S}}$ dividing the modulus)* For every $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$,
$$|\{p : p \mid q_R, \, p \mid \rho_{\boldsymbol{d};\mathcal{S}}\}| \leqslant LJ \operatorname{rank}_{\boldsymbol{d}}(R).$$

(3) *(Combinatorial sieve)* Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$. For all $n \in \mathbb{Z}$, we have

$$\mathbf{1}_{\forall i, \, n+b_i \in Y_L \text{ and } n \in A_{\boldsymbol{d}}} = \sum_{R \in X_{\boldsymbol{d}}} c_{R,\boldsymbol{d}} \mathbf{1}_{n \in R \cap A_{\boldsymbol{d}}} + O\left(3^{3KJ} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \mathbf{1}_{n \in R \cap A_{\boldsymbol{d}}} \right),$$

where the coefficients $c_{R,\boldsymbol{d}}$ are independent of n and satisfy $|c_{R,\boldsymbol{d}}| \leqslant 2^{2KJ}$.

(4) *(Main term bound)* We have
$$\sum_{R \in X_{\boldsymbol{d}}} \prod_{\substack{p \mid q_R \\ p \nmid \rho_{\boldsymbol{d}}}} \frac{1}{p} \ll e^{O(KJ)}.$$

(5) *(Remainder term bound)* Suppose $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$. Then
$$\sum_{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \prod_{p \mid q_R \rho_{\boldsymbol{d}}} \frac{1}{p} \ll 1.$$

Proof of parts (1) *and* (2) *of Lemma 7.6.* Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$ and let $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ be a progression with $\operatorname{rank}_{\boldsymbol{d}}(R) = T < +\infty$. By Lemma 11.4, there are $Q_1, \ldots, Q_T \in \mathcal{Y} - \boldsymbol{b}$ such that
$$\emptyset \neq R \cap A_{\boldsymbol{d}} = \bigcap_{t \in [\![T]\!]} Q_t \cap A_{\boldsymbol{d}}.$$

Property (1) follows, since
$$\omega(q_R) \leqslant \omega(q_{R \cap A_{\boldsymbol{d}}}) \leqslant \sum_{t \in [\![T]\!]} \omega(q_{Q_t}) + \omega(q_{A_{\boldsymbol{d}}}) \leqslant TLJ + KJ.$$

For property (2), write $\omega_{\mathcal{S}}(n) := \sum_{p \mid \rho_{\boldsymbol{d};\mathcal{S}}} \mathbf{1}_{p \mid n}$. We similarly obtain
$$\omega_{\mathcal{S}}(q_R) \leqslant \omega_{\mathcal{S}}(q_{R \cap A_{\boldsymbol{d}}}) \leqslant \sum_{t \in [\![T]\!]} \omega_{\mathcal{S}}(q_{Q_t}) + \omega_{\mathcal{S}}(q_{A_{\boldsymbol{d}}}) \leqslant TLJ + 0$$

as $q_{A_{\boldsymbol{d}}}$ is only divisible by the primes $p \mid \rho_{\boldsymbol{d};\mathcal{L}}$. $\qquad\square$

For part (3) of Lemma 7.6, namely the combinatorial sieve, we just need to use Proposition A.3, checking that the hypotheses are satisfied.

Proof of part (3) *of Lemma 7.6.* We use Proposition A.3 with the initial set of arithmetic progressions being $\mathcal{Y} - \boldsymbol{b}$, and with $X = X_{\boldsymbol{d}}$ being the set of all $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ such that $\operatorname{rank}_{\boldsymbol{d}}(R) < K^{5\varepsilon_1}$. Note that $X_{\boldsymbol{d}} \neq \emptyset$ as $\mathbb{Z} \in X_{\boldsymbol{d}}$.

For any $R, R' \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ with $R \subset R'$, it is clear from Definition 11.3 that $\operatorname{rank}_{\boldsymbol{d}}(R') \leqslant \operatorname{rank}_{\boldsymbol{d}}(R)$. Therefore $X_{\boldsymbol{d}}$ is closed under containment. Furthermore, $\omega(q_R) \leqslant 2KJ$ for all $R \in X_{\boldsymbol{d}}$, by property (1) of Lemma 7.6, as $L = K^{1-10\varepsilon_1}$. For elements $R \in \partial X_{\boldsymbol{d}} \setminus \{\emptyset\}$, we have $\omega(q_R) \leqslant 3KJ$, as any $P \in \mathcal{Y} - \boldsymbol{b}$ has $\omega(q_P) \leqslant LJ \leqslant KJ$ by definition of a prohibited progression. The conclusion follows from Proposition A.3, observing that '$n \notin P$ for all $P \in \mathcal{Y} - \boldsymbol{b}$' is equivalent to '$n + b_i \in Y_L$ for all $i \in [\![K]\!]$'. $\qquad\square$

To prove parts (4) and (5) of Lemma 7.6, we will use the following technical lemma.

Lemma 11.5. *Let $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ be sets such that $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ and let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}}$. Let $T = \lceil K^{5\varepsilon_1} \rceil$. Let $\mathcal{Y}_T$ be a set whose elements are progressions $R \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ for which there exist $Q_1, \ldots, Q_T \in \mathcal{Y} - \boldsymbol{b}$ such that*

$$\text{(82)} \qquad \emptyset \neq R \cap A_{\boldsymbol{d}} = \bigcap_{t \in [\![T]\!]} Q_t \cap A_{\boldsymbol{d}}.$$

Let $\mathcal{P}' \subset \mathcal{P}$ be a set of size $\leqslant 2KJ$ containing the prime divisors of $q_{A_{\boldsymbol{d}}}$. Then

$$\sum_{R \in \mathcal{Y}_T} \prod_{\substack{p | q_R \\ p \notin \mathcal{P}'}} \frac{1}{p} \ll e^{O(KJ)}.$$

Proof of part (4) of Lemma 7.6. This immediately follows from Lemma 11.5, choosing $\mathcal{Y}_T = X_{\boldsymbol{d}}$ and $\mathcal{P}' = \{p : p \mid \rho_{\boldsymbol{d}}\}$. Note that this choice of $\mathcal{Y}_T$ satisfies the required property by definition of $X_{\boldsymbol{d}}$ and Lemma 11.4. $\qquad \square$

Proof of Lemma 11.5. Let $\mathcal{X}$ be the set of all $R_1 \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ of the form $R_1 = R \cap A_{\boldsymbol{d}}$ for some $R \in \mathcal{Y}_T$. Since the prime factors of $q_{A_{\boldsymbol{d}}}$ are in $\mathcal{P}'$, we may rewrite

$$\sum_{R \in \mathcal{Y}_T} \prod_{\substack{p | q_R \\ p \notin \mathcal{P}'}} \frac{1}{p} = \sum_{R_1 \in \mathcal{X}} \prod_{\substack{p | q_{R_1} \\ p \notin \mathcal{P}'}} \frac{1}{p} \sum_{\substack{R \in \mathcal{Y}_T \\ R \cap A_{\boldsymbol{d}} = R_1}} 1.$$

To bound the inner sum, we use the following fact: for any arithmetic progression R_1 and any $d \mid q_{R_1}$, there is a unique arithmetic progression $R \supset R_1$ with $q_{R_1}/q_R = d$, and moreover all progressions $R \supset R_1$ are obtained in this way. Therefore, the inner sum is bounded by the number of divisors of q_{R_1}. For every $R_1 \in \mathcal{X}$, we have $\omega(q_{R_1}) \ll KJ$. This follows from (82) as in the proof of part (1) of Lemma 7.6. Therefore, q_{R_1} has $e^{O(KJ)}$ divisors, and hence the inner sum is $e^{O(KJ)}$.

It remains to show that

$$\text{(83)} \qquad \sum_{R_1 \in \mathcal{X}} \prod_{\substack{p | q_{R_1} \\ p \notin \mathcal{P}'}} \frac{1}{p} \ll e^{O(KJ)}.$$

This is a simple counting problem, similar to Lemma 6.4 or Proposition 8.8. However, the notation is much heavier in this case.

Let $R_1 \in \mathcal{X}$. By definition of $\mathcal{Y}_T$ and $\mathcal{X}$, we can write

$$\text{(84)} \qquad R_1 = \bigcap_{t \in [\![T]\!]} \left(Q_t - b_{k_t}\right) \cap A_{\boldsymbol{d}}$$

for some $Q_t \in \mathcal{Y}$ and some $k_t \in [\![K]\!]$. For $t \in [\![T]\!]$, let $\boldsymbol{d}^{(t)}$ be a primitive prohibited sequence having Q_t as its associated prohibited progression. Let ℓ_t be the length of $\boldsymbol{d}^{(t)}$ and let $\sigma_t \in \{\pm 1\}^{\ell_t}$ be the sequence of signs of the coordinates of $\boldsymbol{d}^{(t)}$. As usual, for $(i, j) \in [\![\ell_t]\!] \times [\![J]\!]$ we write $d_{ij}^{(t)}$ for the unique prime in $\mathcal{P}_j$ dividing $d_i^{(t)}$. Let $\sim$ be the equivalence relation on $\bigsqcup_{t \in [\![T]\!]} \left(\{t\} \times [\![\ell_t]\!] \times [\![J]\!]\right)$ defined by

$$(t_1, i_1, j_1) \sim (t_2, i_2, j_2) \qquad \Longleftrightarrow \qquad d_{i_1 j_1}^{(t_1)} = d_{i_2 j_2}^{(t_2)}.$$

If α is an equivalence class for $\sim$, we write p_α for the prime $d_{ij}^{(t)}$, where (t, i, j) is any element of α. This definition does not depend on the choice of representative, by definition of $\sim$. Let E be the set of all equivalence classes α for $\sim$ such that $p_\alpha \in \mathcal{P}'$. Let $\phi : E \to \mathcal{P}'$ be the map defined by

$\phi(\alpha) = p_\alpha$. We call the tuple $((k_t)_{t\in[\![T]\!]}, (\ell_t)_{t\in[\![T]\!]}, (\sigma_t)_{t\in[\![T]\!]}, \sim, E, \phi)$ a *template* for R_1. Thus, to every progression $R_1 \in \mathcal{X}$ we may associate a template (note that there may not be a canonical choice for the template associated to R_1, as it depends on the choice of a representation of R_1 as in (84)).

