TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 369, Number 12, December 2017, Pages 8369–8388
http://dx.doi.org/10.1090/tran/6903
Article electronically published on May 5, 2017

PRESCRIBING THE BINARY DIGITS OF SQUAREFREE NUMBERS AND QUADRATIC RESIDUES

RAINER DIETMANN, CHRISTIAN ELSHOLTZ, AND IGOR E. SHPARLINSKI

Abstract. We study the equidistribution of multiplicatively defined sets, such as the squarefree integers, quadratic non-residues or primitive roots, in sets which are described in an additive way, such as sumsets or Hilbert cubes. In particular, we show that if one fixes any proportion less than 40% of the digits of all numbers of a given binary bit length, then the remaining set still has the asymptotically expected number of squarefree integers. Next, we investigate the distribution of primitive roots modulo a large prime p, establishing a new upper bound on the largest dimension of a Hilbert cube in the set of primitive roots, improving on a previous result of the authors. Finally, we study sumsets in finite fields and asymptotically find the expected number of quadratic residues and non-residues in such sumsets, given that their cardinalities are big enough. This significantly improves on a recent result by Dartyge, Mauduit and Sárközy. Our approach introduces several new ideas, combining a variety of methods, such as bounds of exponential and character sums, geometry of numbers and additive combinatorics.

1. Introduction

1.1. **Motivation.** Given an integer $n \ge 2$, we denote by $\mathcal{D}_n$ the set of vectors $\mathbf{d} = (\delta_0, \ldots, \delta_{n-1})$ where $\delta_i \in \{*, 0, 1\}$, $i = 0, \ldots, n-1$, and for $\mathbf{d} \in \mathcal{D}_n$ we consider the set

$$\mathcal{N}_n(\mathbf{d}) = \left\{ \sum_{i=0}^{n-1} d_i 2^i \ : \ d_i \in \{0,1\} \text{ if } \delta_i = *, \ d_i = \delta_i \text{ otherwise} \right\}.$$

Furthermore, for integers $n > k \ge 1$ we denote by $\mathcal{D}_{k,n}$ the subset of n-dimensional vectors $\mathbf{d} \in \mathcal{D}_n$ with exactly k components of $\mathbf{d}$ that are fixed as either 0 or 1 and $n-k$ components that are $*$. We also use $\mathcal{D}^*_{k,n}$ to denote the set of vectors $\mathbf{d} \in \mathcal{D}_{k,n}$ with $\delta_0 = 1$. In particular, for $\mathbf{d} \in \mathcal{D}^*_{k,n}$ all elements of $\mathcal{N}_n(\mathbf{d})$ are odd.

Various arithmetic properties of elements from $\mathcal{N}_n(\mathbf{d})$ as well as of other integers with restricted digits have been studied in a number of works.

We first recall that Bourgain [5, 6] has recently obtained several very strong results about prime numbers with prescribed binary digits; see also [24]. For example, the result of [6] gives an asymptotic formula for the number of primes $p \in \mathcal{N}_n(\mathbf{d})$ for very dense vectors $\mathbf{d} \in \mathcal{D}^*_{k,n}$, more precisely, when $k \le cn$ for some absolute constant $c > 0$, which is a dramatic improvement over the previous results of [5, 24].

Received by the editors June 17, 2015 and, in revised form, January 12, 2016.

2010 *Mathematics Subject Classification.* Primary 11A63, 11B30, 11N25; Secondary 11H06, 11L40, 11P70, 11T30.

Key words and phrases. Digital problems, square-free numbers, non-residues, finite fields, Hilbert cubes.

Mauduit and Rivat [33] have recently settled a problem of Gelfond about the distribution of primes with the sums of digits in a prescribed arithmetic progression; see also [17]. Partially motivated by some cryptographic applications, the distribution and construction of RSA moduli and smooth numbers with some binary digits prescribed have been studied in [21, 40]. Various results on prime divisors and other arithmetic properties of numbers with very few non-zero binary digits can be found in [2, 4, 18, 32, 39, 41]. For a diverse variety of results on integers with various restrictions on their digits (for example, palindromes), see [3, 9, 10, 16, 29, 31, 34, 35] and references therein.

1.2. **Our results and methods.** Here we combine a variety of methods, such as bounds on exponential and character sums, geometry of numbers, and additive combinatorics, to derive new results about the arithmetic structure of elements of $\mathcal{N}_n(\mathbf{d})$. Throughout, our goal is to treat $\mathbf{d} \in \mathcal{D}_{k,n}$ with the ratio k/n as large as possible (that is, for thin sets of integers with as large as possible proportion of pre-assigned digits). We believe that our ideas and results can find application to several other problems of this type.

More precisely, in Section 3.1 we first study the distribution of squarefree numbers in sets $\mathcal{N}_n(\mathbf{d})$. Using some combinatorial arguments, the theory of lattice minima and a result of Bourgain [6, Lemma 4] we obtain an asymptotic formula for the number of squarefree integers $s \in \mathcal{N}_n(\mathbf{d})$ for $\mathbf{d} \in \mathcal{D}^*_{k,n}$ provided that $k \le (2/5 - \varepsilon)n$ for any fixed $\varepsilon > 0$. In Section 3.2 we also give an asymptotic for the sums of values of the Euler function in essentially full range $k \le (1 - \varepsilon)n$.

Furthermore, we also estimate multiplicative character sums and in Section 3.3 obtain results about the distribution of quadratic non-residues and primitive roots modulo p among the elements of $\mathcal{N}_n(\mathbf{d})$ for $\mathbf{d} \in \mathcal{D}_{k,n}$ provided that $k \le (1/2 - \varepsilon)n$ for any fixed $\varepsilon > 0$. This result complements those of [1, 14, 15, 36], where similar questions are considered for integers with various restrictions on their binary digits (and also digits in other bases).

Finally, in Section 3.4, we consider a related question about quadratic residues and primitive roots in *Hilbert cubes*. For a prime power $r = p^n$, let $\mathbb{F}_r$ denote the finite field of r elements. For $a_0, a_1, \ldots, a_d \in \mathbb{F}_p$ we define the *Hilbert cube* as

$$\mathcal{H}(a_0; a_1, \ldots, a_d) = \left\{ a_0 + \sum_{i=1}^{d} \vartheta_i a_i \ : \ \vartheta_i \in \{0, 1\} \right\}. \tag{1.1}$$

We define $f(p)$ as the largest d such that there are $a_0, a_1, \ldots, a_d \in \mathbb{F}_p$ with pairwise distinct $a_1, \ldots, a_d$ such that $\mathcal{H}(a_0; a_1, \ldots, a_d)$ does not contain a quadratic non-residue modulo p. Furthermore, we define $F(p)$ as the largest d such that there are $a_0, a_1, \ldots, a_d \in \mathbb{F}_p$ with pairwise distinct $a_1, \ldots, a_d$ such that $\mathcal{H}(a_0; a_1, \ldots, a_d)$ does not contain a primitive root modulo p. Clearly

$$f(p) \le F(p).$$

Hegyvári and Sárközy [25, Theorem 2] give the bound $f(p) < 12p^{1/4}$ for sufficiently large p, which has been improved to

$$F(p) \le p^{1/5 + o(1)}$$

as $p \to \infty$, by Dietmann, Elsholtz and Shparlinski [14, Theorem 1.3]. Here we improve this further to

$$F(p) \le p^{3/19 + o(1)}$$

and recall that reducing the exponent below $1/8$ immediately implies an improvement of the Burgess bound on the least primitive root (note that $3/19 - 1/8 = 0.0328\ldots$).

As a further application of our method of Section 3.3, in Section 3.5, we outline a substantial improvement of one of the results of Dartyge, Mauduit and Sárközy [11]. Namely, given a basis $\omega_1, \ldots, \omega_n$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ and a collection of n sets $\mathfrak{A} = \{\mathcal{A}_i \subseteq \mathbb{F}_p \ : \ i = 1, \ldots, n\}$, we consider the set

$$\mathcal{W}_{\mathfrak{A}} = \{a_1\omega_1 + \ldots + a_n\omega_n \ : \ a_i \in \mathcal{A}_i, \ i = 1, \ldots, n\}, \tag{1.2}$$

which has a natural interpretation of elements in $\mathbb{F}_{p^n}$ "with restricted digits". Dartyge, Mauduit and Sárközy [11, Theorem 2.1] show that if for some fixed $\varepsilon > 0$ the lower bound

$$\min_{1 \le i \le r} \#\mathcal{A}_i \ge \left(\frac{\sqrt{5}-1}{2} + \varepsilon\right) p \tag{1.3}$$

holds, then, as $p \to \infty$, the set $\mathcal{W}_{\mathfrak{A}}$ contains asymptotically equal proportions of quadratic residues and non-residues (note that in [11] only the case of $\mathcal{A}_1 = \ldots = \mathcal{A}_r$ is considered, but the proof immediately extends to different sets).

