Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Low-energy decomposition results over finite fields

A. Mohammadi, S. Stevens

RICAM-Report 2021-20

LOW-ENERGY DECOMPOSITION RESULTS OVER FINITE FIELDS

ALI MOHAMMADI AND SOPHIE STEVENS

Abstract. We prove various low-energy decomposition results, showing that we can decompose a finite set $A \subset \mathbb{F}_p$ satisfying $|A| < p^{5/8}$, into $A = S \sqcup T$ so that, for a non-degenerate quadratic $f \in \mathbb{F}_p[x, y]$, we have

$$|\{(s_1, s_2, s_3, s_4) \in S^4 : s_1 + s_2 = s_3 + s_4\}| \ll |A|^{3-\frac{1}{5}+\epsilon}$$

and

$$|\{(t_1, t_2, t_3, t_4) \in T^4 : f(t_1, t_2) = f(t_3, t_4)\}| \ll |A|^{3-\frac{1}{5}+\epsilon}.$$

Variations include extending this result to large A and a low-energy decomposition involving additive energy of images of rational functions. This gives a quantitative improvement to a result of Roche-Newton, Shparlinski and Winterhof [23] as well as a generalisation of a result of Rudnev, Shkredov and Stevens [26].

We consider applications to conditional expanders, exponential sum estimates and the finite field Littlewood problem. In particular, we improve results of Mirzaei [16], Swaenepoel and Winterhof [31] and Garcia [8].

1. Introduction

In this paper, we show results of the following flavour: for a suitable non-degenerate polynomial $f \in \mathbb{F}_p[x, y]$, $A \subseteq \mathbb{F}_p$ admits a decomposition into disjoint sets $A = S \sqcup T$ so that

$$\mathsf{E}(S) := |\{(s_1, s_2, s_3, s_4) \in S^4 : s_1 + s_2 = s_3 + s_4\}| \ll |A|^{3-\frac{1}{5}+\epsilon}$$

and

$$\mathsf{E}_f(T) := |\{(t_1, t_2, t_3, t_4) \in T^4 : f(t_1, t_2) = f(t_3, t_4)\}| \ll |A|^{3-\frac{1}{5}+\epsilon}.$$

The actual formulation of our results is somewhat technical, and so we defer their formal presentation until Section 2, after we have developed the necessary terminology and context of these results. We give an overview of our techniques in Section 2.2.

1.1. Background.

The sum-product problem over finite fields is a quantitative interpretation of the observation that a set $A \subseteq \mathbb{F}_q$ (where q is some prime power) cannot be both additive and multiplicative, unless the intersection of A with a multiplicative coset of a subfield of $\mathbb{F}_q$ is large. More precisely, the Erdős-Szemerédi conjecture (over $\mathbb{F}_q$) asks if, for every $0 < \epsilon < 1$, the inequality

$$\max\{|A + A|, |AA|\} \gg_\epsilon |A|^{1+\epsilon} \tag{1}$$

holds under some natural conditions on sets $A \subset \mathbb{F}_q$.

Here, we denote by $A+A$ the sum set of A: given $A, B \subseteq \mathbb{F}_q$, we define

$$A+B = \{a+b : (a,b) \in A \times B\}.$$

We similarly define $A-B$, AB and A/B, which we refer to as the difference, product and ratio sets of A and B respectively; we do not consider division by zero in the definition of A/B.

Note, for instance, that no $\epsilon > 0$ exists in the case that $A = cG$ for some subfield G of $\mathbb{F}_q$ and element $c \in \mathbb{F}_q$. The most studied instances of this problem involve either sets that are large in terms of the order of the field (e.g. $|A| > q^{1/2}$; see Garaev [7] for the state-of-the-art in this direction) or sets that are small in terms of the characteristic of the field (e.g. $|A| < p^{1/2}$; see the authors' companion paper [17] for the best results towards this problem). A work of Roche-Newton and Li [21] considers a less restrictive constraint, based on the size of the intersection of the set in question and multiplicative cosets of proper subfields of $\mathbb{F}_q$.

We describe A as e.g. 'additive' if its sum set is small, the canonical example being an arithmetic progression, where $|A+A| = 2|A|-1$. However this classification is fragile, for example, if A is the union of an arithmetic progression of size N, and a random set of N elements, then $|A+A| \sim |A|^2$, but it is clear that A does possess additive structure (in the sense that A contains a large arithmetic progression). A more robust characterisation of structure can be given via the energy of a set: for sets A and B we define the representation functions

$$r_{A\circ B}(\lambda) = |\{(a,b) \in A \times B : a \circ b = \lambda\}| \quad \text{for} \quad \circ \in \{+,-,\times,/\}. \tag{2}$$

The additive and multiplicative energies of A and B are moments of the representation functions: for $k \geq 1$ we define

$$\mathsf{E}_k(A,B) := \sum_{x \in A-B} r^k_{A-B}(x) \text{ and } \mathsf{E}^{\times}_k(A,B) := \sum_{x \in A/B} r^k_{A/B}(x).$$

We write $\mathsf{E}_k(A,A) = \mathsf{E}_k(A)$ and $\mathsf{E}(A,B) = \mathsf{E}_2(A,B)$; we do the same for multiplicative energy. . When $k \in \mathbb{N}$, the energy has a combinatorial interpretation as the number of solutions to (in the case of additive energy) the equation $a_1 - b_1 = \cdots = a_k - b_k$ where $a_i \in A$ and $b_i \in B$. The case $k=2$ is special because we can replace "+" with "−" in the above. Hence, through a simple application of the Cauchy-Schwarz inequality, we have

$$\mathsf{E}(A,B)|A \pm B| \geq |A|^2|B|^2. \tag{3}$$

These inequalities also hold for their multiplicative counterparts.