Let Θ be the set of all tuples $((k_t)_{t\in[\![T]\!]}, (\ell_t)_{t\in[\![T]\!]}, (\sigma_t)_{t\in[\![T]\!]}, \sim, E, \phi)$ which are a template of some element $R_1 \in \mathcal{X}$. Fix some $\theta = ((k_t)_{t\in[\![T]\!]}, (\ell_t)_{t\in[\![T]\!]}, (\sigma_t)_{t\in[\![T]\!]}, \sim, E, \phi) \in \Theta$. Let $\mathcal{X}_\theta$ be the set of all $R_1 \in \mathcal{X}$ for which θ is a template. Suppose that $\mathcal{X}_\theta$ is non-empty. Any $R_1 \in \mathcal{X}_\theta$ is uniquely determined by the sequence of primes $(p_\alpha)_{\alpha\in E'}$, where E' is the set of all equivalence classes of $\sim$ not in E. Thus

$$\sum_{\substack{R_1\in\mathcal{X}_\theta}} \prod_{\substack{p\mid q_{R_1} \\ p\notin\mathcal{P}'}} \frac{1}{p} \leqslant \sum_{(p_\alpha)_{\alpha\in E'}} \prod_{\alpha\in E'} \frac{1}{p_\alpha} \leqslant V^{|E'|} \leqslant V^{TLJ} \leqslant e^{O(KJ)}, \tag{85}$$

where we used that $T \ll K^{5\varepsilon_1}$ and $L < K^{1-10\varepsilon_1}$ in the last inequality.

We proceed to sum (85) over all choices of $\theta \in \Theta$. We will be done provided that the number of possible templates is $e^{O(KJ)}$. The number of choices for $(k_t)_{t\in[\![T]\!]}$, $(\ell_t)_{t\in[\![T]\!]}$ and $(\sigma_t)_{t\in[\![T]\!]}$ is at most K^T, L^T and $(2^L)^T$ respectively. Since $\sim$ is an equivalence relation on a set of size $\leqslant TLJ$, there are $\leqslant (TLJ)^{TLJ}$ choices for $\sim$. There are $\leqslant 2^{TLJ}$ choices for E. Finally, ϕ is a map from a set of size $\leqslant TLJ$ to a set of size $\leqslant 2KJ$, so there are $\leqslant (2KJ)^{TLJ}$ possibilities for ϕ. In summary, the number of templates is

$$\leqslant K^T \cdot L^T \cdot (2^L)^T \cdot (TLJ)^{TLJ} \cdot 2^{TLJ} \cdot (2KJ)^{TLJ} = e^{O(KJ)}.$$

This concludes the proof of Lemma 11.5. $\qquad\square$

Before turning to part (5) of Lemma 7.6, we first prove an intermediate substructure result, related to collections of primitive prohibited sequences.

Lemma 11.6. *Let* $\mathcal{S}$, $\mathcal{L}$, $\mathcal{U}$ *be sets such that* $[\![K]\!] \times [\![J]\!] = \mathcal{S} \sqcup \mathcal{L} \sqcup \mathcal{U}$ *and* $|\mathcal{U}| \leqslant K^{2\varepsilon_1}$, *and let* $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S},\mathcal{L}}$. *Let* $T \geqslant 8K^{2\varepsilon_1}$ *and let* $k_1 \leqslant k_2 \leqslant \cdots \leqslant k_T$ *be elements of* $[\![K]\!]$. *Let* $\boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(T)}$ *be primitive prohibited sequences, and let* $Q_1, \ldots, Q_T \in \mathcal{Y}$ *be the associated prohibited progressions. Suppose that, for each* $t \in [\![T]\!]$, *the modulus* q_{Q_t} *does not divide* $q_{A_{\boldsymbol{d}}} \prod_{s\in[\![T]\!]\setminus\{t\}} q_{Q_s}$, *and that*

$$\bigcap_{t\in[\![T]\!]} (Q_t - b_{k_t}) \cap A_{\boldsymbol{d}} \neq \emptyset.$$

Let $\boldsymbol{v}$ *be the sequence obtained by concatenation of* $\boldsymbol{d}, \boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(T)}, -\boldsymbol{d}^{(1)}, \ldots, -\boldsymbol{d}^{(T)}$. *Then,* $\boldsymbol{v}$ *satisfies a triangular system of complexity* $(\lfloor \frac{1}{16}T\rfloor; 3, 0)$.

Proof. Fix some $n \in \bigcap_{t=1}^T (Q_t - b_{k_t}) \cap A_{\boldsymbol{d}}$.

For every $t \in [\![T]\!]$, let Γ_t be the set of all constraints satisfied by $\boldsymbol{d}^{(t)}$ of the form $C_{I,i_0,j_0,0}(\boldsymbol{d})$ for some discrete interval $I \subset [\![\ell_t]\!]$ and some $(i_0, j_0) \in [\![\ell_t]\!] \times [\![J]\!]$, where ℓ_t is the length of $\boldsymbol{d}^{(t)}$.

Suppose first that there is a set $\mathcal{I} \subset [\![T]\!]$ of size $\geqslant \frac{1}{16}T$ such that, for every $t \in \mathcal{I}$, there is a constraint $C_t \in \Gamma_t$ and a prime p_t which is involved in C_t and does not divide $\prod_{s<t} q_{Q_s}$. Then, clearly, p_t is absent from C_s, for every $s \in \mathcal{I}$ with $s < t$, which means that the constraints $(C_t)_{t\in\mathcal{I}}$ form a triangular system of complexity $(\lfloor \frac{1}{16}T\rfloor; 1, 0)$. The same conclusion holds if there is a set $\mathcal{I} \subset [\![T]\!]$ of size $\geqslant \frac{1}{16}T$ such that, for every $t \in \mathcal{I}$, there is a constraint $C_t \in \Gamma_t$ and a prime p_t which is involved in C_t and does not divide $\prod_{s>t} q_{Q_s}$. We may thus assume that, for $\geqslant \frac{7}{8}T$ values of $t \in [\![T]\!]$, every prime involved in some constraint of Γ_t divides both $\prod_{s<t} q_{Q_s}$ and $\prod_{s>t} q_{Q_s}$.

For every $t \in \llbracket T \rrbracket$, fix a prime p_t dividing q_{Q_t} but not dividing $q_{A_{\boldsymbol{d}}} \prod_{s \in \llbracket T \rrbracket \setminus \{t\}} q_{Q_s}$. This is possible by the assumption in the statement. We apply Lemma 11.2 with this prime p_t. Note that the first case of Lemma 11.2 can only occur for $< \frac{1}{8}T$ values of $t \in \llbracket T \rrbracket$ by definition of p_t and the previous paragraph. Let $\mathcal{I}_1$ be the set of $t \in \llbracket T \rrbracket$ such that the second case holds, i.e. for $t \in \mathcal{I}_1$ there is a prime q_t involved in a constraint $C_t \in \Gamma_t$ such that

$$(86) \qquad \sum_{\substack{1 \leqslant i < i_t \\ p_t \mid d_i^{(t)}}} d_i^{(t)} \not\equiv 0 \pmod{q_t},$$

where $i_t \in \llbracket \ell_t \rrbracket$ is such that $q_t \mid d_{i_t}^{(t)}$. Thus $|\mathcal{I}_1| \geqslant \frac{7}{8}T$.

By our earlier observation, there is a subset $\mathcal{I}_2 \subset \mathcal{I}_1$ of size $\geqslant \frac{1}{2}T$ such that, for all $t \in \mathcal{I}_2$, there are $1 \leqslant s_1(t) < t < s_2(t) \leqslant T$ with $q_t \mid q_{Q_{s_1(t)}}$ and $q_t \mid q_{Q_{s_2(t)}}$.