Here, in Section 3.5 we prove a similar asymptotic equidistribution of quadratic residues and non-residues under a much more relaxed condition than (1.3). Namely, for our result we only assume that for some fixed $\varepsilon > 0$ we have

$$\prod_{1 \le i \le n} \#\mathcal{A}_i \ge p^{(1/2+\varepsilon)n^2/(n-1)} \qquad \text{and} \qquad \min_{1 \le i \le n} \#\mathcal{A}_i \ge p^{\varepsilon}.$$

For $n \ge 3$ this is a much wider range of parameters than the earlier restriction (1.3) that is linear in p. We remark that Gabdullin [19, 20] has recently achieved further progress on this problem.

1.3. **Notation.** Throughout the paper the implied constants in the symbols "O" and "$\ll$" may depend on the real parameter $\varepsilon > 0$ and an integer parameter $\nu \ge 1$. We recall that the expressions $A \ll B$ and $A = O(B)$ are each equivalent to the statement that $|A| \le cB$ for some constant c.

As usual, $\log z$ denotes the natural logarithm of z.

The letter p always denotes a prime.

As we have mentioned, we use $\mathbb{F}_r$ to denote the finite field of r elements.

2. Preparations

2.1. **Bounds of some exponential and character sums.** We need the following result of Bourgain [6, Lemma 4].

Lemma 2.1. *Let $n > k \ge 1$ and $\mathbf{d} \in \mathcal{D}^*_{k,n}$. Then for any integers a and q with $\gcd(2a, q) = 1$ and $3 \le q \le n^{1/10\kappa}$, where $\kappa = k/n$, we have*

$$\left| \sum_{s \in \mathcal{N}_n(\mathbf{d})} \exp(2\pi i a s/q) \right| < \#\mathcal{N}_n(\mathbf{d}) 2^{-\sqrt{n}}.$$

We need the following bound of a double character sum due to Karatsuba [27] (see also [28, Chapter VIII, Problem 9]), which in turn follows from the Weil bound (see [26, Corollary 11.24]) and the Hölder inequality.

We present it in the settings of arbitrary finite fields.

Lemma 2.2. *For any integer $\nu \ge 1$, any sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{F}_r$ and any non-trivial multiplicative character χ of $\mathbb{F}_r$, we have*

$$\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}}\chi(a+b) \ll (\#\mathcal{A})^{1-1/2\nu}\#\mathcal{B}r^{1/4\nu} + (\#\mathcal{A})^{1-1/2\nu}(\#\mathcal{B})^{1/2}r^{1/2\nu},$$

where the implied constant depends only on ν.

In particular, taking $\nu = \lceil \varepsilon^{-1} \rceil$ for a fixed $\varepsilon > 0$ we derive from Lemma 2.2:

Corollary 2.3. *For any $\eta > 0$ there exists $\delta > 0$ such that for any sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{F}_r$ with $\#\mathcal{A} \ge r^{1/2+\eta}$ and $\#\mathcal{B} \ge r^{\eta}$ and any non-trivial multiplicative character χ of $\mathbb{F}_r$, we have*

$$\sum_{a\in\mathcal{A}}\sum_{b\in\mathcal{B}}\chi(a+b) \ll \#\mathcal{A}\#\mathcal{B}r^{-\delta},$$

where the implied constant depends only on η.

We make use of the following special case of a result of Shao [37].

Lemma 2.4. *Let $\nu \ge 1$ be a fixed integer. Let $0 \le w_1 < \ldots < w_J < p$ be $J \ge 1$ arbitrary integers with*

$$w_{j+1} - w_j \ge H, \qquad j = 1, \ldots, J-1,$$

for some real $H \ge p^{1/2\nu}$. Then for any non-principal multiplicative character χ modulo p, we have

$$\sum_{j=1}^{J-1}\max_{h\le H}\left|\sum_{i=1}^{h}\chi\left(i+w_j\right)\right|^{2\nu} \ll p^{1/2+1/2\nu+o(1)}H^{2\nu-2}.$$

2.2. **Bounds of the number of solutions of some congruences.** For $\mathbf{d} \in \mathcal{D}_n$ and an integer $q \ge 2$ we consider the set

$$\mathcal{N}_n(\mathbf{d}, q) = \left\{s \in \mathcal{N}_n(\mathbf{d}) \ : \ s \equiv 0 \pmod q\right\}.$$

Lemma 2.5. *Let $n > k \ge 1$ and $\mathbf{d} \in \mathcal{D}^*_{k,n}$. Then for any odd q with $1 < q \le n^{1/10\kappa}$, where $\kappa = k/n$, we have*

$$\#\mathcal{N}_n(\mathbf{d}, q) = \frac{1}{q}\#\mathcal{N}_n(\mathbf{d}) + O\left(\#\mathcal{N}_n(\mathbf{d})2^{-\sqrt{n}}\right).$$

Proof. Using the orthogonality of exponential functions we write

$$\#\mathcal{N}_n(\mathbf{d}, q) = \sum_{s\in\mathcal{N}_n(\mathbf{d})}\frac{1}{q}\sum_{a=0}^{q-1}\exp(2\pi ias/q) = \frac{1}{q}\sum_{a=0}^{q-1}\sum_{s\in\mathcal{N}_n(\mathbf{d})}\exp(2\pi ias/q).$$

The term corresponding to $a = 0$ is equal to $\#\mathcal{N}_n(\mathbf{d})/q$, while it is easy to see that Lemma 2.1 also applies to exponential sums with denominators $q/\gcd(a, q) \ge 3$ instead of q. □

For larger values of q we only have an upper bound on $\#\mathcal{N}_n(\mathbf{d}, q)$.

For real positive κ and ϱ we define

$$\tau(\kappa, \varrho) = \frac{1 + \varrho - \sqrt{(1-\varrho)^2 + 4\varrho\kappa}}{2} \tag{2.1}$$

as the root of the equation

$$\tau^2 - \tau(1+\varrho) + \varrho(1-\kappa) = 0,$$

which belongs to the interval $[0, \varrho]$. We now set

$$\vartheta(\kappa, \varrho) = \frac{\tau(\kappa, \varrho)}{\varrho}. \tag{2.2}$$

Lemma 2.6. *Let $\varepsilon > 0$ be fixed. Let q be an odd integer and let*

$$n(1-\varepsilon) > k \geq 1 \qquad \text{and} \qquad 2^{n(1-\varepsilon)} \geq q \geq 1.$$

Then for any $\mathbf{d} \in \mathcal{D}^*_{k,n}$ we have

$$\#\mathcal{N}_n(\mathbf{d}, q) \ll \#\mathcal{N}_n(\mathbf{d}) q^{-\vartheta(\kappa,\varrho)},$$

where the implied constant is absolute, κ and ϱ are defined by

$$\kappa = k/n \qquad \text{and} \qquad q = 2^{\varrho n},$$

and $\vartheta(\kappa, \varrho)$ is given by (2.2).

Proof. We refer to the digits of $s \in \mathcal{N}_n(\mathbf{d})$ on positions j with $\delta_j = *$ as *free positions* and we refer to other digits as *fixed positions*.

We set

$$r = \left\lceil \frac{\log q}{\log 2} \right\rceil - 1.$$

Let $\mathbf{d} = (\delta_0, \dots, \delta_{n-1})$.

We set $\delta_i = 0$ for all integers $i \notin [0, n-1]$.

Now, for $j \in \mathbb{Z}$, we denote by w_j the number of free positions amongst the positions $j, \dots, j+r-1$ and let χ^* be the characteristic function of the symbol '$*$' defined on the set $\{*, 0, 1\}$. Then

$$\sum_{j=-r+1}^{n-1} w_j = \sum_{j=-r+1}^{n-1} \sum_{i=0}^{r-1} \chi^*(\delta_{i+j}) = r(n-k).$$

We now set $t = \lfloor \tau(\kappa, \varrho) n \rfloor$, where $\tau(\kappa, \varrho)$ is given by (2.1). We count the total number W of free positions which appear in each of the $n + r - 2t + 1$ blocks of width r starting at the points $j = -r + t, \dots, n - t$. Then we have

$$W = \sum_{j=-r+t}^{n-t} w_j = r(n-k) - \sum_{j=-r+1}^{-r+t-1} w_j - \sum_{j=n-t+1}^{n-1} w_j. \tag{2.3}$$

We now note that for $j < 0$ we have $w_j \leq r - |j|$ and for $j < r$ we have $w_{n-j} \leq j$. Hence,

$$\sum_{j=-r+1}^{-r+t-1} w_j + \sum_{j=n-t+1}^{n-1} w_j \leq 2 \sum_{i=1}^{t-1} i = t^2 + O(t).$$

Therefore, we conclude from (2.3) that

$$W \geq r(n-k) - t^2 + O(t).$$

Hence, for some $h \in [-r+t, n-t]$ we have

$$\begin{aligned} w_h &\ge \frac{W}{n+r-2t+1} \ge \frac{r(n-k)-t^2+O(t)}{n+r-2t+1} \\ &= n\frac{\varrho(1-\kappa)-\tau(\kappa,\varrho)^2}{1+\varrho-2\tau(\kappa,\varrho)}+O(1) \\ &= n\tau(\kappa,\varrho)+O(1) = n\varrho\vartheta(\kappa,\varrho)+O(1), \end{aligned} \tag{2.4}$$

where the implied constant is absolute.