1.2. **Energy formulation of the sum product problem.** Balog and Wooley [1] raised the question of whether, in line with the sum-product problem, a similar duality statement exists for the energies of a given set, i.e. additive and multiplicative structure cannot coexist in a set, as measured by the energy. In [1, Theorem 1.3], the authors prove that for $A \subset \mathbb{F}_p$, there exist disjoint subsets $B, C \subseteq A$ such that $A = B \cup C$ and

$$\max\{\mathsf{E}(B), \mathsf{E}^{\times}(C)\} \lesssim |A|^{3-\frac{4}{101}} + |A|^{3+\frac{1}{15}}p^{-\frac{1}{15}}. \tag{4}$$

This decomposition formulation is necessary: for example, if A is the union of an arithmetic progression and a geometric progression of the same size, then both of its energies are essentially maximally

large. Since $\max\{|B|,|C|\} \geq |A|/2$, the Cauchy-Schwarz estimate (3) (or its multiplicative counterpart) converts a statement of the form (4) into a sum-product inequality. A significant quantitative improvement to (4) was obtained by Rudnev, Shkredov and Stevens [26, Theorem 2.8], who proved that when $|A| \leq p^{5/8}$, one can take $|A|^{3-1/5}$ as the upper bound.

1.3. **Energy formulation of sums and function images problem.** A variation of the sum-product problem, considered for example by Bukh and Tsimerman [4], Solymosi [29] and Cilleruelo et al. [6], is to establish a non-trivial lower bound for the quantity $\max\{|A+A|, |f(A)+f(A)|\}$ under some natural conditions on A and some function f. The energy-variant of this problem was considered by Roche-Newton, Shparlinksi and Winterhof [23], which we now describe.

Given a rational function $f \not\equiv 0$ in $\mathbb{F}_q(x)$, we write $f = g/h$, for coprime polynomials g and $h \not\equiv 0$ in $\mathbb{F}_q[x]$ and define the degree of f to be $d = \max\{\deg(g), \deg(h)\}$. We say that a rational function f in $\mathbb{F}_q(x)$ of degree d is *non-degenerate* if

$$f(x) \notin \{(a(g(x)^p - g(x)) + bx + c : g(x) \in \mathbb{F}_q(x); a, b, c \in \mathbb{F}_q\}. \tag{5}$$

Roche-Newton et al. [23, Theorem 1.1] show that for any set $A \subseteq \mathbb{F}_q$ and degree d rational function f in $\mathbb{F}_q(x)$, there exists a decomposition of A into disjoint subsets $S, T \subseteq A$ such that

$$\max\{\mathsf{E}(S), \mathsf{E}(f(T))\} \ll_d \frac{|A|^3}{M_A},$$

where

$$M_A = \min\left\{\frac{q^{1/2}}{|A|^{1/2}(\log|A|)^{11/4}}, \frac{|A|^{4/5}}{q^{2/5}(\log|A|)^{31/10}}\right\}. \tag{6}$$

This estimate is non-trivial when $|A| \gtrsim q^{1/2}$; no similar results are known for smaller sets. Macourt [15] has studied variations of [23, Theorem 1.1]. The non-degeneracy condition on f ensures that it is not a linearised permutation polynomial (see [23, Section 1.3]).

1.4. **Function energy formulation.** Both the sum-product problem and its function-image variant discussed above are instances of the more general problem of the growth of $|A+A| + |f(A,A)|$ for some bivariate function f. See for example [4], [12], [16] and [34] for results in this direction. Note that, if $f = g(ax+by)$ for some polynomial g and A is an arithmetic progression, then $|A+A| \approx |f(A,A)| \approx |A|$. Hence, we require the following notion of non-degeneracy.

Definition 1. *We say a polynomial $f \in \mathbb{F}_q[x,y]$ is non-degenerate if it depends on each variable and is not of the form $g(ax+by)$, for some univariate polynomial g with coefficients in $\mathbb{F}_q$.*

In this paper, we consider the energy formulation for this more general problem.

Given a bivariate polynomial $f \in \mathbb{F}_q[x,y]$, define

$$\mathsf{E}_f(A,B) := |\{(a_1,a_2,b_1,b_2) \in A^2 \times B^2 : f(a_1,b_1) = f(a_2,b_2)\}|. \tag{7}$$

For $\lambda \in \mathbb{F}_q$, let

$$r_{f(A,B)}(\lambda) = |\{(a,b) \in A \times B : f(a,b) = \lambda\}|.$$

We record the following identities and consequence of the Cauchy-Schwarz inequality:

$$\sum_{\lambda\in\mathbb{F}_q} r_{f(A,B)}(\lambda) = |A||B| \ , \quad \sum_{\lambda\in\mathbb{F}_q} r_{f(A,B)}(\lambda)^2 = \mathsf{E}_f(A,B) \ , \quad \mathsf{E}_f(A,B)|f(A,B)| \geq |A|^2|B|^2.$$

Roche-Newton et al. [23, Question 5.1] ask if one can prove the existence of a decomposition $A = S \cup T$, such that

$$\max\{\mathsf{E}_f(S), \mathsf{E}_g(T)\} \ll |A|^{3-\delta}, \tag{8}$$

for some $\delta > 0$, where $A \subset \mathbb{F}_q$ and $f, g \in \mathbb{F}_q[x,y]$ satisfy some natural non-degeneracy conditions. In this paper, incorporating some ideas from [12], we prove results of this nature.

1.5. **Applications of low-energy decomposition results.** Progress on low-energy decomposition results has led to direct improvements to the Erdős-Szemerédi sum-product problem (1). Due to the increased level of attention that this problem attracts, we have recorded the improvements to the finite field sum-product problem due to the techniques here in a companion paper [17].