By definition of p_t, we know that $p_t \nmid q_{A_{\boldsymbol{d}}}$, i.e. $p_t \nmid \rho_{\boldsymbol{d};\mathcal{L}}$. In other words, p_t does not appear in $\boldsymbol{d}$ at a lit index. Moreover, there are at most $\frac{1}{8}T$ values of $t \in \mathcal{I}_2$ such that $p_t \mid \rho_{\boldsymbol{d};\mathcal{U}}$, since $|\mathcal{U}| \leqslant K^{2\varepsilon_1} \leqslant \frac{1}{8}T$ and all p_t are distinct. We may thus find a subset $\mathcal{I}_3 \subset \mathcal{I}_2$ of size $\geqslant \frac{1}{4}T$ such that $p_t \nmid \rho_{\boldsymbol{d};\mathcal{L} \sqcup \mathcal{U}}$ for every $t \in \mathcal{I}_3$.

Let $\mathcal{I}_4$ be the set of all $t \in \mathcal{I}_3$ such that $p_t \nmid \rho_{\boldsymbol{d};\llbracket 1, k_t - 1 \rrbracket \times \llbracket J \rrbracket}$. Let $\mathcal{I}_5$ be the set of all $t \in \mathcal{I}_3$ such that $p_t \nmid \rho_{\boldsymbol{d};\llbracket k_t, K \rrbracket \times \llbracket J \rrbracket}$. By definition of $\mathcal{I}_3$, we know that for every $t \in \mathcal{I}_3$ there is at most one index $(i, j) \in \llbracket K \rrbracket \times \llbracket J \rrbracket$ (a single index) such that $p_t = d_{ij}$. In particular, $\mathcal{I}_4 \cup \mathcal{I}_5 = \mathcal{I}_3$, so one of $\mathcal{I}_4$ and $\mathcal{I}_5$ has size $\geqslant \frac{1}{8}T$. We will only treat the case where $|\mathcal{I}_4| \geqslant \frac{1}{8}T$; the proof for the case $|\mathcal{I}_5| \geqslant \frac{1}{8}T$ is the same up to symmetry.

Let $t \in \mathcal{I}_4$. Since $n \in R$, we have $n + b_{k_t} \in Q_t$ and thus, by definition of Q_t being the prohibited progression associated to $\boldsymbol{d}^{(t)}$,

$$q_t \ \bigg| \ n + b_{k_t} + \sum_{1 \leqslant i < i_t} d_i^{(t)},$$

with i_t as defined earlier. Since $q_t \mid q_{Q_{s_1(t)}}$, the same reasoning shows that

$$q_t \ \bigg| \ n + b_{k_{s_1(t)}} + \sum_{1 \leqslant i < i_t'} d_i^{(s_1(t))},$$

where $i_t' \in \llbracket \ell_{s_1(t)} \rrbracket$ is such that $q_t \mid d_{i_t'}^{(s_1(t))}$. Subtracting the two divisibility relations, we obtain

$$q_t \ \bigg| \ \sum_{k_{s_1(t)} \leqslant i < k_t} d_i + \sum_{1 \leqslant i < i_t} d_i^{(t)} - \sum_{1 \leqslant i < i_t'} d_i^{(s_1(t))}.$$

This is now a genuine constraint on $\boldsymbol{v}$, which we call C_t. By (86), and since $p_t \nmid q_{Q_{s_1(t)}}$ (by definition of p_t) and $p_t \nmid \rho_{\boldsymbol{d};\llbracket 1, k_t - 1 \rrbracket \times \llbracket J \rrbracket}$ (by definition of $\mathcal{I}_4$), we see that p_t is (iii)-involved in this constraint C_t. In addition, for $t_1, t_2 \in \mathcal{I}_4$ with $t_1 < t_2$, the prime p_{t_2} is absent from C_{t_1} since none of $\rho_{\boldsymbol{d};\llbracket 1, k_{t_2} - 1 \rrbracket \times \llbracket J \rrbracket}$, $q_{Q_{t_1}}$ and $q_{Q_{s_1(t_1)}}$ are divisible by p_{t_2}. Therefore, the family $(C_t)_{t \in \mathcal{I}_4}$ forms a triangular system of complexity $(\lfloor \frac{1}{8}T \rfloor; 3, 0)$.

The case $|\mathcal{I}_5| \geqslant \frac{1}{8}T$ is analogous, where this time $s_2(t)$ takes the role of $s_1(t)$. $\square$

Using this technical Lemma 11.6, we can finally prove part (5) of Lemma 7.6.

Proof of part (5) of Lemma 7.6. Let $T := \lceil K^{5\varepsilon_1} \rceil$. Let $\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}, \mathcal{L}}$ and let $R \in \partial X_{\boldsymbol{d}}$ such that $R \cap A_{\boldsymbol{d}} \neq \emptyset$. Observe that R satisfies $T \leqslant \mathrm{rank}_{\boldsymbol{d}}(R) < +\infty$ by definition of $\partial X_{\boldsymbol{d}}$. Thus, by Definition 11.3, we can find progressions $Q_1, \ldots, Q_T \in \mathcal{Y} - \boldsymbol{b}$ containing R such that, for each

$t \in \llbracket T \rrbracket$, the modulus q_{Q_t} does not divide $q_{A_{\boldsymbol{d}}} \prod_{s \in \llbracket T \rrbracket \setminus \{t\}} q_{Q_s}$. We will first sum over all possibilities for $R_1 := \bigcap_{t \in \llbracket T \rrbracket} Q_t \cap A_{\boldsymbol{d}}$.

Let $\mathcal{X}$ be the set of all $R_1 \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ which are of the form $R_1 = \bigcap_{t \in \llbracket T \rrbracket} Q_t \cap A_{\boldsymbol{d}} \neq \emptyset$ for some $Q_1, \ldots, Q_T \in \mathcal{Y} - \boldsymbol{b}$ with the property that, for all $t \in \llbracket T \rrbracket$, q_{Q_t} does not divide $q_{A_{\boldsymbol{d}}} \prod_{s \in \llbracket T \rrbracket \setminus \{t\}} q_{Q_s}$. We have

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}, \mathcal{L}}}} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \prod_{p | q_R \rho_{\boldsymbol{d}}} \frac{1}{p} \leqslant \sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}, \mathcal{L}}}} \sum_{R_1 \in \mathcal{X}} \prod_{p | q_{R_1} \rho_{\boldsymbol{d}}} \frac{1}{p} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \prod_{\substack{p | q_R \\ p \nmid q_{R_1} \rho_{\boldsymbol{d}}}} \frac{1}{p}.$$

For the innermost sum, we apply Lemma 11.5 with the choices $\mathcal{Y}_T = \{R \in \partial X_{\boldsymbol{d}} : R \cap A_{\boldsymbol{d}} \neq \emptyset\}$ and $\mathcal{P}' = \{p : p \mid q_{R_1} \rho_{\boldsymbol{d}}\}$. The assumptions on $\mathcal{P}'$ are satisfied since $\mathcal{P}'$ contains the prime divisors of $q_{A_{\boldsymbol{d}}}$ and

$$\left| \mathcal{P}' \right| \leqslant KJ + \left| \{p : p \mid q_{R_1}, p \nmid q_{A_{\boldsymbol{d}}}\} \right| \leqslant KJ + TLJ \leqslant 2KJ$$

(recalling that $L \leqslant K^{1-10\varepsilon_1}$ and $T \leqslant K^{5\varepsilon_1} + 1$). We also need to check that $\mathcal{Y}_T$ satisfies the assumption in Lemma 11.5. Let $R \in \mathcal{Y}_T$. By definition of $\partial X_{\boldsymbol{d}}$, we know that $R = R' \cap P$ for some $P \in \mathcal{Y} - \boldsymbol{b}$ and $R' \in (\mathcal{Y} - \boldsymbol{b})^{\cap}$ with $\mathrm{rank}_{\boldsymbol{d}}(R') < K^{5\varepsilon_1}$. Thus $\mathrm{rank}_{\boldsymbol{d}}(R') \leqslant T - 1$ and by Lemma 11.4, there are $P_1, \ldots, P_{T-1} \in \mathcal{Y} - \boldsymbol{b}$ such that

$$R' \cap A_{\boldsymbol{d}} = \bigcap_{t \in \llbracket T-1 \rrbracket} P_t \cap A_{\boldsymbol{d}}.$$

Therefore,

$$\emptyset \neq R \cap A_{\boldsymbol{d}} = P \cap \bigcap_{t \in \llbracket T-1 \rrbracket} P_t \cap A_{\boldsymbol{d}},$$

which is what we wanted to show. By Lemma 11.5, we obtain that

$$\sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \prod_{\substack{p | q_R \\ p \nmid q_{R_1} \rho_{\boldsymbol{d}}}} \frac{1}{p} = e^{O(KJ)}.$$

It remains to bound the sum

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}, \mathcal{L}}}} \sum_{R_1 \in \mathcal{X}} \prod_{p | q_{R_1} \rho_{\boldsymbol{d}}} \frac{1}{p}.$$