Now fixing the digits on the remaining $n-k-w_h$ free positions $j \notin [h, h+r-1]$ of the numbers

$$\sum_{i=0}^{n-1} d_i 2^i \in \mathcal{N}_n(\mathbf{d})$$

and recalling that $2^r < q$, we see that the number

$$s = \sum_{i=h}^{h+r-1} d_i 2^{i-h}$$

belongs to a prescribed residue class modulo q, and since $0 \le s < 2^r < q$, s is uniquely defined. Hence, using (2.4), we obtain

$$\#\mathcal{N}_n(\mathbf{d},q) \le 2^{n-k-w_h} \le \#\mathcal{N}_n(\mathbf{d},q)2^{-w_h} \ll \#\mathcal{N}_n(\mathbf{d},q)q^{-\vartheta(\kappa,\varrho)},$$

and the result follows. □

Lemma 2.7. *Let $\varepsilon > 0$ be sufficiently small and*

$$\frac{1}{2} \ge \varrho \ge \frac{1}{4}. \tag{2.5}$$

Moreover, let

$$\kappa = \frac{k}{n} < \frac{3}{7} - 2\varepsilon \tag{2.6}$$

and

$$2^{\varrho n} \ll A \ll 2^{\varrho n},$$

and suppose that

$$(3+4\varepsilon)\varrho \le 2(1-\kappa) \tag{2.7}$$

and

$$\varrho(1+5\varepsilon) < 4(1-\kappa) - 2 - \varepsilon. \tag{2.8}$$

Then

$$\sum_{A<q\le 2A} \#\mathcal{N}_n(\mathbf{d},q^2) \ll \#\mathcal{N}_n(\mathbf{d})A^{-\varepsilon/2}. \tag{2.9}$$

Proof. We follow the definition of free and fixed positions as in the proof of Lemma 2.6.

Let us divide the set of all n positions into three blocks W_1, W_2, W_3 of consecutive positions (from the left to the right) in the following way: The number of positions in $W_1 \cup W_2$ as well as in $W_2 \cup W_3$ is $2\varrho n + O(1)$. This is certainly possible since we have (2.5) (more explicitly, W_1 and W_3 contain $n(1-2\varrho)+O(1)$ positions and

W_2 contains $n(4\varrho-1)+O(1)$ positions). Let w_i be the number of free positions in block W_i, $i=1,2,3$. Since the total number of free positions is $(1-\kappa)n$, we obtain

$$w_1+w_2+w_3=(1-\kappa)n. \tag{2.10}$$

Now let α be the number of free positions in W_1 and W_2 together, that is, $\alpha = w_1+w_2$, and analogously let $\beta=w_1+w_3$ and $\gamma=w_2+w_3$. Then (2.10) implies that

$$\alpha+\beta+\gamma=2(1-\kappa)n. \tag{2.11}$$

Now regarding the neighbouring blocks W_1 and W_2 as one block with α free positions, as in the proof of Lemma 2.6 we obtain

$$\#\mathcal{N}_n(\mathbf{d},q^2)\ll \#\mathcal{N}_n(\mathbf{d})2^{-\alpha} \tag{2.12}$$

whenever $A<q\le 2A$. Note that here we use the fact that $A\gg 2^{\varrho n}$, whence $q^2\gg 2^{2\varrho n}$, so a congruence modulo q^2 fixes all the free positions in the block composed of W_1 and W_2. Analogously, we obtain the alternative bound

$$\#\mathcal{N}_n(\mathbf{d},q^2)\ll \#\mathcal{N}_n(\mathbf{d})2^{-\gamma}, \tag{2.13}$$

using the block composed of W_2 and W_3. Our first observation is that we can assume that $\alpha<(1+\varepsilon)\varrho n$, as otherwise (2.12) implies

$$\#\mathcal{N}_n(\mathbf{d},q^2)\ll \#\mathcal{N}_n(\mathbf{d})A^{-1-\varepsilon}$$

and the result follows. Similarly, using (2.13), we can assume that $\gamma<(1+\varepsilon)\varrho n$. Hence, by (2.11) we can also assume that

$$\beta\ge 2n(1-\kappa-(1+\varepsilon)\varrho). \tag{2.14}$$

Note that by (2.7), this implies that

$$2^{-\beta}\ll A^{-1-\varepsilon}. \tag{2.15}$$

Moreover, as trivially $\beta\le(1-\kappa)n$ where $(1-\kappa)n$ is the total number of free positions, we obtain

$$\varrho\ge\frac{1-\kappa}{2(1+\varepsilon)}.$$

By (2.6), for sufficiently small $\varepsilon>0$ this implies that

$$\varrho>\frac{2}{3}\kappa+\varepsilon,$$

whence

$$\varrho n-\beta+(2-6\varrho)n\le n(2\kappa-3\varrho+2\varepsilon\varrho)\le-\varepsilon n \tag{2.16}$$

for sufficiently small $\varepsilon>0$.

Working with W_1 and W_3 is more difficult, as we are no longer dealing with one, but rather with two intervals. Writing r for the bit position at the right of W_1 and s for the position at the right of W_2, we are now considering congruences of the form

$$2^r a+2^s b+c\equiv 0 \pmod{q^2}. \tag{2.17}$$

Note that from the construction of W_1, W_2, W_3 it follows that

$$r\ge 2\varrho n+O(1).$$

Once b, corresponding to W_2, has been fixed, the solution set of (2.17) is of the form

$$(a, c) = (a_0, c_0) + (a_1, c_1), \tag{2.18}$$

where $(a_0, c_0) \in \mathbb{Z}^2$ is a fixed solution of (2.17) and (a_1, c_1) runs over all solutions of the homogeneous congruence

$$2^r a_1 + c_1 \equiv 0 \pmod{q^2}. \tag{2.19}$$

By construction of the W_i, we see that a and c are non-negative integers with $a, c \ll 2^{(1-2\varrho)n}$, whence also $|a_1|, |c_1| \ll 2^{(1-2\varrho)n}$. Moreover, the congruence (2.19) describes a two-dimensional lattice with a basis $\{(1, -2^r), (0, q^2)\}$ and of determinant q^2. Let $2^{\lambda_1(q)}, 2^{\lambda_2(q)}$ be its successive minima, where $\lambda_1(q) \le \lambda_2(q)$. For the general background on lattices we refer to [22].

Then

$$q^2 \ll 2^{\lambda_1(q)+\lambda_2(q)} \ll q^2.$$

Let us first discuss the case that

$$\lambda_2(q) \le (1 - 2\varrho)n.$$

Then the number of solutions to (2.19) with $|a_1|, |c_1| \ll 2^{(1-2\varrho)n}$ can be estimated as

$$O\left(\left(2^{(1-2\varrho)n-\lambda_1(q)} + 1\right)\left(2^{(1-2\varrho)n-\lambda_2(q)} + 1\right)\right) = O\left(2^{2(1-2\varrho)n} q^{-2}\right)$$

(note that $q^2 \ll 2^{\lambda_1(q)+\lambda_2(q)} \le 2^{2\lambda_2(q)} \le 2^{2(1-2\varrho)n}$). Furthermore, since $q^2 \ge A^2 \gg 2^{2\varrho n}$, we obtain the bound $O\left(2^{(2-6\varrho)n}\right)$ for the number of solutions to (2.19).

Considering all the possible

$$2^{w_2} = 2^{n-k-\beta} = \#\mathcal{N}_n(\mathbf{d}) 2^{-\beta}$$

choices for b, we therefore obtain

$$\#\mathcal{N}_n(\mathbf{d}, q^2) \ll \#\mathcal{N}_n(\mathbf{d}) 2^{(2-6\varrho)n-\beta}.$$

By (2.16), this contribution, when summed over $A < q \le 2A$, is negligible with respect to (2.9).