The (bivariate) function variation of low-energy decomposition results similarly leads to progress on the bivariate function variation of the sum-product problem. Current progress on this problem is summarised by a result of Mirzaei [16] who shows that $A \subseteq \mathbb{F}_p$ with $|A| < p^{1/2}$ satisfies

$$\max\{|A \pm A|, |f(A,A)|\} \gtrsim |A|^{\frac{6}{5}+\frac{4}{305}} .$$

Using low-energy decomposition statements, we improve the exponent in the case of difference sets. We expect these techniques to lead to further progress on related problems in the expanders literature, for example, the unconditional growth of the images of sets under specific polynomials.

As observed by Balog and Wooley [1] as well as in the later works [23, 31], low-energy decomposition results provide useful tools in showing cancellation amongst various types of character sums.

Let $q = p^n$ and define

$$e_p(x) := \exp(2\pi i x/p) \quad \text{and} \quad \psi(x) = e_p(\mathsf{Tr}(x)) ,$$

where $\mathsf{Tr}(x) = x + x^p + \cdots + x^{p^{n-1}}$ is the trace of $x \in \mathbb{F}_q$ over $\mathbb{F}_p$.

Given sets $S, T \subseteq \mathbb{F}_q$, Vinogradov (see [33, p. 92]) showed that

$$\left|\sum_{s\in S}\sum_{t\in T} \psi(st)\right| \leq \sqrt{|S||T|q}. \tag{9}$$

This bound is non-trivial if $|S||T| > q^{1/2}$. There are many results that improve (9), either in terms of strength of the bound or its effective range (see e.g. [2, 3, 10]). Typically, this is achieved by considering sets S and T endowed with a particular structure. We focus on a recent application provided by Swaenepoel and Winterhof [31]. In [31, Theorem 1], the authors show the following: for a rational function $f \in \mathbb{F}_q(x)$ of degree d satisfying (5) so that $f(T) \subseteq T$, there exists $U \subseteq T$ with $|U| \geq |T|/(d+1)$ such that

$$\left|\sum_{s\in S}\sum_{u\in U} \psi(su)\right| \ll \left(\frac{|S|^3|T|^3 q}{M_T}\right)^{1/4}, \tag{10}$$

where M_T is defined by (6). See [31, p. 3] for a discussion of the strength of (10).

Following the work of Shkredov [27], we consider consequences of low-energy decomposition theorems to the finite field Littlewood problem of establishing non-trivial lower bounds on the l_1 norm of exponential sums over various sets. See for instance [8, 27] for a background on this problem, and also for results showing that the l_1 norm of exponential sums, over images of intervals under various functions, is large.

Finally, we mention that orbits of dynamical systems generated by functions $f \in \mathbb{F}_q(x)$ provide natural examples of sets $T \subset \mathbb{F}_q$, with $f(T) \subseteq T$. Namely, sets defined by

$$\mathrm{Orb}_f(u) = \{f^{(n)}(u) : n \geq 0\}, \tag{11}$$

where $u \in \mathbb{F}_q$, $f^{(0)}(u) = u$ and $f^{(n)}(u) = f(f^{(n-1)}(u))$ for $n \geq 1$. Various arithmetical properties of such dynamical systems have been investigated, in particular, in [5, 6, 9, 18, 22].

Notation. For $p \neq 2$ prime and $q = p^n$ for $n \in \mathbb{N}$, $\mathbb{F}_q$ denotes the finite field of order q and characteristic p. We write $\mathbb{F}_q^\times$ to denote the multiplicative group $\mathbb{F}_q \setminus \{0\}$.

We write $\alpha \ll \beta$ or $\beta \gg \alpha$ if there exists an absolute constant $c > 0$ such that $|\alpha| \leq c\beta$. If the constant c, depends on some parameter ϵ, then we write, for example, $\alpha \ll_\epsilon \beta$. If $\alpha \ll \beta$ and $\alpha \gg \beta$, we use $\alpha \approx \beta$. We also write $\alpha \lesssim \beta$ or $\beta \gtrsim \alpha$, if there exist $c_1, c_2 > 0$ such that $|\alpha| \leq c_1(\log \beta)^{c_2}\beta$. If $\alpha \lesssim \beta$ and $\beta \lesssim \alpha$, we write $\alpha \sim \beta$.
For disjoint sets S and T we denote their union as $S \sqcup T$; a decomposition of A into S and T exclusively refers to $A = S \sqcup T$ where S and T are disjoint.

2. Main results

2.1. Low energy decomposition.

Our first result is for 'small' sets A.

Theorem 1. *Let $A \subset \mathbb{F}_p$, with $|A| \leq p^{5/8}$ and let $f \in \mathbb{F}_p[x, y]$ denote a non-degenerate quadratic polynomial. There exist disjoint subsets $S, T \subseteq A$ such that $A = S \sqcup T$ and*

$$\max\{\mathsf{E}(S), \mathsf{E}_f(T)\} \lesssim |A|^{3-1/5}. \tag{12}$$

We remark that this result automatically extends to sets and polynomials over arbitrary fields $\mathbb{F}$. In this setting, A must satisfy the above size constraint in terms of p, the characteristic of $\mathbb{F}$; if the characteristic is zero, then there is no size constraint on A.

By following essentially the same arguments as in the proof of Theorem 1 but using [12, Lemma 5.2] in place of the forthcoming Lemma 3, we obtain the following analogue for 'large' subsets of $\mathbb{F}_q$.

Theorem 2. *Let $A \subseteq \mathbb{F}_q$ and let $f \in \mathbb{F}_q[x, y]$ denote a non-degenerate quadratic polynomial. There exist disjoint subsets $S, T \subseteq A$ such that $A = S \cup T$ and*

$$\max\{\mathsf{E}(S), \mathsf{E}_f(T)\} \lesssim \frac{|A|^3}{M_A},$$

where

$$M_A = \min\left\{\frac{q^{1/3}}{|A|^{1/3}}, \frac{|A|}{q^{1/2}}\right\}. \tag{13}$$

In the cases $f(x,y) = xy$ and $f(x,y) = g(x) + g(y)$, for certain rational functions $g \in \mathbb{F}_q(x)$, we obtain the following improvements.