For every non-decreasing sequence $(k_t)_{t \in \llbracket T \rrbracket}$ of elements of K, let $\mathbf{T}_{(k_t)}$ be the set of all pairs $(\boldsymbol{d}, R_1) \in \mathbf{D}_K^{\mathcal{S}, \mathcal{L}} \times \mathcal{X}$ such that

$$\emptyset \neq R_1 \cap A_{\boldsymbol{d}} = \bigcap_{t \in \llbracket T \rrbracket} \left(Q_t - b_{k_t} \right) \cap A_{\boldsymbol{d}}$$

for some prohibited progressions $Q_t \in \mathcal{Y}$ with $q_{Q_t} \nmid q_{A_{\boldsymbol{d}}} \prod_{s \in \llbracket T \rrbracket \setminus \{t\}} q_{Q_s}$ for all t. By Lemma 11.6, for any $(\boldsymbol{d}, R_1) \in \mathbf{T}_{(k_t)}$ and any choice $\boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(T)}$ of prohibited sequences used in the definition of R_1, the concatenation of $\boldsymbol{d}, \boldsymbol{d}^{(1)}, \ldots, \boldsymbol{d}^{(T)}, -\boldsymbol{d}^{(1)}, \ldots, -\boldsymbol{d}^{(T)}$ satisfies a triangular system of complexity $(\lfloor \frac{1}{16} T \rfloor; 3, 0)$. This concatenation has length $\leqslant K + 2TL \leqslant 2K$. By Lemma 9.6, we get

$$\sum_{(\boldsymbol{d}, R_1) \in \mathbf{T}_{(k_t)}} \prod_{p | q_{R_1} \rho_{\boldsymbol{d}}} \frac{1}{p} \ll K^{22KJ} H_0^{-\lfloor T/16 \rfloor / 2}.$$

Summing over all choices for $(k_t) \in \llbracket K \rrbracket^T$ and recalling our bound for the inner sum, we obtain

$$\sum_{\substack{\boldsymbol{d} \in \mathbf{D}_K^{\mathcal{S}, \mathcal{L}}}} \sum_{\substack{R \in \partial X_{\boldsymbol{d}} \\ R \cap A_{\boldsymbol{d}} \neq \emptyset}} \prod_{p | q_R \rho_{\boldsymbol{d}}} \frac{1}{p} \ll e^{O(KJ)} K^{22KJ+T} H_0^{-T/50},$$

which is $\ll 1$ since $J \leqslant \log K$, $\log H_0 \gg K^{1-\varepsilon_1}$ and $T \geqslant K^{5\varepsilon_1}$. $\qquad\square$

Appendix A. Combinatorial sieve for composite moduli

Let $\mathcal{Y}$ be a finite set of arithmetic progressions in $\mathbb{Z}$. By the inclusion-exclusion principle, we can write

$$\mathbf{1}_{n \notin P \; \forall P \in \mathcal{Y}} = 1 - \sum_{P_1 \in \mathcal{Y}} \mathbf{1}_{n \in P_1} + \sum_{\substack{P_1, P_2 \in \mathcal{Y} \\ \text{distinct}}} \mathbf{1}_{n \in P_1 \cap P_2} - \sum_{\substack{P_1, P_2, P_3 \in \mathcal{Y} \\ \text{distinct}}} \mathbf{1}_{n \in P_1 \cap P_2 \cap P_3} + \cdots = \sum_{S \subset \mathcal{Y}} (-1)^{|S|} \mathbf{1}_{n \in \cap S}.$$

For $S = \emptyset$, we used the convention $\cap \emptyset := \mathbb{Z}$. The last sum contains $2^{|\mathcal{Y}|}$ terms. We wish to replace this exact identity with an approximate version having far fewer terms. To do so, we truncate the above sum and restrict S to a smaller collection $\mathcal{X}$ of subsets of $\mathcal{Y}$.

Lemma A.1. *Let $\mathcal{Y}$ be a finite set of arithmetic progressions in $\mathbb{Z}$. Let $\mathcal{X}$ be a non-empty collection of subsets of $\mathcal{Y}$ which is closed under containment, i.e. if $S \in \mathcal{X}$ and $S' \subset S$ then $S' \in \mathcal{X}$.*

(1) If $n \notin P$ for all $P \in \mathcal{Y}$, then

$$\mathbf{1}_{n \notin P \; \forall P \in \mathcal{Y}} = 1 = \sum_{S \in \mathcal{X}} (-1)^{|S|} \mathbf{1}_{n \in \cap S}.$$

(2) If $n \in P_0$ for some progression $P_0 \in \mathcal{Y}$, then

$$(87) \qquad \mathbf{1}_{n \notin P \; \forall P \in \mathcal{Y}} = 0 = \sum_{S \in \mathcal{X}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} + \sum_{\substack{P_0 \in S \subset \mathcal{Y} \\ S \notin \mathcal{X}, \, S \setminus \{P_0\} \in \mathcal{X}}} (-1)^{|S|} \mathbf{1}_{n \in \cap S}.$$

Proof. (1) If n does not belong to any $P \in \mathcal{Y}$, all the terms in the sum are zero except for $S = \emptyset$.

(2) Suppose $n \in P_0 \in \mathcal{Y}$. By inclusion-exclusion, we know that

$$0 = \mathbf{1}_{n \notin P \; \forall P \in \mathcal{Y}} = \sum_{S \subset \mathcal{Y}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} = \left(\sum_{S \in \mathcal{X}} + \sum_{\substack{S \notin \mathcal{X} \\ P_0 \notin S}} + \sum_{\substack{S \notin \mathcal{X} \\ P_0 \in S \\ S \setminus \{P_0\} \in \mathcal{X}}} + \sum_{\substack{S \notin \mathcal{X} \\ P_0 \in S \\ S \setminus \{P_0\} \notin \mathcal{X}}} \right) (-1)^{|S|} \mathbf{1}_{n \in \cap S}.$$

To obtain the conclusion, note that the second and fourth sums on the right-hand side cancel each other out, since

$$\sum_{\substack{P_0 \in S \subset \mathcal{Y} \\ S \notin \mathcal{X} \\ S \setminus \{P_0\} \notin \mathcal{X}}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} = \sum_{\substack{P_0 \notin T \subset \mathcal{Y} \\ T \notin \mathcal{X}}} (-1)^{|T \cup \{P_0\}|} \mathbf{1}_{n \in \cap T} \mathbf{1}_{n \in P_0} = - \sum_{\substack{P_0 \notin S \subset \mathcal{Y} \\ S \notin \mathcal{X}}} (-1)^{|S|} \mathbf{1}_{n \in \cap S},$$

using that $\mathcal{X}$ is closed under containment in the first equality. $\qquad\square$

The next lemma shows some cancellation for combinatorial sums having up to $2^{2^{|\Omega|}}$ terms. The short proof below is due to Helfgott and Radziwiłł [5].

Lemma A.2. *Let $\mathcal{A}$ be any collection of subsets of a finite set Ω. Then*

$$\left| \sum_{\substack{\mathcal{B} \subset \mathcal{A} \\ \cup \mathcal{B} = \Omega}} (-1)^{|\mathcal{B}|} \right| \leqslant 2^{|\Omega|}.$$

Proof. Observe that, given two finite sets $\Omega_1 \subset \Omega_2$, we have

$$(88) \qquad (-1)^{|\Omega_1|} \sum_{\Omega_1 \subset W \subset \Omega_2} (-1)^{|W|} = \mathbf{1}_{\Omega_1 = \Omega_2}.$$

Indeed, this is obvious if $\Omega_1 = \Omega_2$, and if $\Omega_1 \neq \Omega_2$ the left-hand side is the expanded form of $(1-1)^{|\Omega_2 \setminus \Omega_1|}$.

This allows us to write

$$\left| \sum_{\substack{\mathcal{B} \subset \mathcal{A} \\ \cup \mathcal{B} = \Omega}} (-1)^{|\mathcal{B}|} \right| = \left| \sum_{\mathcal{B} \subset \mathcal{A}} (-1)^{|\mathcal{B}|} \sum_{\cup \mathcal{B} \subset W \subset \Omega} (-1)^{|W|} \right| = \left| \sum_{W \subset \Omega} (-1)^{|W|} \sum_{\substack{\mathcal{B} \subset \mathcal{A} \\ \cup \mathcal{B} \subset W}} (-1)^{|\mathcal{B}|} \right|.$$

The inner sum has the shape of (88), with $\Omega_1 = \emptyset$ and $\Omega_2 = \{A \in \mathcal{A} \mid A \subset W\}$, so is at most 1 in absolute value. Since the outer sum has $\leqslant 2^{|\Omega|}$ terms, the claim follows. $\square$

Assuming that the progressions in $\mathcal{Y}$ have square-free moduli, and with an additional hypothesis on the shape of $\mathcal{X}$, we can use Lemma A.2 to show that the two sums in (87) exhibit some cancellation.