We may therefore without loss of generality assume that

$$\lambda_2(q) > (1 - 2\varrho)n.$$

Again, from (2.18) we conclude that the number of solutions of (2.17) with $|a_1|$, $|c_1| \ll 2^{(1-2\varrho)n}$ is

$$O\left(\left(2^{(1-2\varrho)n-\lambda_1(q)} + 1\right)\left(2^{(1-2\varrho)n-\lambda_2(q)} + 1\right)\right) = O\left(2^{(1-2\varrho)n-\lambda_1(q)} + 1\right),$$

and the number of possible choices for b is bounded by $\#\mathcal{N}_n(\mathbf{d}) 2^{-\beta}$, so

$$\#\mathcal{N}_n(\mathbf{d}, q^2) \ll \#\mathcal{N}_n(\mathbf{d}) 2^{-\beta}(2^{(1-2\varrho)n-\lambda_1(q)} + 1). \tag{2.20}$$

Let us define a real parameter

$$\lambda = (1 - 2\varrho)n - 2(1 - \kappa)n + 3(1 + \varepsilon)\varrho n. \tag{2.21}$$

If $\lambda_1(q) > \lambda$, then (2.14) and (2.15) give

$$2^{-\beta}\left(2^{(1-2\varrho)n-\lambda_1(q)} + 1\right) \le 2^{-(1+\varepsilon)\varrho n} + 2^{-\beta} \ll A^{-1-\varepsilon},$$

so

$$\sum_{\substack{A<q\le 2A:\\ \lambda_1(q)\ge\lambda}} \#\mathcal{N}_n(\mathbf{d},q^2) \ll \#\mathcal{N}_n(\mathbf{d})A^{-\varepsilon}.$$

It now remains to estimate the contribution from q with $\lambda_1(q)\le\lambda$. Furthermore, it is enough to show that for any real positive $\mu<\lambda$ we have

$$\sum_{\substack{A<q\le 2A:\\ \mu\le\lambda_1(q)<\mu+1}} \#\mathcal{N}_n(\mathbf{d},q^2) \ll \#\mathcal{N}_n(\mathbf{d})A^{-\varepsilon}. \tag{2.22}$$

Now $\lambda_1(q)\le\mu+1$ means that there exist $a_1,c_1\in\mathbb{Z}$, not both zero, such that $|a_1|,|c_1|\ll 2^{\mu}$ and (2.19) holds true. Note that $2^r a_1+c_1=0$ is impossible, as it implies that $2^r\mid c_1$, so $|c_1|\ge 2^r\gg 2^{2\varrho n}$, contradicting $|c_1|\ll 2^{\mu}$ as by (2.5) and (2.7) we have

$$\mu<\lambda\le(1-2\varrho)n-\varepsilon\varrho n\le 2\varrho n-\varepsilon\varrho n.$$

Therefore, $2^r a_1+b_1\ne 0$, so by (2.19) for each fixed pair (a_1,c_1) there are only $2^{o(n)}$ possibilities for q that are integer divisors of $2^r a_1+b_1=O(2^n)$; see [23, Theorem 317]. The number of possible (a_1,c_1) can be bounded by $O(2^{2\mu})$, and $\#N_n(\mathbf{d},q^2)$, by (2.20), is at most of order of magnitude

$$\#\mathcal{N}_n(\mathbf{d})(2^{-\beta+(1-2\varrho)n-\mu}+2^{-\beta})\ll\#\mathcal{N}_n(\mathbf{d})(2^{-\beta+(1-2\varrho)n-\mu}+A^{-1-\varepsilon}).$$

We therefore obtain

$$\begin{aligned}\sum_{\substack{A<q\le 2A:\\ \mu\le\lambda_1(q)<\mu+1}} \#\mathcal{N}_n(\mathbf{d},q^2) &\ll \#\mathcal{N}_n(\mathbf{d})(2^{-\beta+(1-2\varrho)n+\mu+n\varepsilon}+A^{-\varepsilon})\\ &\ll \#\mathcal{N}_n(\mathbf{d})(2^{-\beta+(1-2\varrho)n+\lambda+o(n)}+A^{-\varepsilon}).\end{aligned}$$

Now by (2.8), (2.14) and (2.21), we have

$$-\beta+(1-2\varrho)n+\lambda<-\varepsilon n,$$

completing the proof. □

In particular, covering an interval $[A,B]$ by dyadic intervals, we derive from Lemma 2.7 the following result suitable for our applications.

Corollary 2.8. *Let $\varepsilon>0$ and*

$$\kappa=\frac{k}{n}<\frac{3}{7}-2\varepsilon.$$

Suppose that

$$\frac12\ge\zeta\ge\xi\ge\frac14,\qquad (3+4\varepsilon)\zeta\le 2(1-\kappa),\qquad \zeta(1+5\varepsilon)<4(1-\kappa)-2-\varepsilon.$$

Then for any A and B with

$$2^{\xi n}\ll A\le B\ll 2^{\zeta n}$$

we have

$$\sum_{A<q\le B}\#\mathcal{N}_n(\mathbf{d},q^2)\ll\#\mathcal{N}_n(\mathbf{d})A^{-\varepsilon/2}\log B.$$

Lemma 2.9. *Keeping the notation of Lemma* 2.6*, suppose that*

$$\frac{\kappa}{2} \le \varrho \le \frac{1}{2}.$$

Then

$$\#\mathcal{N}_n(\mathbf{d}, q) \ll \#\mathcal{N}_n(\mathbf{d}) q^{-1+\kappa/2\varrho}.$$

Proof. We use a similar but simpler argument as in the proof of Lemma 2.7. As $\varrho \le \frac{1}{2}$, we can divide the n bits into three blocks W_1, W_2, W_3 (from the left to the right), such that W_1 and W_3 have size $\varrho n + O(1)$ each, and W_2 has size $(1-2\varrho)n + O(1)$. Then in one of W_1 and W_3, say W_1, there must be at least

$$\frac{1}{2}\left((1-\kappa)n - (1-2\varrho)n\right) + O(1) = n(\varrho - \kappa/2) + O(1)$$

many free positions. Once all the bits outside W_1 have been chosen, a congruence modulo q fixes all $\gg n(\varrho - \kappa/2)$ remaining free positions in W_1, whence

$$\#\mathcal{N}_n(\mathbf{d}, q) \ll \#\mathcal{N}_n(\mathbf{d}) 2^{-n(\varrho-\kappa/2)} \ll \#\mathcal{N}_n(\mathbf{d}) q^{-1+\kappa/2\varrho},$$

which concludes the proof. □

To apply Lemma 2.6 we also need the following technical statement.

Lemma 2.10. *For* $1 > \kappa > 0$ *the function* $\vartheta(\kappa, \varrho)$ *given by* (2.2) *is monotonically decreasing as* ϱ *is increasing.*

Proof. The result follows from the observation that the derivative

$$\frac{\partial \vartheta}{\partial \varrho} = \frac{1 + (-1+2\kappa)\varrho - \sqrt{1 + (-2+4\kappa)\varrho + \varrho^2}}{2\varrho^2\sqrt{1 + (-2+4\kappa)\varrho + \varrho^2}}$$

is negative, as $1 + 2c\varrho + \varrho^2 > 0$ and $1 + c\varrho - \sqrt{1 + 2c\varrho + \varrho^2} < 0$ when $|c| < 1$. □

2.3. **Some results from additive combinatorics.** We now recall a recent result by Schoen [38, Theorem 3.3] in additive combinatorics. As in [14], we note that [38, Theorem 3.3] is only stated for subset sums but can be easily extended to Hilbert cubes.

Lemma 2.11. *For any* $a_0 \in \mathbb{F}_p$ *and pairwise distinct* $a_1, \ldots, a_d \in \mathbb{F}_p$ *such that* $d \ge 8(p/\log p)^{1/D}$*, where* D *is an integer satisfying*

$$0 < D \le \sqrt{\frac{\log p}{2 \log\log p}},$$

the Hilbert cube (1.1) *contains an arithmetic progression of length* L *where*

$$L \ge 2^{-10}(d/\log p)^{1+1/(D-1)}.$$

For a set $\mathcal{S} \subseteq \mathbb{F}_p$ we use $\Sigma_k(\mathcal{S})$ to denote the set of all k-elements subset sums of $\mathcal{S}$, that is,

$$\Sigma_k(\mathcal{S}) = \left\{ \sum_{t \in \mathcal{T}} t \ : \ \mathcal{T} \subseteq \mathcal{S}, \ \#\mathcal{T} = k \right\}.$$

We make use of the following result of Dias da Silva and Hamidoune [13, Theorem 4.1].