Theorem 3. *Let $A \subseteq \mathbb{F}_q$. There exist disjoint subsets $S, T \subseteq A$ such that $A = S \cup T$ and*

$$\max\{\mathsf{E}(S), \mathsf{E}^{\times}(T)\} \ll \frac{|A|^3}{M_A},$$

where

$$M_A = \min\left\{ \frac{q^{1/2}}{|A|^{1/2}(\log|A|)^{1/2}}, \frac{|A|}{q^{1/2}(\log|A|)} \right\}. \tag{14}$$

The first term in M_A dominates when $q^2 < |A|^3(\log(|A|))^{-\frac{1}{2}}$.

We can use the Cauchy-Schwarz inequality (3) to convert Theorem 3 into a sum-product inequality. This reproduces, up to logarithmic factors Garaev's result [7, Theorem 1]. Through a construction provided in the same paper, this is optimal in the range $|A| \gtrsim q^{2/3}$.

Finally, we obtain a result on rational functions:

Theorem 4. *Let $A \subseteq \mathbb{F}_q$ and let $f \in \mathbb{F}_q(x)$ denote a rational function, with deg$(f) = d$, satisfying* (5)*. There exist disjoint subsets $S, T \subseteq A$ such that $A = S \cup T$ and*

$$\max\{\mathsf{E}(S), \mathsf{E}(f(T))\} \ll_d \frac{|A|^3}{M_A},$$

where M_A is given by (14)*.*

This improves the main result of Roche-Newton et al. [23, Theorem 1.1] for both minimands in M_A. Based on a construction in [23, Section 1.3], Theorem 4 is sharp up to constants in the range $|A| > q^{2/3}(\log|A|)^{1/6}$, whereas [23, Theorem 1.1] is sharp in a range of the form $|A| \gg q^{9/13}(\log|A|)^{7/26}$.

2.2. **Techniques.** The proofs of our main results all follow a similar scheme, present in the original paper of Balog and Wooley [1]. To prove the decomposition results, we prove an auxiliary 'mixed energy' result, extracting from A subsets $C \subseteq B$ with $|C| \gtrsim |B| \gg |A|$ that admit an upper bound on e.g. $\mathsf{E}(B)$ in terms of $\mathsf{E}_f(C)$. In particular, one of e.g. $\mathsf{E}(B), \mathsf{E}_f(C)$ must be 'small'.

The auxiliary decomposition is particular to the theorem in question and typically is proved as follows. First, we apply a regularisation lemma of Xue [35], who also credits Rudnev. This regularisation lemma enables us to extract subsets B and C, where C is endowed with an additive structure relating to the additive energy of B. We then use results from incidence geometry (or Weil bounds) – for instance Rudnev's incidence theorem between points and planes [24] – to yield the auxiliary decomposition result.

To decompose A, we first find subsets $C \subseteq B \subseteq A$ and identify which of e.g. $\mathsf{E}(B), \mathsf{E}_f(C)$ is small. We then remove the subset pertaining to small energy from A and repeat this argument. In this way, we decompose A. We describe this algorithmic approach in Section 3.3.

Previous Balog-Wooley decomposition arguments proceeded in a subtly different way: the auxiliary 'mixed energy result' extracted a single subset $B \subseteq A$ for which one of e.g. $\mathsf{E}(A), \mathsf{E}_f(B)$ is small. Unlike the regularisation afforded by Xue's lemma, it is not guaranteed that $|B| \gtrsim |A|$. Instead, an algorithm at the end, present in the original paper of Balog and Wooley [1] performs this regularisation.

Using Lemma 2, we streamline and simplify the analysis.

2.3. **Applications.** Our first application is a quantitative improvement to an expansion result of Mirzaei [16].

Theorem 5. *Let $A \subset \mathbb{F}_p$, with $|A| \ll p^{23/52}$ and let $f \in \mathbb{F}_p[x,y]$ be a non-degenerate quadratic polynomial. Then*

$$\max\{|A-A|, |f(A,A)|\} \gtrsim |A|^{28/23}.$$

This improves the exponent of $6/5 + 4/305$ attained by Mirzaei to the exponent $6/5 + 2/115$. We note in particular that this result demonstrates the efficiencies that low-energy decomposition results yield. To see how low-energy decomposition results are used in the sum-product problem, we refer the reader to [17].

As a second application, we give improvements of the main result of Swaenepoel and Winterhof [31, Theorem 1]:

Theorem 6. *Let $S, T \subset \mathbb{F}_q$ and suppose $f(T) \subseteq T$ for some rational function $f \in \mathbb{F}_q(x)$, satisfying (5). Then T contains a large subset U with $|U| \geq |T|/(d+1)$, where $d = deg(f)$, such that*

$$\left|\sum_{s\in S}\sum_{u\in U}\psi(su)\right| \ll \left(\frac{|S|^3|T|^3 q}{M_T}\right)^{1/4},$$

where M_T is given by (14).

This result is non-trivial in a range of the form $|S||T|^2 \gtrsim q^{3/2}$ and its strength increases as $|T|$ becomes larger than $|S|$; when $|S| = |T|$, this result is weaker than (9). Theorem 6 is a strict improvement over [31, Theorem 1], as can be seen through a comparison of the quantities M_T appearing in the two theorems (given by (14) and (6)). Furthermore, Theorem 6 yields quantitative improvements to [31, Theorems 5 and 9].

As a third application, we provide a variant of Theorem 6 concerning small subsets of $\mathbb{F}_p$ and quadratic polynomials.