Proposition A.3. *Let $\mathcal{Y}$ be a finite set of arithmetic progressions in $\mathbb{Z}$ with square-free moduli. Let*

$$\mathcal{Y}^{\cap} := \{\cap S : S \subset \mathcal{Y}\}.$$

Fix a non-empty subset $X \subset \mathcal{Y}^{\cap}$ that is closed under containment, i.e. if a progression $P \in \mathcal{Y}^{\cap}$ is an element of X, then so are all $P' \in \mathcal{Y}^{\cap}$ with $P' \supset P$. Let $\mathcal{X}$ be the collection of subsets of $\mathcal{Y}$ defined by[14]

$$\mathcal{X} = \{S \subset \mathcal{Y} : \cap S \in X\}.$$

Then

$$(89) \qquad \mathbf{1}_{n \notin P \ \forall P \in \mathcal{Y}} = \sum_{S \in \mathcal{X}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} + O\left(\sum_{R \in \partial X} 3^{\omega(q_R)} \mathbf{1}_{n \in R} \right)$$

where

$$\partial X := \{R \in \mathcal{Y}^{\cap} : R \notin X \text{ and } R = P \cap P' \text{ for some } P \in X \text{ and } P' \in \mathcal{Y}\}.$$

Moreover, the first sum can be rewritten as

$$\sum_{S \in \mathcal{X}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} = \sum_{P \in X} c_P \mathbf{1}_{n \in P}$$

for some coefficients $c_P \in \mathbb{Z}$ satisfying $|c_P| \leqslant 2^{\omega(q_P)}$.

Proof. If the left-hand side of (89) is 1, then the equality (89) is true by Lemma A.1. On the other hand, if the left-hand side is 0, then by Lemma A.1 we have

$$\mathbf{1}_{n \notin P \ \forall P \in \mathcal{Y}} = \sum_{S \in \mathcal{X}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} + \sum_{\substack{P_0 \in S \subset \mathcal{Y} \\ S \notin \mathcal{X}, S \setminus \{P_0\} \in \mathcal{X}}} (-1)^{|S|} \mathbf{1}_{n \in \cap S},$$

where, in the last sum, $P_0 \in \mathcal{Y}$ is an arbitrary progression containing n. We will bound the second sum at the end of this proof.

Let us analyse the first sum. We have

$$\sum_{S \in \mathcal{X}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} = \sum_{P \in X} c_P \mathbf{1}_{n \in P}$$

---

[14]Note that $\emptyset \in \mathcal{X}$ since $\cap \emptyset = \mathbb{Z}$ by convention.

where, for $P \in X$,

$$c_P := \sum_{\substack{S \in \mathcal{X} \\ \cap S = P}} (-1)^{|S|} = \sum_{\substack{S \subset \{P' \in \mathcal{Y}: P' \supset P\} \\ \cap S = P}} (-1)^{|S|}.$$

Fix some $P \in X$. If S is a set of progressions containing P, the condition $\cap S = P$ is equivalent to

$$\operatorname{lcm}\{q_{P'} : P' \in S\} = q_P.$$

Since all progressions in $\mathcal{Y}$ have square-free moduli, this is in turn equivalent to

$$\bigcup_{P' \in S} \{p : p \mid q_{P'}\} = \{p : p \mid q_P\}.$$

Let $\Omega_P := \{p : p \mid q_P\}$,

$$\mathcal{A}_P := \left\{ \{p : p \mid q_{P'}\} : P' \in \mathcal{Y}, P' \supset P \right\}$$

and, for every set S of progressions containing P, let

$$\mathcal{B}_P(S) := \left\{ \{p : p \mid q_{P'}\} : P' \in S \right\}.$$

Note that $\mathcal{B}_P(S)$ determines S, since a progression $P' \in \mathcal{Y}$ with $P' \supset P$ is uniquely determined by its modulus $q_{P'}$, which in turn is uniquely determined by its set of prime factors. Therefore,

$$c_P = \sum_{\substack{S \subset \{P' \in \mathcal{Y}: P' \supset P\} \\ \cap S = P}} (-1)^{|S|} = \sum_{\substack{\mathcal{B} \subset \mathcal{A}_P \\ \cup \mathcal{B} = \Omega_P}} (-1)^{|\mathcal{B}|}.$$

By Lemma A.2, we obtain $|c_P| \leqslant 2^{|\Omega_P|} = 2^{\omega(q_P)}$.

We now turn to the remainder term. We suppose that $n \in P_0$ for some $P_0 \in \mathcal{Y}$. We operate a change of variables and write $S' = S \setminus \{P_0\}$, $P = \cap S'$ and $R = \cap S$. The conditions $S \notin \mathcal{X}$ and $S \setminus \{P_0\} \in \mathcal{X}$ become $R \notin X$ and $P \in X$, respectively. Hence, we have

$$\sum_{\substack{P_0 \in S \subset \mathcal{Y} \\ S \notin \mathcal{X}, S \setminus \{P_0\} \in \mathcal{X}}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} = \sum_{\substack{R \in \mathcal{Y}^\cap \\ R \notin X}} \mathbf{1}_{n \in R} \sum_{\substack{P \in X \\ R = P \cap P_0}} \sum_{\substack{S' \in \mathcal{X} \\ \cap S' = P}} (-1)^{|S'|+1}.$$

The inner sum is exactly $-c_P$, which is $O\left(2^{\omega(q_P)}\right)$. Recalling that, for fixed R, a progression $P \supset R$ is uniquely determined by its modulus q_P, which divides q_R, we have

$$\sum_{\substack{P_0 \in S \subset \mathcal{Y} \\ S \notin \mathcal{X}, S \setminus \{P_0\} \in \mathcal{X}}} (-1)^{|S|} \mathbf{1}_{n \in \cap S} = O\left(\sum_{R \in \partial X} \mathbf{1}_{n \in R} \sum_{d \mid q_R} 2^{\omega(d)} \right).$$

The observation that $\sum_{d \mid m} 2^{\omega(d)} = 3^{\omega(m)}$ for all square-free $m \geqslant 1$ concludes the proof. $\qquad \square$

Appendix B. Sum without divisibility conditions

In this section we prove Proposition 2.6, which quickly follows from the next proposition.

Proposition B.1. *Let ε_1, H, J, H_0, $(\mathcal{P}_i)$ and (V_i) be as in Theorem 2.1. Let $V := \max_i V_i$.*

Let $N \geqslant \exp\left((\log H)^2\right)$ and let $I_N := \mathbb{N} \cap (N, 2N]$.

For $\mathcal{I} \subset [\![J]\!]$, define $\mathcal{D}_{\mathcal{I}}$ to be the set of all products $\prod_{i \in \mathcal{I}} p_i$ with $p_i \in \mathcal{P}_i$ for all i. Then, for all non-empty $\mathcal{I} \subset [\![J]\!]$, we have

$$\sum_{n \in I_N} \sum_{d \in \mathcal{D}_{\mathcal{I}}} \frac{1}{d} \lambda(n) \lambda(n+d) \ll \frac{V^J N}{(\log H)^{1/2000}}.$$

Proof of Proposition 2.6 assuming Proposition B.1. We can expand the difference $S_2 - S_1$ as

$$\sum_{\emptyset \neq \mathcal{I} \subset [\![J]\!]} S(\mathcal{I}),$$

where

$$S(\mathcal{I}) := \sum_{n \in (N, 2N]} \sum_{(p_1, \ldots, p_J) \in \mathcal{P}_1 \times \cdots \times \mathcal{P}_J} \left(\prod_{i \in \mathcal{I}} \frac{1}{p_i} \right) \left(\prod_{i \in [\![J]\!] \setminus \mathcal{I}} \mathbf{1}_{p_i | n} \right) \lambda(n) \lambda(n + p_1 \cdots p_J).$$

Changing variables $n = md$ with $d = \prod_{i \in [\![J]\!] \setminus \mathcal{I}} p_i$ gives

$$S(\mathcal{I}) = \sum_{d \in \mathcal{D}_{[\![J]\!] \setminus \mathcal{I}}} \lambda(d)^2 \sum_{\frac{N}{d} < m \leqslant \frac{2N}{d}} \sum_{d' \in \mathcal{D}_{\mathcal{I}}} \frac{1}{d'} \lambda(m) \lambda(m + d').$$

By Proposition B.1, the double sum over m and d' is

$$\ll \frac{V^J N / d}{(\log H)^{1/2000}}$$

Hence,

$$S(\mathcal{I}) \ll \frac{V^J N}{(\log H)^{1/2000}} \sum_{d \in \mathcal{D}_{[\![J]\!] \setminus \mathcal{I}}} \frac{1}{d} \ll \frac{V^{2J} N}{(\log H)^{1/2000}}$$

for every non-empty $\mathcal{I} \subset [\![J]\!]$. Therefore

$$|S_2 - S_1| \ll 2^J V^{2J} \frac{N}{(\log H)^{1/2000}}.$$

Note that $2^J V^{2J} \ll (\log H)^{\varepsilon_1}$ by Lemma 2.4, so $|S_2 - S_1| \ll \frac{N}{(\log H)^{1/2500}}$ if ε_1 is sufficiently small. $\square$

Lemma B.2. *Fix a non-empty $\mathcal{I} \subset [\![J]\!]$ and let $\mathcal{D}_{\mathcal{I}}$ be as in Proposition B.1. Let $M \in [H_0, H]$. Define*

$$Q(\alpha) := \sum_{\substack{d \in \mathcal{D}_{\mathcal{I}} \\ d \in (M/2, M]}} \frac{e(\alpha d)}{d},$$

where, as usual, $e(x) := \exp(2\pi i x)$. Then,

$$\int_0^1 |Q(\alpha)|^4 \, d\alpha \ll \frac{V^{4J}}{M(\log M)^4}.$$

Proof. By Parseval's identity, we can expand

$$\int_0^1 |Q(\alpha)|^4 d\alpha = \int_0^1 \left| \sum_{|m| \leqslant M} \left(\sum_{\substack{d_1, d_2 \in \mathcal{D}_{\mathcal{I}} \\ d_1, d_2 \in (M/2, M] \\ d_1 - d_2 = m}} \frac{1}{d_1 d_2} \right) e(m\alpha) \right|^2 d\alpha = \sum_{|m| \leqslant M} \left| \sum_{\substack{d_1, d_2 \in \mathcal{D}_{\mathcal{I}} \\ d_1, d_2 \in (M/2, M] \\ d_1 - d_2 = m}} \frac{1}{d_1 d_2} \right|^2.$$

For $m = 0$, the inner sum is trivially $\ll 1/M$.