Lemma 2.12. *For a set $\mathcal{S} \subseteq \mathbb{F}_p$ and an integer $k \geq 1$, we have*

$$\#\Sigma_k(\mathcal{S}) \geq \min\{p, k\#\mathcal{S} - k^2 + 1\}.$$

We now define

$$\Sigma_*(\mathcal{S}) = \bigcup_{k=0}^{\#\mathcal{S}} \Sigma_k(\mathcal{S}).$$

Taking $k = \lfloor \#\mathcal{S}/2 \rfloor$ in Lemma 2.12 we immediately derive:

Corollary 2.13. *For a set $\mathcal{S} \subseteq \mathbb{F}_p$ and an integer $k \geq 1$ we have*

$$\#\Sigma_*(\mathcal{S}) \gg \min\{p, (\#\mathcal{S})^2 + 1\}.$$

3. Main results

3.1. **Squarefree integers with fixed digits.** Let $S_n(\mathbf{d})$ be the number of squarefree integers $s \in \mathcal{N}_n(\mathbf{d})$.

Theorem 3.1. *For any $\varepsilon > 0$, uniformly over integer $k < (2/5 - \varepsilon)n$ and $\mathbf{d} \in \mathcal{D}_{k,n}^*$, we have*

$$S_n(\mathbf{d}) = \left(\frac{8}{\pi^2} + o(1)\right) \#\mathcal{N}_n(\mathbf{d}).$$

Proof. The inclusion-exclusion principle yields

$$S_n(\mathbf{d}) = \sum_{q=1}^{\infty} \mu(q) \#\mathcal{N}_n(\mathbf{d}, q^2),$$

where $\mu(q)$ is the Möbius function; see [23, Section 16.3].

As before, we define $\kappa = k/n$ and we also use the function $\vartheta(\kappa, \varrho)$ that is given by (2.2).

For $\kappa < 2/5 - \varepsilon$ we have

$$\vartheta(\kappa, 2/5) > 1/2 \tag{3.1}$$

as $\vartheta(2/5, 2/5) = 1/2$ and for fixed ϱ, the function $\vartheta(\kappa, \varrho)$ given by (2.2) is obviously decreasing in κ. Choose $\zeta > 2/5$ such that

$$(3 + 4\varepsilon)\zeta \leq 2(1 - \kappa), \qquad \zeta(1 + 5\varepsilon) < 4(1 - \kappa) - 2 - \varepsilon, \tag{3.2}$$

which for sufficiently small $\varepsilon > 0$ is possible since $\kappa < 2/5 - \varepsilon$. Note that in particular, $\zeta > \kappa$.

We set

$$T = n^{1/20\kappa}, \qquad U = 2^{n/5}, \qquad V = 2^{n/4}, \qquad W = 2^{\zeta n},$$

and write

$$S_n(\mathbf{d}) = S_1 + S_2 + S_3 + S_4 + S_5, \tag{3.3}$$

where

$$S_1 = \sum_{q \le T} \mu(q) \# \mathcal{N}_n(\mathbf{d}, q^2),$$
$$S_2 = \sum_{T < q \le U} \mu(q) \# \mathcal{N}_n(\mathbf{d}, q^2),$$
$$S_3 = \sum_{U < q \le V} \mu(q) \# \mathcal{N}_n(\mathbf{d}, q^2),$$
$$S_4 = \sum_{V < q \le W} \mu(q) \# \mathcal{N}_n(\mathbf{d}, q^2),$$
$$S_5 = \sum_{q > W} \mu(q) \# \mathcal{N}_n(\mathbf{d}, q^2).$$

We use Lemma 2.5 for $q \le T$, getting the main term

$$\begin{aligned} S_1 &= \# \mathcal{N}_n(\mathbf{d}) \sum_{\substack{q \le T \\ q \text{ odd}}} \frac{\mu(q)}{q^2} + O\left(\# \mathcal{N}_n(\mathbf{d}) T 2^{-\sqrt{n}}\right) \\ &= \# \mathcal{N}_n(\mathbf{d}) \sum_{q \text{ odd}} \frac{\mu(q)}{q^2} + O\left(\# \mathcal{N}_n(\mathbf{d}) T 2^{-\sqrt{n}} + \# \mathcal{N}_n(\mathbf{d}) T^{-1}\right) \\ &= \# \mathcal{N}_n(\mathbf{d}) \prod_{\substack{\ell \ge 3 \\ \ell \text{ prime}}} \left(1 - \frac{1}{\ell^2}\right) + O\left(\# \mathcal{N}_n(\mathbf{d}) T 2^{-\sqrt{n}} + \# \mathcal{N}_n(\mathbf{d}) T^{-1}\right) \\ &= \frac{4}{3} \# \mathcal{N}_n(\mathbf{d}) \prod_{\ell \text{ prime}} \left(1 - \frac{1}{\ell^2}\right) + O\left(\# \mathcal{N}_n(\mathbf{d}) T 2^{-\sqrt{n}} + \# \mathcal{N}_n(\mathbf{d}) T^{-1}\right). \end{aligned}$$

So we now obtain the main term

$$S_1 = \left(\frac{8}{\pi^2} + o(1)\right) \# \mathcal{N}_n(\mathbf{d}); \tag{3.4}$$

see [23, Theorem 280].

To estimate S_2, we use Lemma 2.6. First we note that by Lemma 2.10, for $T < q \le U$, we have

$$\vartheta(\kappa, 2\varrho) \ge \vartheta(\kappa, 2/5),$$

where, in analogy to Lemma 2.6, ϱ is defined by $q = 2^{\varrho n}$. Hence in this range we have

$$\# \mathcal{N}_n(\mathbf{d}, q^2) \ll \# \mathcal{N}_n(\mathbf{d}) q^{-2\vartheta(\kappa, 2\varrho)} \ll \# \mathcal{N}_n(\mathbf{d}) q^{-2\vartheta(\kappa, 2/5)}.$$

Since by (3.1) we have $2\vartheta(\kappa, 2/5) > 1$, we now derive

$$S_2 \ll \# \mathcal{N}_n(\mathbf{d}) T^{1 - 2\vartheta(\kappa, 2/5)} = o\left(\# \mathcal{N}_n(\mathbf{d})\right). \tag{3.5}$$

For S_3 we use a similar argument as for S_2, this time with Lemma 2.9 instead of Lemma 2.6, noting that with $\kappa < 2/5$ and $\varrho \ge 1/5$ we obtain

$$\# \mathcal{N}_n(\mathbf{d}, q^2) \ll \# \mathcal{N}_n(\mathbf{d}) q^{2(-1 + \kappa/4\varrho)} \ll \# \mathcal{N}_n(\mathbf{d}) q^{-1-\delta}$$

for some sufficiently small $\delta > 0$, so again

$$S_3 = o\left(\# \mathcal{N}_n(\mathbf{d})\right). \tag{3.6}$$

To estimate S_4, we use Corollary 2.8, which by (3.2) applies with some sufficiently small $\varepsilon > 0$, getting

$$S_4 \ll \#\mathcal{N}_n(\mathbf{d}) V^{-\varepsilon/2} \log W = o\left(\#\mathcal{N}_n(\mathbf{d})\right). \tag{3.7}$$

Finally, we use the trivial bound $\#\mathcal{N}_n(\mathbf{d}, q^2) \le 2^n/q^2$ for $q > W$ and using $\zeta > \kappa$ we derive

$$S_5 \ll 2^n W^{-1} \ll \#\mathcal{N}_n(\mathbf{d}) 2^{\kappa n - \zeta n} = o\left(\#\mathcal{N}_n(\mathbf{d})\right). \tag{3.8}$$

Substituting (3.4), (3.5), (3.6), (3.7) and (3.8) into (3.3), we now conclude the proof. □

3.2. **Average values of the Euler function.** We now consider the average value

$$F_n(\mathbf{d}) = \sum_{s \in \mathcal{N}_n(\mathbf{d})} \frac{\varphi(s)}{s}$$

with the Euler function $\varphi(s)$; see [23, Section 16.3].

Theorem 3.2. *For any $\varepsilon > 0$, uniformly over integers $k < (1-\varepsilon)n$ and $\mathbf{d} \in \mathcal{D}^*_{k,n}$, we have*

$$F_n(\mathbf{d}) = \left(\frac{8}{\pi^2} + o(1)\right) \#\mathcal{N}_n(\mathbf{d}).$$

Proof. Using the the well-known formula

$$\frac{\varphi(s)}{s} = \sum_{q \mid s} \frac{\mu(q)}{q}$$

(see [23, equation (16.3.1)]) and changing the order of summation, we write

$$F_n(\mathbf{d}) = \sum_{q=1}^{\infty} \frac{\mu(q)}{q} \#\mathcal{N}_n(\mathbf{d}, q).$$

We now proceed very similarly to the proof of Theorem 3.1.