Theorem 7. *Let $f \in \mathbb{F}_p[x]$ be a quadratic polynomial and let $T \subseteq \mathbb{F}_p$. Suppose that $f(T) \subseteq T$ and $|T| \leq p^{5/8}$. Then there exists a subset $U \subseteq T$, with $|U| \gg |T|$ such that for any set $S \subseteq \mathbb{F}_p$, we have*

$$\left|\sum_{s\in S}\sum_{u\in U} e_p(su)\right| \lesssim (|S|^3|T|^{3-\frac{1}{5}}p)^{1/4}. \tag{15}$$

Moreover, if $|S| \leq p^{5/8}$ and $f(S) \subseteq S$, then there exists $V \subseteq S$, with $|V| \gg |S|$, such that

$$\left|\sum_{u\in U}\sum_{v\in V} e_p(uv)\right| \lesssim p^{1/8}(|S||T|)^{17/20}. \tag{16}$$

To allow for a rough comparison between the estimates of Theorem 7 and Vinogradov's estimate (9), suppose we have sets $S, T, V, U \subset \mathbb{F}_p$ as given by Theorem 7, with $|S| = |T| = N$ and let $f \in \mathbb{F}_p[x]$ be quadratic. Then

$$\left|\sum_{u\in U}\sum_{v\in V} e_p(uv)\right| \leq \begin{cases} Np^{\frac{1}{2}} & \text{for } p^{\frac{1}{2}+\frac{1}{18}} < N \leq p\ ; \\ N^{\frac{29}{20}}p^{\frac{1}{4}+o(1)} & \text{for } p^{\frac{1}{2}} < N \leq p^{\frac{1}{2}+\frac{1}{18}}\ ; \\ N^{\frac{17}{10}}p^{\frac{1}{8}+o(1)} & \text{for } p^{\frac{1}{2}-\frac{1}{22}} < N \leq p^{\frac{1}{2}}\ ; \\ N^2 & \text{for } N < p^{\frac{1}{2}-\frac{1}{22}}. \end{cases}$$

That is, in this demonstrative setting, Theorem 7 is superior to both Vinogradov's estimate and the trivial bound in the range $p^{\frac{1}{2}-\frac{1}{22}} < N \leq p^{\frac{1}{2}+\frac{1}{18}}$.

Finally, we turn our attention to the finite field Littlewood problem and prove lower bounds on the l_1 norm of exponential sums over certain types of sets.

Theorem 8. *Let $f \in \mathbb{F}_p[x]$ denote a quadratic polynomial and let $A \subseteq \mathbb{F}_p$ be any set with $|A| \ll p^{2/3}$ and $|A+A| \ll |A|$, then*

$$\frac{1}{p}\sum_{\lambda\in\mathbb{F}_p}\left|\sum_{a\in f(A)} e_p(\lambda a)\right| \gtrsim |A|^{1/4}. \tag{17}$$

Let $T \subseteq \mathbb{F}_p$, with $|T| \ll p^{5/8}$ and suppose $f(T) \subseteq T$. Then

$$\frac{1}{p}\sum_{\lambda\in\mathbb{F}_p}\left|\sum_{t\in T} e_p(\lambda t)\right| \gtrsim |T|^{1/10}. \tag{18}$$

Estimate (17) provides a quantitative improvement to a result of Garcia [8, Corollaries 8 and 9] for quadratic polynomials. Using Proposition 1, one can recover, under a more favourable p-constraint on the set in question, the estimate of Shkredov [27, Corollary 2] concerning a lower bound on the l_1 norm of exponential sums over multiplicatively structured sets. We further note that estimate (18) appears to be new in the sense that it provides a new class of examples toward the modular Littlewood problem. Finally, we mention that orbits of dynamical systems, defined in (11), provide natural examples of sets T to which Theorems 6, 7 and 8 apply.

3. Preliminaries

3.1. **Energy preliminaries.** We record the following energy sub-additivity lemma, which is a consequence of the Cauchy-Schwarz inequality.

Lemma 1. *Let $f \in \mathbb{F}_q[x,y]$ and $V_1, \dots, V_k \subseteq \mathbb{F}_q$. Then*

$$E_f\Big(\bigcup_{i=1}^k V_i\Big) \leq \Big(\sum_{i,j=1}^k \mathsf{E}_f(V_i,V_j)^{1/2}\Big)^2.$$

Furthermore, if f has the property that, for any $X, Y \subset \mathbb{F}_q$,

$$\mathsf{E}_f(X,Y) \ll \mathsf{E}_f(X)^{1/2}E_f(Y)^{1/2}, \tag{19}$$

then we have

$$\mathsf{E}_f\Big(\bigcup_{i=1}^k V_i\Big) \ll \Big(\sum_{i=1}^k \mathsf{E}_f(V_i)^{1/4}\Big)^4. \tag{20}$$

Proof. Without loss of generality, we may assume that the sets V_i, $1 \le i \le k$ are pairwise disjoint. Thus, by an application of the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\mathsf{E}_f\Big(\bigcup_{i=1}^k V_i\Big) &= \sum_{i,j,k,l=1}^k \sum_{\lambda\in\mathbb{F}_q} r_{f(V_i,V_j)}(\lambda)\cdot r_{f(V_k,V_l)}(\lambda)\\ &\le \sum_{i,j,k,l=1}^k \Big(\sum_{\lambda\in\mathbb{F}_q} r_{f(V_i,V_j)}(\lambda)^2\Big)^{1/2}\Big(\sum_{\lambda\in\mathbb{F}_q} r_{f(V_k,V_l)}(\lambda)^2\Big)^{1/2} = \Big(\sum_{i,j=1}^k \mathsf{E}_f(V_i,V_j)^{1/2}\Big)^2.\end{aligned}$$

Applying (19) to the above inequality gives (20). □

Key to our results is the following regularisation lemma as recorded by Xue [35]. Although Xue formulates the regularisation in $\mathbb{R}$, the proof is valid over abelian groups. This regularisation has origins in [26, Proposition 16].