Fix $m > 0$. Let $N(m, b, \mathcal{I}, M)$ denote the number pairs $(d_1, d_2) \in \mathcal{D}_{\mathcal{I}} \times \mathcal{D}_{\mathcal{I}}$ such that $d_1 - d_2 = m$, $d_1 \in (M/2, M]$ and $\gcd(d_1, d_2, m) = b$. Observe that $N(m, b, \mathcal{I}, M) = 0$ unless $b \mid m$ and $b \in \mathcal{D}_{\mathcal{I}'}$ for some $\mathcal{I}' \subset \mathcal{I}$, in which case we have

$$N(m, b, \mathcal{I}, M) = N\left(\frac{m}{b}, 1, \mathcal{I} \setminus \mathcal{I}', \frac{M}{b} \right).$$

We thus are led to bound the number of coprime solutions $(e_1, e_2) \in \mathcal{D}_{\mathcal{I} \setminus \mathcal{I}'} \times \mathcal{D}_{\mathcal{I} \setminus \mathcal{I}'}$ to the equation $e_1 - e_2 = m/b$ with $e_1 \in (M/2b, M/b]$. Let i_+ be the largest element of $\mathcal{I} \setminus \mathcal{I}'$. We can rewrite

$e_i = n_i p_i$ where $n_i \in \mathcal{D}_{(\mathcal{I} \setminus \mathcal{I}') \setminus \{i_+\}}$ and $p_i \in \mathcal{P}_{i_+}$ for $i \in \{1, 2\}$. For fixed n_1, n_2, the number of solutions $(p_1, p_2) \in \mathcal{P}_{i_+} \times \mathcal{P}_{i_+}$ to the linear equation

$$n_1 p_1 - n_2 p_2 = \frac{m}{b}$$

with $n_1 p_1 \in (M/2b, M/b]$ is

$$\ll \frac{M/b}{\varphi(n_1)\varphi(n_2)(\log M/b)^2} \cdot \frac{m/b}{\varphi(m/b)}$$

by classical sieve theoretic methods, such as [3, Proposition 6.22]. To apply this particular result, we used the fact that $\max(n_1, n_2) \leqslant (M/b)^{1/10}$, which holds by property (c) of Lemma 2.4.

Note that $\varphi(n) \gg n$ if n is a product of $\leqslant J$ primes, each $\geqslant H_0$. This is the case for n_1 and n_2. Therefore,

$$N(m, b, \mathcal{I}, M) \ll \sum_{n_1, n_2 \in \mathcal{D}_{(\mathcal{I} \setminus \mathcal{I}') \setminus \{i_+\}}} \frac{M/b}{n_1 n_2 (\log M/b)^2} \cdot \frac{m/b}{\varphi(m/b)} \ll V^{2J} \frac{M/b}{(\log M/b)^2} \cdot \frac{m}{\varphi(m)}.$$

We conclude that the total number of solutions (d_1, d_2) to $d_1 - d_2 = m$ with $d_1, d_2 \in (M/2, M]$ is

$$\ll \sum_{\substack{b|m \\ b < m}} V^{2J} \frac{M/b}{(\log M/b)^2} \cdot \frac{m}{\varphi(m)} \ll \frac{V^{2J} M}{(\log M)^2} \cdot \frac{\sigma_1(m)}{\varphi(m)}.$$

We thus obtain

$$\int_0^1 |Q(\alpha)|^4 d\alpha \ll \frac{1}{M^2} + \left(\frac{V^{2J}}{M(\log M)^2} \right)^2 \sum_{m=1}^{M} \left(\frac{\sigma_1(m)}{\varphi(m)} \right)^2 \ll \frac{V^{4J}}{M(\log M)^4},$$

where we used the elementary estimate [19, Corollary 3.6] in the last inequality. $\square$

Proof of Proposition B.1. Let $V_{[M]} := \sum_{d \in \mathcal{D}_{\mathcal{I}} \cap (M/2, M]} 1/d$. It suffices to show that

$$(90) \qquad T_M := \sum_{n \in I_N} \sum_{\substack{d \in \mathcal{D}_{\mathcal{I}} \\ d \in (M/2, M]}} \frac{1}{d} \lambda(n) \lambda(n+d) \ll \left(\frac{V^J}{\log M} + V_{[M]} \right) \frac{N}{(\log H)^{1/1750}}$$

holds for all $M \in [H_0, H]$. Indeed, summing this inequality for $M \in \{H 2^{-j} : j \geqslant 0\} \cap [H_0, H]$ gives the desired upper bound

$$\sum_{n \in I_N} \sum_{d \in \mathcal{D}_{\mathcal{I}}} \frac{1}{d} \lambda(n) \lambda(n+d) \ll \left(V^J (\log \log H) + V^J \right) \frac{N}{(\log H)^{1/1750}} \ll \frac{V^J N}{(\log H)^{1/2000}}.$$

To prove (90), we start by introducing a new average over shifts $m \leqslant M$ and use the circle method:

$$T_M = \frac{1}{M} \sum_{m \leqslant M} \sum_{n \in I_N} \sum_{\substack{d \in \mathcal{D}_{\mathcal{I}} \\ d \in (M/2, M]}} \frac{1}{d} \lambda(n+m) \lambda(n+m+d) + O(MV^J)$$

$$= \frac{1}{M} \sum_{n \in I_N} \int_0^1 Q(\alpha) F_n(\alpha) G_n(\alpha) d\alpha + O(MV^J),$$

with $F_n(\alpha) := \sum_{m \leqslant M} \lambda(n+m) e(\alpha m)$, $G_n(\alpha) := \sum_{k \leqslant 2M} \lambda(n+k) e(-\alpha k)$ and $Q(\alpha)$ as in Lemma B.2. The error term $O(MV^J)$ is clearly negligible.

Let $\varepsilon > 0$ be a parameter that will be fixed later, and let $E_\varepsilon := \{\alpha \in [0,1] : |Q(\alpha)| > \varepsilon\}$. Outside of E_ε, the function $|Q|$ is small and we have

$$\sum_{n \in I_N} \int_{[0,1] \setminus E_\varepsilon} |Q||F_n||G_n| \leqslant \varepsilon N \, \|F_n\|_2 \, \|G_n\|_2 \ll \varepsilon M N.$$

On the other hand, the Lebesgue measure of E_ε is $\ll V^{4J}/(\varepsilon^4 M (\log M)^4)$ by Lemma B.2 and Markov's inequality. Hence

$$\sum_{n \in I_N} \int_{E_\varepsilon} |Q||F_n||G_n| \ll \frac{V^{4J} \|Q\|_\infty}{\varepsilon^4 M (\log M)^4} \left\| \sum_{n \in I_N} |F_n||G_n| \right\|_\infty \ll \frac{V^{4J} V_{[M]}}{\varepsilon^4 M (\log M)^4} M \left\| \sum_{n \in I_N} |F_n| \right\|_\infty.$$

We now make crucial use of [9, Theorem 1.3] to obtain

$$\left\| \sum_{n \in I_N} |F_n| \right\|_\infty = \sup_{\alpha \in \mathbb{R}} \sum_{n \in I_N} \left| \sum_{n \leqslant n' \leqslant n+M} \lambda(n') e(\alpha n') \right| \ll \left((\log M)^{-1/2} + (\log N)^{-1/700} \right) M N.$$

Since $M \geqslant H_0$ and $\log N \geqslant (\log H)^2$, this upper bound is $\ll (\log H)^{-c'}$ where $c' = 1/350$.

Putting everything together, we conclude that

$$T_M \ll \left(\varepsilon + \frac{V^{4J} V_{[M]}}{\varepsilon^4 (\log M)^4 (\log H)^{c'}} \right) N.$$

We choose $\varepsilon = V^J (\log M)^{-1} (\log H)^{-c'/5}$ and deduce the claimed bound (90). $\qquad\square$

Appendix C. Smooth cut-off

Lemma C.1. *There exists a C^∞ function $W : \mathbb{R} \to [0,1]$ such that*

- $W(x) = 1$ *for* $x \in \left[\frac{1}{2} JV, \frac{3}{2} JV\right]$;
- $W(x) = 0$ *for* $x \notin [0, 2JV]$;
- *(Bound a-th derivative of m-th power) For any integers $a \geqslant 1$ and $m \geqslant 1$,*

$$\left\| (W^m)^{(a)} \right\|_\infty \leqslant 2^m \left(\frac{Ca}{JV} \right)^a,$$

  where C is an absolute constant.