Again we define $\kappa = k/n$ and we also use the function $\vartheta(\kappa, \varrho)$ that is given by (2.2).

Clearly, for $0 < \kappa < 1$ we have

$$\vartheta(\kappa, 1) = 1 - \sqrt{\kappa} > 0.$$

Thus, using Lemma 2.10, we see that for any $\kappa < 1$ we can find ξ to satisfy

$$1 > \xi > \kappa \tag{3.9}$$

and

$$\vartheta(\kappa, \xi) > 0. \tag{3.10}$$

We set

$$Q = n^{1/10\kappa} \qquad \text{and} \qquad W = 2^{\xi n}$$

and write

$$F_n(\mathbf{d}) = T_1 + T_2 + T_3, \tag{3.11}$$

where

$$T_1 = \sum_{q \le Q} \frac{\mu(q)}{q} \#\mathcal{N}_n(\mathbf{d}, q),$$

$$T_2 = \sum_{Q < q \le W} \frac{\mu(q)}{q} \#\mathcal{N}_n(\mathbf{d}, q),$$

$$T_3 = \sum_{q > W} \frac{\mu(q)}{q} \#\mathcal{N}_n(\mathbf{d}, q).$$

We use Lemma 2.5 for $q \le Q$, and exactly as in the proof of Theorem 3.1 we obtain the main term

$$T_1 = \#\mathcal{N}_n(\mathbf{d}) \sum_{\substack{q \le Q \\ q \text{ odd}}} \frac{\mu(q)}{q^2} + O\left(\#\mathcal{N}_n(\mathbf{d}) Q 2^{-\sqrt{n}}\right) = \left(\frac{8}{\pi^2} + o(1)\right) \#\mathcal{N}_n(\mathbf{d}); \tag{3.12}$$

see [23, Theorem 280].

To estimate T_2, we use Lemma 2.6 for $Q < q \le W$. First we note that by Lemma 2.10, for $Q < q \le W$, we have

$$\vartheta(\kappa, \varrho) \ge \vartheta(\kappa, \xi),$$

where, in analogy to Lemma 2.6, ϱ is defined by $q = 2^{\varrho n}$. Hence in this range we have

$$\#\mathcal{N}_n(\mathbf{d}, q) \ll \#\mathcal{N}_n(\mathbf{d}) q^{-\vartheta(\kappa,\varrho)} \ll \#\mathcal{N}_n(\mathbf{d}) q^{-\vartheta(\kappa,\xi)}.$$

Since by (3.10) we have $\vartheta(\kappa, \xi) > 0$, we now derive

$$T_2 \ll \#\mathcal{N}_n(\mathbf{d}) Q^{-\vartheta(\kappa,\xi)} = o\left(\#\mathcal{N}_n(\mathbf{d})\right). \tag{3.13}$$

Finally, we use the trivial bound $\#\mathcal{N}_n(\mathbf{d}, q) \le 2^n/q$ for $q > W$ and using (3.9) derive

$$T_3 \ll 2^n W^{-1} \ll \#\mathcal{N}_n(\mathbf{d}) 2^{\kappa n - \xi n} = o\left(\#\mathcal{N}_n(\mathbf{d})\right). \tag{3.14}$$

Substituting (3.12), (3.13) and (3.14) into (3.11), we conclude the proof. □

3.3. **Non-residues with fixed digits.** For a prime p, we use $\mathcal{N}_n^+(\mathbf{d}, p)$ and $\mathcal{N}_n^-(\mathbf{d}, p)$ to denote the sets of $s \in \mathcal{N}_n(\mathbf{d})$ that are quadratic residues and non-residues, respectively (we also use $\mathcal{N}_n^\pm(\mathbf{d}, p)$ to denote either of these sets).

Theorem 3.3. *For any $\varepsilon > 0$ there exists some $\delta > 0$ such that for $k < (1/2 - \varepsilon)n$, $\mathbf{d} \in \mathcal{D}_{k,n}$ and any prime p with $2^n < p < 2^{n+1}$ we have*

$$\#\mathcal{N}_n^\pm(\mathbf{d}, p) = \left(\frac{1}{2} + O(p^{-\delta})\right) \#\mathcal{N}_n(\mathbf{d}).$$

Proof. We select arbitrary $s = \lceil \varepsilon n/2 \rceil$ free positions and denote by $\mathcal{B}$ the set of 2^s integers with all possible combinations of digits on these positions and zeros on all other positions. We also define by $\mathcal{A}$ the subset of 2^{n-k-s} elements of $\mathcal{N}_n(\mathbf{d})$ which also have zero digits on the positions that are allocated to $\mathcal{B}$. Clearly each element of $\mathcal{N}_n(\mathbf{d})$ has a unique representation as $a + b$ with $a \in \mathcal{A}$, $b \in \mathcal{B}$. The result is now instant from Corollary 2.3. □

3.4. **Primitive roots in Hilbert cubes.** We now present an improvement of [14, Theorem 1.3].

Theorem 3.4. *We have*

$$F(p) \le p^{3/19+o(1)}.$$

Proof. Let $\mathcal{H}(a_0; a_1, \dots, a_d)$ be a Hilbert cube, with d distinct base elements $a_0, \dots, a_d \in \mathbb{F}_p$. Suppose that $\mathcal{H}(a_0; a_1, \dots, a_d)$ does not contain primitive roots modulo p. We show that $d = O(p^{3/19+o(1)})$.

As in the proof of [14, Theorem 1.3], we fix some $\varepsilon > 0$ and discarding some extra elements we assume that

$$d = 2\lceil 0.5p^{3/19+\varepsilon}\rceil \tag{3.15}$$

(as it is convenient to assume that d is even).

Let $\mathcal{U} = \mathcal{H}(a_0; a_1, \dots, a_{d/2})$ and $\mathcal{V} = \mathcal{H}(0; a_{d/2+1}, \dots, a_d)$, both $\mathcal{U}$ and $\mathcal{V}$ understood as subsets of $\mathbb{F}_p$. It follows from Lemma 2.11, applied with $D = 7$, that $\mathcal{U}$ contains an arithmetic progression $\mathcal{A} \subseteq \mathbb{F}_p$ of length

$$L = \#\mathcal{A} \ge p^{7/38+o(1)}. \tag{3.16}$$

Let Δ be the difference between consecutive terms of the progression. Let us consider the interval

$$\mathcal{I} = \{\Delta^{-1}a \ : \ a \in \mathcal{A}\} \subseteq \mathbb{F}_p$$

of L consecutive residues modulo p.

On the other hand, it follows from Corollary 2.13 that $\#\mathcal{V} \gg d^2$. Now let

$$\mathcal{W} = \{\Delta^{-1}a \ : \ a \in \mathcal{V}\} \subseteq \mathbb{F}_p.$$

We now take an arbitrary element $w_1 \in \mathcal{W}$ and remove from $\mathcal{W}$ at most $O(L)$ elements w with $|w - w_1| \le L$ and denote the remaining set as $\mathcal{W}_1$. We now choose an arbitrary element $w_2 \in \mathcal{W}_1$ and remove from $\mathcal{W}_1$ at most $O(L)$ elements v with $|w - w_2| \le L$ and denote the remaining set as $\mathcal{W}_2$. Continuing, we obtain a set $\{w_1, \dots, w_J\}$ of

$$J \gg \#\mathcal{W}/L = \#\mathcal{V}/L \gg d^2/L \tag{3.17}$$

elements, which after renumbering satisfy the condition of Lemma 2.4.

By Lemma 2.4 and (3.17), taking a sufficiently large ν after simple calculations we obtain that for any non-trivial multiplicative character χ of $\mathbb{F}_p$ we have

$$\sum_{j=1}^{J-1}\left|\sum_{i=1}^{L}\chi\left(i+w_j\right)\right|^{2\nu} \ll p^{1/2+1/2\nu+o(1)}L^{2\nu-2} \ll JL^{2\nu}p^{-\eta},$$

provided that

$$L^{1/2}d \gg p^{1/4+\varepsilon}.$$

Recalling (3.15) and (3.16), we find that the latter condition is satisfied. Expressing, in a standard fashion, the counting function for primitive roots among the elements $i+w_j$, $i = 1, \dots, L$, $j = 1, \dots, J$, via multiplicative characters (see, for example, [14, Lemma 2.4]), we see that it is positive, provided p is large enough. Since ε is arbitrary, we obtain the desired result. □

3.5. **Elements with restricted digits in finite fields.** Our next result improves [11, Theorem 2.1] for any $n \ge 3$.