Lemma 2. *Let A be a subset of an abelian group. There exist subsets $C \subseteq B \subseteq A$, with $|A| \ll |B| \ll (\log|A|)^2|C|$, a number $1 \le t \le |B|$ and a set $D = \{x \in B - B : t \le r_{B-B}(x) < 2t\}$ such that*

$$|D|t^2 \ll \mathsf{E}(B) \ll (\log|B|)|D|t^2$$

and for any $c \in C$,

$$r_{D+B}(c) \approx (\log|A|)^2\frac{|D|t}{|B|}.$$

3.2. **Energy estimates and incidence theorems.** A typical application of incidence geometry is to energy estimates.

Using Rudnev's [24] incidence theorem between points and planes in $\mathbb{F}^3$, Koh, Mirzaei, Pham and Shen [11, Theorem 2.1] obtained the following energy estimate, generalising a result of Pham, Vinh and de Zeeuw [20].

Lemma 3. *Given sets $U, V, W \subseteq \mathbb{F}_p^\times$, with $|U||V||W| \ll p^2$ and a quadratic polynomial $f \in \mathbb{F}_p[x,y,z]$, which depends on each variable and is not of the form $g(h(x)+k(y)+l(z))$, we have*

$$\begin{aligned}&|\{(u_1,u_2,v_1,v_2,w_1,w_2) \in U^2\times V^2\times W^2 : f(u_1,v_1,w_1) = f(u_2,v_2,w_2)\}|\\ &\qquad\ll (|U||V||W|)^{3/2} + \max\{|U|^2|V|^2, |U|^2|W|^2, |V|^2|W|^2\}.\end{aligned}$$

For large sets we have a similar result, relying instead on point-line incidence estimates due to Vinh [32]:

Lemma 4. *Let P denote a set of points and L a collection of lines over $\mathbb{F}_q^2$. Let $I(P,L) = |\{(p,l) \in P\times L : p \in l\}|$ denote the number of incidences between P and L. Then*

$$\left|I(P,L) - \frac{|P||L|}{q}\right| \le \sqrt{q|P||L|}.$$

Vinh's bound enables us to obtain a bound on the representation function of $C \cdot X$, which we recycle into an energy estimate using standard techniques.

Lemma 5. *Let $B, C, D, X \subset \mathbb{F}_q^\times$ and let $\kappa > 0$ be such that $r_{D+B}(x) \geq \kappa$ for all $x \in C$. Suppose that*

$$n \geq 2\frac{|B||D||X|}{\kappa q}. \tag{21}$$

Then, writing $S = \{x \in \mathbb{F}_q : r_{C\cdot X}(x) \geq n\}$, we have

$$|S| \ll \frac{|B||D||X|q}{\kappa^2 n^2}.$$

Proof. We have

$$\begin{aligned} n|S| &\leq \sum_{z\in S} r_{C\cdot X}(z) = |\{(c,x,z) \in C \times X \times S : cx = z\}| \\ &\leq \kappa^{-1} \cdot |\{(d,b,x,z) \in D \times B \times X \times S : (d+b)x = z\}|. \end{aligned}$$

Now, note that the cardinality in the last line above can be interpreted as the number of incidences between the set of points $P = D \times S$ and the collection of lines $L = \{l_{x,b} : (x,b) \in X \times B\}$, where $l_{x,b} = \{(s,t) \in \mathbb{F}_q^2 : t = x(s+b)\}$. Thus, applying Lemma 4, we get

$$\kappa \cdot n \cdot |S| \leq \frac{|A||D||S||X|}{q} + \sqrt{q|A||D||S||X|}.$$

Applying (21) gives the desired result. □

We record an analogous result, obtained by Roche-Newton, Shparlinksi and Winterhof [23, Lemma 2.3] using Weil bounds in place of Vinh's incidence estimate.

Lemma 6. *Let $B, C, D \subset \mathbb{F}_q^\times$ and let $f \in \mathbb{F}_q(x)$ denote a rational function of degree d, satisfying (5). Let $\kappa > 0$ be such that $r_{D-B}(c) \geq \kappa$ for all $c \in C$. Suppose that*

$$n \geq 2d\frac{|B||C||D|}{\kappa q}.$$

Writing $S = \{x \in \mathbb{F}_q : r_{g(C,C)}(x) \geq n\}$, where $g(x,y) = f(x) + f(y)$, we have

$$|S| \ll_d \frac{|B||C||D|q}{\kappa^2 n^2}.$$

3.3. **The decomposition algorithm.** We now describe the decomposition algorithm introduced in Section 2.2.

Let $A \subseteq \mathbb{F}$ and let $1 \leq M \leq |A|$ denote a parameter. The algorithm outputs two disjoint sets $S, T \subseteq A$ so that $A = S \sqcup T$ and $\mathsf{E}(S) \leq |A|^3 M^{-1}$. The set T inherits certain properties, so that, using the appropriate energy bounds that we obtain in this note, we can control $\mathsf{E}_f(T)$.