Proof. The first step is to bound the derivatives of the test function $\varphi(x) := \mathbf{1}_{[-1,1]}(x) f(x)$, where

$$f(z) := \exp\left(\frac{2}{z^2 - 1} \right).$$

This can be done using Cauchy's inequality for holomorphic functions.

For $0 < x < 1$, we choose the radius $R(x) = (1-x)/2$. Note that $\frac{2}{z^2 - 1} = \frac{1}{z-1} - \frac{1}{z+1}$. For any $z \in \mathbb{C}$ with $|z - x| = R(x)$, we have

$$|f(z)| = \exp\left(\mathrm{Re}\left(\frac{1}{z-1} \right) \right) \exp\left(\mathrm{Re}\left(\frac{-1}{z+1} \right) \right) \ll \exp\left(\frac{1}{x - R(x) - 1} \right) = \exp\left(\frac{-1}{3R(x)} \right).$$

In particular, for any integer $a \geqslant 1$ we have $|f(z)| \ll a!(3R(x))^a$. Cauchy's inequality then gives

$$\left| f^{(a)}(x) \right| \ll \frac{a!}{R(x)^a} a! (3R(x))^a \ll (O(a))^a.$$

Therefore $\left\| \varphi^{(a)} \right\|_\infty \leqslant (O(a))^a$.

Let $T := JV$. We now define W as the convolution
$$W(x) := \tfrac{4}{T}\varphi\left(\tfrac{4}{T}x\right) * \mathbf{1}_{\left[\frac{T}{4}, \frac{7T}{4}\right]}(x).$$

Using $\frac{d}{dx}(F * G) = \left(\frac{d}{dx}F\right) * G$ and $\|F * G\|_\infty \leqslant \|F\|_\infty \|G\|_1$ we get the bound
$$\left\|W^{(a)}\right\|_\infty \leqslant \left(\frac{Ca}{T}\right)^a$$
for the derivatives of W, where $C > 0$ is an absolute constant.

For powers of W, we use the generalised Leibniz rule to get
$$\left\|(W^m)^{(a)}\right\|_\infty \leqslant \sum_{\substack{b_1 + \ldots + b_m = a \\ b_i \in \mathbb{Z}^{\geqslant 0}}} \frac{a!}{b_1! \cdots b_m!} \prod_{i=1}^m \left\|W^{(b_i)}\right\|_\infty.$$

This sum has $\binom{a+m-1}{m-1} \leqslant 2^{a+m}$ terms, and each of them is
$$\leqslant \frac{a!}{b_1! \cdots b_m!} \prod_{i=1}^m \left(\frac{Cb_i}{T}\right)^{b_i} \leqslant \left(\frac{O(a)}{T}\right)^a.$$

The inequality follows, and the other properties of W are clear. $\qquad\square$

Appendix D. Probabilistic model for the integers

This section is devoted to Lemma 6.2, which replaces the integer n with a random variable $\mathbf{n}$, in the spirit of Kubilius' work on probabilistic number theory [7]. The proof uses standard sieve techniques.

Lemma D.1 (Fundamental Lemma of sieve theory). *Let $z, D, \kappa > 0$. Let P be a set of primes $p \leqslant z$. Let (a_n) be a sequence of non-negative real numbers. Suppose that, for every square-free $d \leqslant D$ all of whose prime factors are in P, we have*
$$\sum_{d|n} a_n = g(d)M + R_d$$
with g a non-negative multiplicative function such that, for all $2 \leqslant w \leqslant z$,
$$\prod_{\substack{w \leqslant p < z \\ p \in P}} (1 - g(p))^{-1} \leqslant A\left(\frac{\log z}{\log w}\right)^\kappa,$$
where $A > 1$ is a constant. Let $s = \log D / \log z$, and assume $9\kappa - s < -1$. Then
$$\sum_{p \nmid n\, \forall p \in P} a_n = \prod_{p \in P}(1 - g(p)) \cdot M \cdot \left(1 + O(e^{9\kappa - s}A^{10})\right) + O\left(\sum_{\substack{d \leqslant D \\ p | d \Rightarrow p \in P}} \mu(d)^2 |R_d|\right),$$
where the implied constants are $\leqslant 1$ in absolute value.

Proof. This is [3, Theorem 6.9]. $\qquad\square$

Corollary D.2. *Let $z, \kappa > 0$. Let P be a set of primes $p \leqslant z$. Let (a_n) be a of sequence of non-negative real numbers.*

Suppose that, for every square-free d all of whose prime factors are in P, we have
$$\sum_{d|n} a_n = \frac{\rho(d)}{d}M + R_d,$$

where

- ρ *is a non-negative multiplicative function;*
- $|R_d| \ll \rho(d)$ *for all square-free d all of whose prime factors are in P;*
- $\rho(p) \leqslant \min(p-1, \kappa)$ *for every $p \in P$.*

There exists an absolute constant $C > 1$ such that the following holds. Let $u = \log M / \log z$ and assume that $\log M \geqslant C\kappa \log z$. Then

$$\sum_{p \nmid n \,\forall p \in P} a_n = M \prod_{p \in P} \left(1 - \frac{\rho(p)}{p}\right) \cdot \left(1 + O\left(e^{-u/2}\right)\right),$$

where the implied constant is absolute.

Proof. On the one hand, note that

$$\prod_{p \in P} \left(1 - \frac{\rho(p)}{p}\right)^{-1} \leqslant \prod_{p \leqslant 2\kappa} p \prod_{2\kappa < p \leqslant z} \left(1 - \frac{\kappa}{p}\right)^{-1} \leqslant \exp\left(O(\kappa) + \kappa(\log\log z + O(1))\right) \leqslant e^{O(\kappa)}(\log z)^{\kappa}.$$

It is an easy exercise to adapt this computation and obtain, for any $2 \leqslant w \leqslant z$,

$$\prod_{\substack{p \in P \\ p \geqslant w}} \left(1 - \frac{\rho(p)}{p}\right)^{-1} \leqslant e^{O(\kappa)} \left(\frac{\log z}{\log w}\right)^{\kappa}.$$

On the other hand, for any $D \geqslant 1$ we have

$$\sum_{\substack{d \leqslant D \\ p|d \Rightarrow p \in P}} \mu(d)^2 |R_d| \ll D \sum_{\substack{d \leqslant D \\ p|d \Rightarrow p \in P}} \mu(d)^2 \frac{|\rho(d)|}{d} \leqslant D \prod_{p \in P} \left(1 + \frac{\rho(p)}{p}\right) \leqslant D \exp\left(\kappa(\log\log z + O(1))\right),$$

which is $\ll D(\log z)^{\kappa}$.

Choose $D = M^{2/3}$ and apply Lemma D.1 with $g(d) = \rho(d)/d$ and $A = e^{O(\kappa)}(\log z)^{\kappa}$. Note that

$$s = \frac{\frac{2}{3}\log M}{\log z} \geqslant C\kappa$$

by assumption. So, if C is sufficiently large,

$$\sum_{p \nmid n \,\forall p \in P} a_n = M \prod_{p \in P} \left(1 - \frac{\rho(p)}{p}\right) \cdot \left(1 + O\left(e^{-u/2}\right)\right) + O\left(M^{2/3}(\log z)^{\kappa}\right).$$

To obtain the desired conclusion, it remains to check that

$$M^{2/3}(\log z)^{\kappa} \ll \frac{M}{e^{O(\kappa)}(\log z)^{\kappa}} e^{-u/2},$$

which also follows from our assumption $\log M \geqslant C\kappa \log z$. $\qquad\square$

Lemma D.3. *Let X be a subset of $\mathcal{P} \times [\![K]\!]$. Suppose that*

$$|\mathrm{pr}_1(X)| \leqslant 2KJV.$$

Let $\mathbf{n}$ be a random variable taking values in $\prod_{p \in \mathcal{P}} \mathbb{Z}/p\mathbb{Z}$ with the uniform distribution. Then

$$\frac{1}{N} \sum_{\substack{n \in I_N \\ p|n+b_i \,\forall(p,i) \in X \\ p \nmid n+b_i \,\forall(p,i) \notin X}} \mathbf{1}_{p|n+b_i \,\forall(p,i) \in X} = \mathbb{P}\left(\begin{matrix} \forall(p,i) \in X, \, p \mid \mathbf{n}+b_i \text{ and} \\ \forall(p,i) \notin X, \, p \nmid \mathbf{n}+b_i \end{matrix}\right) \cdot \left(1 + O\left(e^{-\sqrt{\log N}}\right)\right).$$

Proof. We can assume that X satisfies the following consistency constraints (otherwise both sides are zero and there is nothing to prove):

- If $b_i, b_j \in A$ are congruent modulo $p \in \mathcal{P}$, then $(p, i) \in X$ if and only if $(p, j) \in X$.
- If $b_i, b_j \in A$ are not congruent modulo $p \in \mathcal{P}$, and $(p, i) \in X$, then $(p, j) \notin X$.
- If $\boldsymbol{b}$ covers all residue classes modulo $p \in \mathcal{P}$, then $(p, i) \in X$ for at least one $i \in [\![K]\!]$.