For $\mathcal{W}_{\mathfrak{A}}$, given by (1.2), we use $\mathcal{W}_{\mathfrak{A}}^+$ and $\mathcal{W}_{\mathfrak{A}}^-$ to denote the sets of $w \in \mathcal{W}_{\mathfrak{A}}$ that are quadratic residues and non-residues, respectively (we also use $\mathcal{W}_{\mathfrak{A}}^{\pm}$ to denote either of these sets).

Theorem 3.5. *Let $n \ge 2$. For any $\varepsilon > 0$ there is some $\delta > 0$ such that for an arbitrary basis $\omega_1, \ldots, \omega_n$ of $\mathbb{F}_{p^n}$ over $\mathbb{F}_p$ and a collection of n sets*

$$\mathfrak{A} = \{\mathcal{A}_i \subseteq \mathbb{F}_p \ : \ i = 1, \ldots, n\},$$

satisfying

$$\prod_{1 \le i \le n} \#\mathcal{A}_i \ge p^{(1/2+\varepsilon)n^2/(n-1)} \tag{3.18}$$

and

$$\min_{1 \le i \le n} \#\mathcal{A}_i \ge p^{\varepsilon}, \tag{3.19}$$

the following holds: for the set

$$\mathcal{W}_{\mathfrak{A}} = \{a_1\omega_1 + \ldots + a_n\omega_n \ : \ a_i \in \mathcal{A}_i, \ i = 1, \ldots, n\},$$

we have

$$\#\mathcal{W}_{\mathfrak{A}}^{\pm} = \left(\frac{1}{2} + O(p^{-\delta})\right) \#\mathcal{W}_{\mathfrak{A}}$$

uniformly over n and p.

Proof. Without loss of generality, we can assume that $\varepsilon \le 1/2$. We also set

$$n_0(\varepsilon) = \left\lceil 4\varepsilon^{-1} \right\rceil.$$

We first consider the case when

$$n > n_0(\varepsilon). \tag{3.20}$$

Assuming that (3.20) holds, we set

$$m = \left\lceil \frac{1+\varepsilon}{1+2\varepsilon} n \right\rceil.$$

By choosing $\mathcal{I}$ such that the sets $\mathcal{A}_i$ with $i \in \mathcal{I}$ are the m sets of largest cardinality we see that

$$\prod_{i \in \mathcal{I}} \#\mathcal{A}_i \ge \left(\prod_{1 \le i \le n} \#\mathcal{A}_i \right)^{m/n}. \tag{3.21}$$

We then define $\mathcal{J} = \{1, \ldots, n\} \setminus \mathcal{I}$ and

$$\mathcal{A} = \left\{ \sum_{i \in \mathcal{I}} a_i\omega_i \ : \ a_i \in \mathcal{A}_i, \ i \in \mathcal{I} \right\},$$

$$\mathcal{B} = \left\{ \sum_{i \in \mathcal{J}} a_i\omega_i \ : \ a_i \in \mathcal{A}_i, \ i \in \mathcal{J} \right\}.$$

Thus, recalling (3.18) and (3.21), we see that

$$\#\mathcal{A} \ge \left(p^{(1/2+\varepsilon)n^2/(n-1)}\right)^{m/n} \ge \left(p^{(1/2+\varepsilon)n}\right)^{m/n} \ge p^{(1/2+\varepsilon)m} \ge p^{(1/2+\varepsilon/2)n}. \tag{3.22}$$

Furthermore for $n > n_0(\varepsilon)$ we have

$$\#\mathcal{J} = n - m \geq n - \frac{1+\varepsilon}{1+2\varepsilon}n - 1 = n - \frac{(1+\varepsilon)n + 1 + 2\varepsilon}{1+2\varepsilon}$$
$$\geq n - \frac{(1+\varepsilon)n+2}{1+2\varepsilon} = \frac{\varepsilon n - 2}{1+2\varepsilon} \geq \frac{\varepsilon n - 2}{2} \geq \varepsilon n/4.$$

Therefore, recalling (3.19) again, we see that

$$\#\mathcal{B} \geq p^{\varepsilon^2 n/4}. \tag{3.23}$$

As $\{\omega_1, \ldots, \omega_n\}$ is a basis, every element $w = \mathcal{W}_{\mathfrak{A}}$ has a unique representation $w = a + b$ with $a \in \mathcal{A}$, $b \in \mathcal{B}$. Hence for any multiplicative character χ of $\mathbb{F}_r$ we have

$$\sum_{w \in \mathcal{W}_{\mathfrak{A}}} \chi(w) = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} \chi(a+b).$$

Now using the bounds (3.22) and applying (3.23) and applying Corollary 2.3 with $\eta = \varepsilon^2/4$, we obtain

$$\sum_{w \in \mathcal{W}_{\mathfrak{A}}} \chi(w) \ll \#\mathcal{A}\#\mathcal{B}p^{-\delta} = \#\mathcal{W}_{\mathfrak{A}}p^{-\delta} \tag{3.24}$$

for any non-trivial multiplicative character χ of $\mathbb{F}_{p^n}$ where $\delta > 0$ depends only on ε. Thus, taking the quadratic character χ, we obtain the desired result if the inequality (3.20) holds.

Now, for small values of n, for which inequality (3.20) fails, we simply take $m = n - 1$ and choose $\mathcal{I}$ as before to satisfy (3.21). Hence, instead of (3.22) we have

$$\#\mathcal{A} \geq p^{(1/2+\varepsilon)n}$$

and also trivially, we have

$$\#\mathcal{B} \geq p^{\varepsilon} \geq p^{\varepsilon n/n_0(\varepsilon)}.$$

Applying Corollary 2.3 with $\eta = \varepsilon/n_0(\varepsilon)$, we obtain the bound (3.24) again, which concludes the proof. □

4. Comments

We remark that motivated by a question of Erdős, Mauduit and Sárközy [18, Problem 5], Banks and Shparlinski [2] have studied the average value of $\varphi(s)/s$ over integers with some digital restrictions, different from that of Theorem 3.2; see also [3]. As in [3], one can also study the average value of $\sigma(s)/s$ for the sums of divisors function and obtain a full analogue of Theorem 3.2 for this function. Furthermore, our argument can be used to give an asymptotic formula for the number of pairs (s, r) with $r, s \in \mathcal{N}_n(\mathbf{d})$ such that $\gcd(s, r) = 1$.

Clearly, the bound (3.24) also allows us to study the distribution of primitive roots in the set $\mathcal{W}_{\mathfrak{A}}$. Finally, we note that the case when the sets $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are sets of consecutive residues corresponds to the settings of [11, Theorem 3.1]. In this case, one can use the recent generalisations of the Burgess bounds that are due to Chang [7, 8] and Konyagin [30] (when n is bounded) together with a classical bound of Davenport and Lewis [12] (when $n \to \infty$) and improve the result of [11, Theorem 3.1].

Acknowledgements

The authors would like to thank Sergei Konyagin for helpful discussions and in particular for his suggestion to use the result of Shao [35] in the proof of Theorem 3.4.

The authors are also very grateful to Cécile Dartyge and to the referee for careful reading of the manuscript and useful suggestions.

Most of this work originated at the Erwin Schrödinger International Institute for Mathematical Physics, Vienna, and then was continued during a very enjoyable stay of the authors at the CIRM, Luminy. The second author was partially supported by the FWF (Austria) grant W1230. The third author was partially supported by the CIRM-CNRS-SMF (France) as a Jean-Morlet Chair and ARC (Australia) grant DP140100118.