The algorithm will produce the following sequences of sets:

$$A_0 = A \supseteq A_1 \supseteq \dots$$
$$S_0 = \emptyset \subseteq S_1 \subseteq \cdots \subseteq A$$
$$T_0 = \emptyset \subseteq T_1 \subseteq \cdots \subseteq A$$

The algorithm relies on a *black box*, which for each A_i, outputs two subsets $C_i \subseteq B_i \subset A$, with $|C_i| \gtrsim |B_i| \gg |A_i|$:

Algorithm 1: Decomposition Algorithm

Input : $A \subseteq \mathbb{F}$, suitably chosen $M \in [1, |A|]$
Output: Decomposition $A = S \sqcup T$ so that $\max\{\mathsf{E}(S), \mathsf{E}_f(T)\} \leq |A|^3 M^{-1}$

Initialisation $A_0 = A$, $S_0 = \emptyset$, $T_0 = \emptyset$;
if $\mathsf{E}(A) \leq |A|^3 M^{-1}$ **then**
  $A_0 = \emptyset$ $S = A$;
  $T = \emptyset$;
end
if $E_f(A) \leq |A|^3 M^{-1}$ **then**
  $A_0 = \emptyset$ $S = \emptyset$;
  $T = A$
end
while $A_i \neq \emptyset$ **do**
  instructions;
  if $|A_i| < (|A|^3/M)^{\frac{1}{3}}$ **then**
    $A_{i+1} = \emptyset$;
    $S_{i+1} = S_i \sqcup A_i$;
    $T_{i+1} = T_i$;
  else
    Apply *black box* to A_i to get $C_i \subseteq B_i \subseteq A_i$;
    if $\mathsf{E}(B_i) < |A|^{3-\frac{1}{4}} |B_i|^{\frac{1}{4}} / M$ **then**
      $A_{i+1} = A_i \setminus B_i$;
      $S_{i+1} = S_i \sqcup B_i$;
      $T_{i+1} = T_i$;
    else
      $A_{i+1} = A_i \setminus C_i$;
      $S_{i+1} = S_i$;
      $T_{i+1} = T_i \sqcup C_i$;
    end
  end
end

Notice that this algorithm terminates since $|A_i|$ is a monotonically decreasing sequence. Since we remove a subset of size $\gtrsim |A|M^{-1/3}$ at each step, the algorithm terminates after at most M iterations.

Suppose that the algorithm terminates in k iterations so that $A_{k+1} = \emptyset$. Let $S = S_{k+1}$ and $T = T_{k+1}$. We introduce the notation $\mathcal{I}_S$ to denote the stages of the algorithm in which B_i is added to S_i. Similarly, $\mathcal{I}_T$ indexes the iterations in which T_i is augmented.

We justify now that $\mathsf{E}(S) \leq |A|^3 M^{-1}$; the analogous justification for $\mathsf{E}_f(T)$ will require tools that we develop in the subsequent. Observe that S is a disjoint union of a subset of the sets B_i arising at stages of the algorithm, together with possibly a small set A_k.

Applying Lemma 1, we have

$$\begin{aligned}\mathsf{E}(S) = \mathsf{E}(\bigsqcup_{i\in\mathcal{I}_S} B_i \sqcup A_k) &\leq \left(\sum_{i\in\mathcal{I}_S} \mathsf{E}(B_i)^4 + \mathsf{E}(A_k)^4\right)^{1/4} \\ &< \left(\sum_{i\in\mathcal{I}_S}\left(\frac{|A|^{3-\frac{1}{4}}|B_i|^{\frac{1}{4}}}{M}\right)^4 + \left(\frac{|A|^3}{M}\right)^4\right)^{1/4} \\ &= \left(\frac{|A|^{11}}{M^4}\sum_{i\in\mathcal{I}_S}|B_i| + |A_i|^{12}\right)^{1/4} \leq \frac{|A|^3}{M}.\end{aligned} \tag{22}$$

We record two further identities key to our applications of this algorithm. For $i \in \mathcal{I}_T$, we have

$$\mathsf{E}(B_i) \geq |A|^{3-1/4}|B_i|^{1/4}M^{-1} \tag{23}$$

and

$$T = \bigsqcup_{i\in\mathcal{I}_T} C_i. \tag{24}$$

4. Proof of Theorem 1

We first prove a preliminary decomposition result. This is the black box that we use in Algorithm 1 to prove Theorem 1.

Proposition 1. *Let $A \subseteq \mathbb{F}_p^\times$. Let $f \in \mathbb{F}_p[x,y]$ denote a non-degenerate quadratic polynomial. Then there exist subsets $C \subseteq B \subseteq A$ with $|C| \gtrsim |B| \gg |A|$ so that*

$$\mathsf{E}(B)^3\mathsf{E}_f(C,X)^2 \lesssim |A|^{11}|X|^3$$

and

$$\mathsf{E}(B)^3\mathsf{E}_f(X,C)^2 \lesssim |A|^{11}|X|^3$$

for any $X \subseteq \mathbb{F}_p^\times$ satisfying $|B|^5|X|(\mathsf{E}(B))^{-1} \ll p^2$ and $|X| \gg |B|$.

Note that Proposition 1 is trivially true if $|A|^2 \lesssim |X|$.

Proof. We prove only the first energy bound the second follows almost identically due to the symmetry on x and y in the definition of a non-degenerate polynomial.

We first apply Lemma 2, to get subsets $C \subseteq B \subseteq A$ and a set $D \subseteq B-B$ so that $r_{B-B}(d) \in [t, 2t)$ for each $d \in D$ and $\mathsf{E}(B) \gtrsim |D|t^2$. Moreover, $r_{D+B}(c) \sim |D|t|A|^{-1}$ for all $c \in C$.

Without loss of generality, we may assume that $0 \notin D$, removing it if necessary. Indeed, if $|D| \ll 1$, then $\mathsf{E}(B) \ll |B|^2$ and so we are done using the trivial bound $\mathsf{E}_f(C,X) \leq |C|^2|X|$. Otherwise, if $|D| \gg 1$, then $\mathsf{E}(B) \gg (|D|-1)t^2$, and so removing an element from D is without consequence.

Since the polynomial $f \in \mathbb{F}_p[x,y]$ is non-degenerate, one can readily verify that the associated polynomial $\tilde{f} \in \mathbb{F}_p[x,y,z]$ defined by $\tilde{f}(x,y,z) = f(x+y,z)$ is not of the form $g(h(x)+k(y)+l(z))$.