We now split $\mathcal{P}$ into two subsets: $\mathcal{P}^+$, the set of $p \in \mathcal{P}$ such that $(p, i) \in X$ for at least one $i \in [\![K]\!]$; and its complement $\mathcal{P}^- := \mathcal{P} \setminus \mathcal{P}^+$.

By the Chinese remainder theorem, there is a progression $a + q\mathbb{Z}$, with $q = \prod_{p \in \mathcal{P}^+} p$ such that

$$\frac{1}{N} \sum_{\substack{n \in I_N \\ p \mid n+b_i \ \forall (p,i) \in X \\ p \nmid n+b_i \ \forall (p,i) \notin X}} \mathbf{1}_{p \mid n+b_i \ \forall (p,i) \in X} = \frac{1}{N} \sum_{n \in I_N \cap (a+q\mathbb{Z})} \mathbf{1}_{p \nmid n+b_i \ \forall p \in \mathcal{P}^- \ \forall i \in [\![K]\!]}.$$

We can rewrite the latter sum as $\sum_{p \nmid m \ \forall p \in \mathcal{P}^-} a_m$ where

$$a_m := \sum_{n \in I_N \cap (a+q\mathbb{Z})} \mathbf{1}_{m = \prod_{i \in [\![K]\!]} (n+b_i)}.$$

We wish to use Corollary D.2. It is easy to show that

$$\sum_{d \mid m} a_m = \sum_{n \in I_N \cap (a+q\mathbb{Z})} \mathbf{1}_{d \mid \prod_{i \in [\![K]\!]} (n+b_i)} = \frac{\rho(d)}{d} \frac{N}{q} + O(\rho(d)),$$

where the multiplicative function $\rho(d)$ counts the number of solutions to $\prod_{i \in [\![K]\!]} (x + b_i) \equiv 0 \pmod{d}$. Note that $\rho(p) \leqslant K$ for all $p \in \mathcal{P}^-$ and $\rho(p) \leqslant p - 1$ since $\boldsymbol{b}$ does not cover all residue classes modulo $p \in \mathcal{P}^-$ by one of our preliminary assumptions.

We now apply the Fundamental Lemma in the form of Corollary D.2, with $g(d) = \rho(d)/d$, $M = N/q$, $z = H$ and $\kappa = K$. The hypothesis $\log M \geqslant C\kappa \log z$ is satisfied, since it can be rewritten as $\log N \geqslant \log q + CK \log H$ and we know that $q \leqslant H^{2KJV} \leqslant N^{1/2}$ by our choice of parameters. We conclude that

$$\frac{1}{N} \sum_{\substack{n \in I_N \\ p \mid n+b_i \ \forall (p,i) \in X \\ p \nmid n+b_i \ \forall (p,i) \notin X}} \mathbf{1}_{p \mid n+b_i \ \forall (p,i) \in X} = \prod_{p \in \mathcal{P}^+} \frac{1}{p} \prod_{p \in \mathcal{P}^-} \left(1 - \frac{\rho(p)}{p}\right) \cdot \left(1 + O\left(e^{-\sqrt{\log N}}\right)\right),$$

which is exactly what we wanted by definition of the random variable $\mathbf{n}$ and the consistency constraints above. $\qquad \square$

Proof of Lemma 6.2. We start by removing, in (31), the condition that $n + b_i \in I_N$ for every i, and only require the starting vertex n to be in I_N. Since $|b_i| \leqslant KH$ for all $i \in [\![K]\!]$, we have

$$\sum_{\boldsymbol{d} \in \mathbf{D}_K} \sum_{\substack{n \in I_N \\ \exists i, \ n+b_i \notin I_N}} 1 \ll (2H)^K \cdot KH \ll N.$$

Given $\boldsymbol{d} \in \mathbf{D}_K$ and a subset X of $\mathcal{P} \times [\![K]\!]$, write $I_N(X, \boldsymbol{d})$ for the set of all $n \in I_N$ such that, for all $(p, i) \in \mathcal{P} \times [\![K]\!]$,

$$p \mid n + b_i \iff (p, i) \in X.$$

Summing over all possibilities for X, the expression (31) becomes

$$\text{Tr}\left((\text{Ad}_G)^K\right) = \sum_{\boldsymbol{d} \in \mathbf{D}_K} \sum_{X \subset \mathcal{P} \times [\![K]\!]} \sum_{n \in I_N(X, \boldsymbol{d})} w_{\boldsymbol{d}}(n) \ + O(N).$$

The important observation is that, for fixed $\boldsymbol{d}$ and X, the term $w_{\boldsymbol{d}}(n)$ is independent of $n \in I_N(X, \boldsymbol{d})$. We may thus call it $w_{\boldsymbol{d}}(X)$, and rewrite the triple sum as

$$(91) \qquad \sum_{\boldsymbol{d} \in \mathbf{D}_K} \sum_{X \subset \mathcal{P} \times \llbracket K \rrbracket} w_{\boldsymbol{d}}(X) \left| I_N(X, \boldsymbol{d}) \right|.$$

If $|\mathrm{pr}_1(X)| > 2KJV$ (pr_1 being the projection on the first coordinate), the coefficient $w_{\boldsymbol{d}}(X)$ is zero, since in that case one of the factors W vanishes. Otherwise, by Lemma D.3,

$$|I_N(X, \boldsymbol{d})| = N \cdot \mathbb{P} \begin{pmatrix} \forall (p, i) \in X,\ p \mid \mathbf{n} + b_i \ \text{ and} \\ \forall (p, i) \notin X,\ p \nmid \mathbf{n} + b_i \end{pmatrix} \cdot \left(1 + O\left(e^{-\sqrt{\log N}} \right) \right).$$

Hence, we can interpret the sum over X as the expected value of $w_{\boldsymbol{d}}(\mathbf{n})$, with a small error term. More precisely, (91) is

$$N \sum_{\boldsymbol{d} \in \mathbf{D}_K} \mathbb{E}\left[w_{\boldsymbol{d}}(\mathbf{n}) \right] + O\left(N e^{-\sqrt{\log N}} \right) \sum_{\boldsymbol{d} \in \mathbf{D}_K} \mathbb{E}\left[|w_{\boldsymbol{d}}(\mathbf{n})| \right],$$

which concludes the proof of Lemma 6.2. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

References

[1] Rajendra Bhatia, *Matrix analysis*, vol. 169, Springer, 1997.

[2] Sarvadaman Chowla, *The Riemann hypothesis and Hilbert's tenth problem*, Mathematics and Its Applications, vol. 4, Gordon and Breach Science Publishers, New York-London-Paris, 1965.

[3] John B. Friedlander and Henryk Iwaniec, *Opera de Cribro*, vol. 57, American Mathematical Society, 2010.

[4] Harald A. Helfgott, *Expansion, divisibility and parity: an explanation*, Combinatorial and Additive Number Theory, New York Number Theory Seminar, Springer, 2021, pp. 199–237.

[5] Harald A. Helfgott and Maksym Radziwiłł, *Expansion, divisibility and parity*, arXiv:2103.06853 (2021).

[6] Harald A. Helfgott and Adrián Ubis, *Primos, paridad y análisis*, arXiv:1812.08707 (2018).

[7] Jonas Kubilius, *Probabilistic methods in the theory of numbers*, Translations of Mathematical Monographs, vol. 11, American Mathematical Society, 1964.

[8] Kaisa Matomäki and Maksym Radziwiłł, *Multiplicative functions in short intervals*, Annals of Mathematics (2016), 1015–1056.

[9] Kaisa Matomäki, Maksym Radziwiłł, and Terence Tao, *An averaged form of Chowla's conjecture*, Algebra Number Theory **9** (2015), no. 9, 2167–2196.

[10] ———, *Sign patterns of the Liouville and Möbius functions*, Forum of Mathematics, Sigma **4** (2016), e14.

[11] Karl K. Norton, *On the number of restricted prime factors of an integer. I*, Illinois Journal of Mathematics **20** (1976), no. 4, 681–705.

[12] Bruce E. Sagan, Yeong-Nan Yeh, and Günter M. Ziegler, *Maximizing Möbius functions on subsets of Boolean algebras*, Discrete Mathematics **126** (1994), no. 1-3, 293–311.

[13] Bernd S. W. Schröder, *Ordered sets: an introduction*, vol. 29, Springer, 2003.

[14] Terence Tao, *The Erdős discrepancy problem*, Discrete Analysis (2016), 609.

[15] ———, *The logarithmically averaged Chowla and Elliott conjectures for two-point correlations*, Forum of Mathematics, Pi **4** (2016).

[16] Terence Tao and Joni Teräväinen, *Odd order cases of the logarithmically averaged Chowla conjecture*, Journal de Théorie des Nombres de Bordeaux **30** (2018), no. 3, 997–1015.

[17] ———, *The structure of correlations of multiplicative functions at almost all scales, with applications to the Chowla and Elliott conjectures*, Algebra & Number Theory **13** (2019), no. 9, 2103–2150.

[18] ———, *The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures*, Duke Mathematical Journal **168** (2019), no. 11, 1977–2027.

[19] Gérald Tenenbaum, *Introduction to analytic and probabilistic number theory*, third ed., vol. 163, American Mathematical Society, 2015.

Mathematical Institute, University of Oxford.

Email address: `cedric.pilatte@maths.ox.ac.uk`