References

[1] William D. Banks, Alessandro Conflitti, and Igor E. Shparlinski, *Character sums over integers with restricted g-ary digits*, Illinois J. Math. **46** (2002), no. 3, 819–836. MR1951242

[2] William D. Banks and Igor E. Shparlinski, *Arithmetic properties of numbers with restricted digits*, Acta Arith. **112** (2004), no. 4, 313–332, DOI 10.4064/aa112-4-1. MR2046944

[3] William D. Banks and Igor E. Shparlinski, *Average value of the Euler function on binary palindromes*, Bull. Pol. Acad. Sci. Math. **54** (2006), no. 2, 95–101, DOI 10.4064/ba54-2-1. MR2266140

[4] J. Bourgain, *Estimates on exponential sums related to the Diffie-Hellman distributions*, Geom. Funct. Anal. **15** (2005), no. 1, 1–34, DOI 10.1007/s00039-005-0500-4. MR2140627

[5] Jean Bourgain, *Prescribing the binary digits of primes*, Israel J. Math. **194** (2013), no. 2, 935–955, DOI 10.1007/s11856-012-0104-2. MR3047097

[6] Jean Bourgain, *Prescribing the binary digits of primes, II*, Israel J. Math. **206** (2015), no. 1, 165–182, DOI 10.1007/s11856-014-1129-5. MR3319636

[7] Mei-Chu Chang, *On a question of Davenport and Lewis and new character sum bounds in finite fields*, Duke Math. J. **145** (2008), no. 3, 409–442, DOI 10.1215/00127094-2008-056. MR2462111

[8] Mei-Chu Chang, *Burgess inequality in* $\mathbb{F}_{p^2}$, Geom. Funct. Anal. **19** (2009), no. 4, 1001–1016, DOI 10.1007/s00039-009-0031-5. MR2570312

[9] Sylvain Col, *Diviseurs des nombres ellipséphiques* (French, with French summary), Period. Math. Hungar. **58** (2009), no. 1, 1–23, DOI 10.1007/s10998-009-9001-9. MR2487242

[10] Sylvain Col, *Palindromes dans les progressions arithmétiques* (French), Acta Arith. **137** (2009), no. 1, 1–41, DOI 10.4064/aa137-1-1. MR2481980

[11] Cécile Dartyge, Christian Mauduit, and András Sárközy, *Polynomial values and generators with missing digits in finite fields*, Funct. Approx. Comment. Math. **52** (2015), no. 1, 65–74, DOI 10.7169/facm/2015.52.1.5. MR3326124

[12] H. Davenport and D. J. Lewis, *Character sums and primitive roots in finite fields*, Rend. Circ. Mat. Palermo (2) **12** (1963), 129–136. MR0167482

[13] J. A. Dias da Silva and Y. O. Hamidoune, *Cyclic spaces for Grassmann derivatives and additive theory*, Bull. London Math. Soc. **26** (1994), no. 2, 140–146, DOI 10.1112/blms/26.2.140. MR1272299

[14] Rainer Dietmann, Christian Elsholtz, and Igor E. Shparlinski, *On gaps between primitive roots in the Hamming metric*, Q. J. Math. **64** (2013), no. 4, 1043–1055, DOI 10.1093/qmath/has022. MR3151603

[15] Rainer Dietmann, Christian Elsholtz, and Igor E. Shparlinski, *On gaps between quadratic non-residues in the Euclidean and Hamming metrics*, Indag. Math. (N.S.) **24** (2013), no. 4, 930–938, DOI 10.1016/j.indag.2013.02.005. MR3124809

[16] Michael Drmota and Christian Mauduit, *Weyl sums over integers with affine digit restrictions*, J. Number Theory **130** (2010), no. 11, 2404–2427, DOI 10.1016/j.jnt.2010.04.006. MR2678855

[17] Michael Drmota, Christian Mauduit, and Joël Rivat, *Primes with an average sum of digits*, Compos. Math. **145** (2009), no. 2, 271–292, DOI 10.1112/S0010437X08003898. MR2501419

[18] Paul Erdős, Christian Mauduit, and András Sárközy, *On arithmetic properties of integers with missing digits. II. Prime factors*, Paul Erdős memorial collection, Discrete Math. **200** (1999), no. 1-3, 149–164, DOI 10.1016/S0012-365X(98)00331-8. MR1692287

[19] M. Gabdullin, *On squares in special sets of finite fields* (in Russian), Chebyshevskii Sbornik **17** (2016), no. 2, 56–63.

[20] M. R. Gabdullin, *On the squares in the set of elements of a finite field with constraints on the coefficients of its basis expansion* (Russian, with Russian summary), Mat. Zametki **100** (2016), no. 6, 807–824, DOI 10.4213/mzm11091. MR3588906

[21] Sidney W. Graham and Igor E. Shparlinski, *On RSA moduli with almost half of the bits prescribed*, Discrete Appl. Math. **156** (2008), no. 16, 3150–3154, DOI 10.1016/j.dam.2007.12.012. MR2462121

[22] Martin Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*, 2nd ed., Algorithms and Combinatorics, vol. 2, Springer-Verlag, Berlin, 1993. MR1261419

[23] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 5th ed., The Clarendon Press, Oxford University Press, New York, 1979. MR568909

[24] Glyn Harman and Imre Kátai, *Primes with preassigned digits. II*, Acta Arith. **133** (2008), no. 2, 171–184, DOI 10.4064/aa133-2-5. MR2417463

[25] N. Hegyvári and A. Sárközy, *On Hilbert cubes in certain sets*, Ramanujan J. **3** (1999), no. 3, 303–314, DOI 10.1023/A:1009883404485. MR1714944

[26] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR2061214

[27] A. A. Karatsuba, *Distribution of values of Dirichlet characters on additive sequences* (Russian), Dokl. Akad. Nauk SSSR **319** (1991), no. 3, 543–545; English transl., Soviet Math. Dokl. **44** (1992), no. 1, 145–148. MR1148968

[28] Anatolij A. Karatsuba, *Basic analytic number theory*, Springer-Verlag, Berlin, 1993. Translated from the second (1983) Russian edition and with a preface by Melvyn B. Nathanson. MR1215269

[29] Sergei Konyagin, *Arithmetic properties of integers with missing digits: distribution in residue classes*, Period. Math. Hungar. **42** (2001), no. 1-2, 145–162, DOI 10.1023/A:1015256809636. MR1832701

[30] S. V. Konyagin, *Estimates for character sums in finite fields* (Russian, with Russian summary), Mat. Zametki **88** (2010), no. 4, 529–542, DOI 10.1134/S0001434610090221; English transl., Math. Notes **88** (2010), no. 3-4, 503–515. MR2882215

[31] Sergei Konyagin, Christian Mauduit, and András Sárközy, *On the number of prime factors of integers characterized by digit properties*, Period. Math. Hungar. **40** (2000), no. 1, 37–52, DOI 10.1023/A:1004887821978. MR1774933

[32] Florian Luca, *Arithmetic properties of positive integers with fixed digit sum*, Rev. Mat. Iberoam. **22** (2006), no. 2, 369–412, DOI 10.4171/RMI/461. MR2294785

[33] Christian Mauduit and Joël Rivat, *Sur un problème de Gelfond: la somme des chiffres des nombres premiers* (French, with English and French summaries), Ann. of Math. (2) **171** (2010), no. 3, 1591–1646, DOI 10.4007/annals.2010.171.1591. MR2680394

[34] Christian Mauduit and Zaid Shawket, *Sommes d'exponentielles associées aux fonctions digitales restreintes* (French, with English summary), Unif. Distrib. Theory **7** (2012), no. 1, 105–133. MR2943163

[35] N. G. Moshchevitin and I. D. Shkredov, *On the multiplicative properties modulo m of numbers with missing digits* (Russian, with Russian summary), Mat. Zametki **81** (2007), no. 3, 385–404, DOI 10.1134/S000143460703008X; English transl., Math. Notes **81** (2007), no. 3-4, 338–355. MR2333944

[36] Alina Ostafe and Igor E. Shparlinski, *Multiplicative character sums and products of sparse integers in residue classes*, Period. Math. Hungar. **64** (2012), no. 2, 247–255, DOI 10.1007/s10998-012-6771-2. MR2925171

[37] Xuancheng Shao, *Character sums over unions of intervals*, Forum Math. **27** (2015), no. 5, 3017–3026, DOI 10.1515/forum-2013-0080. MR3393387

[38] Tomasz Schoen, *Arithmetic progressions in sums of subsets of sparse sets*, Acta Arith. **147** (2011), no. 3, 283–289, DOI 10.4064/aa147-3-7. MR2773206
[39] Igor E. Shparlinski, *Prime divisors of sparse integers*, Period. Math. Hungar. **46** (2003), no. 2, 215–222, DOI 10.1023/A:1025996312037. MR2004674
[40] Igor E. Shparlinski, *On RSA moduli with prescribed bit patterns*, Des. Codes Cryptogr. **39** (2006), no. 1, 113–122, DOI 10.1007/s10623-005-3137-2. MR2201387
[41] Igor E. Shparlinski, *Exponential sums and prime divisors of sparse integers*, Period. Math. Hungar. **57** (2008), no. 1, 93–99, DOI 10.1007/s10998-008-7093-3. MR2448400

Department of Mathematics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom
E-mail address: rainer.dietmann@rhul.ac.uk

Institute of Analysis and Number Theory, Graz University of Technology, Kopernikusgasse 24/II, A-8010 Graz, Austria
E-mail address: elsholtz@math.tugraz.at

Department of Pure Mathematics, University of New South Wales, Sydney, New South Wales 2052, Australia
E-mail address: igor.shparlinski@unsw.edu.au