Let us now apply Lemma 3, deferring the justification of the p-constraint until the end of the proof. We have

$$\begin{aligned}
\mathsf{E}_f(C,X) &:= |\{(c_1,c_2,x_1,x_2) \in C^2 \times X^2 \colon f(c_1,x_1) = f(c_2,x_2)\}| \\
&\ll \frac{|A|^2}{|D|^2t^2}|\{(b_1,b_2,d_1,d_2,x_1,x_2) \in B^2 \times D^2 \times X^2 \colon f(b_1+d_1,x_1) = f(b_2+d_2,x_2)\}| \\
&\lesssim \frac{|A|^2}{|D|^2t^2} \max\left\{|D|^{3/2}|X|^{3/2}|B|^{3/2}, |B|^2|X|^2, |B|^2|D|^2, |X|^2|D|^2\right\} \\
&\ll \frac{|A|^2}{|D|^2t^2} \max\left\{|D|^{3/2}|X|^{3/2}|B|^{3/2}, |B|^2|X|^2, |X|^2|D|^2\right\} \\
&:= \frac{|A|^2}{|D|^2t^2} M\,.
\end{aligned}$$

We split into cases according to the maximand in the above expression.

Case 1: $M = |D|^{3/2}|X|^{3/2}|B|^{3/2}$**.** Multiplying by $(|D|t)^{1/2}$ and recalling that $|D|t \leq |B|^2$ completes the proof in this case.

Case 2: $M = |B|^2|X|^2$**.** In particular, we have that $|D|^3 < |B||X|$, and so, using the trivial bound $t \leq |B|$, we have $\mathsf{E}(B)^3\mathsf{E}_f(C,X)^2 < |B||X| \cdot |B|^6 \cdot (|X||C|^2)^2 \leq |B|^7|X|^3|A|^4$.

Case 3: $M = |X|^2|D|^2$**.** We have $|B| \leq |D|, |X|$ and so we bound $\mathsf{E}_f(C,X) \leq |B|^2|X|$. Similar to Case 3, we have

$$\begin{aligned}
|D|\mathsf{E}(B)^3\mathsf{E}_f(C,X)^2 \sim (t|D|)^4(t^2\ \mathsf{E}_f(C,X))\mathsf{E}_f(C,X) &\lesssim |B|^8(|A|^2|X|^2)|C|^2|X| \\
&\ll |B|^7|A|^4|X|^3 \cdot |B| \leq |B|^7|A|^4|X|^3|D|\,.
\end{aligned}$$

Finally, let us justify the p-constraint necessary for our application of Lemma 3. Note that by Lemma 2, we have

$$\mathsf{E}(B)|D| \lesssim (|D|t)^2 \ll |B|^4.$$

Hence if $|B|^5|X|/\mathsf{E}(B) \ll p^2$, our use of Lemma 3 is justified. □

4.1. **Proof of Theorem 1.** We use Proposition 1 to run Algorithm 1 with a parameter $M \geq 1$ to be determined. We must show that, with an appropriate choice of M, $\mathsf{E}_f(T) \lesssim |A|^3/M$.

We recall that, by (24), T can be be written as a disjoint union of T_i for $i \in \mathcal{I}_T$, and so we can apply the subadditivity lemma, Lemma 1. From (23) and Proposition 1, we deduce an upper bound on E_f:

$$\mathsf{E}_f(T) = \mathsf{E}_f(\bigsqcup_{i\in\mathcal{I}_T} C_i) \leq \left(\sum_{i,j} \mathsf{E}_f(C_i,C_j)^2\right)^{1/2} \lesssim \left(\sum_{i,j\in\mathcal{I}_T} M^3|A|^3|C_j||C_i|\right)^{1/2} \leq M^{3/2}|A|^{5/2}\,.$$

When applying Proposition 1, we take X to be the larger of C_j, C_i. Hence

$$\max\{\mathsf{E}(T), \mathsf{E}_f(S)\} \lesssim \frac{|A|^3}{M} + M^{3/2}|A|^{5/2}\,,$$

and so we choose $M = |A|^{1/5}$.

Finally, it remains to justify our use of Proposition 1 in the algorithm. Note that Proposition 1 is used when

$$\mathsf{E}(B_i) > |A|^{3-9/20}|B_i|^{1/4} \tag{25}$$

and recall that we may use Proposition 1 when the condition $|B_i|^5|X|(\mathsf{E}(B_i))^{-1} \ll p^2$ is satisfied. Suppose to the contrary that $|B_i|^5|C_j| \gg \mathsf{E}(B_i)p^2$. Then $|B_i|^5|C_j| \gg |A|^{3-9/20}|B_i|^{1/4}|A|^{16/5}$ and so $\max\{|B_i|,|C_j|\}^{23/4} > |A|^{23/4}$. Since $|B_i|, |C_j| < |A|$, this is a contradiction.

Using e.g. [30, Proposition 1] in place of Lemma 2, we can obtain the following variant of Proposition 1:

Proposition 2. *Let $A, V, X \subseteq \mathbb{F}_p$, with $|V|, |X| \gg |A|$. Let $f \in \mathbb{F}_p[x,y]$ denote a non-degenerate quadratic polynomial. Then there exist subsets $C \subseteq B \subseteq A$ with $|C| \gtrsim |B| \gg |A|$ so that, if $|X||B|^2|V|^3/\mathsf{E}(B,V) \ll p^2$, then*

$$\mathsf{E}(B,V)^3\mathsf{E}_f(C,X)^2 \lesssim |A|^6|X|^3|V|^5.$$

Sketch. The proof proceeds almost identically to that of Proposition 1. Using the regularisation [30, Proposition 1] applied to sets A and V yields subsets $C \subseteq B \subseteq A$ with $|C| \gtrsim |A|$ depending on V so that $r_{D+V}(c) \sim |D|t|A|^{-1}$ for each $c \in C$. Here $D \subseteq B - V$ supports the additive energy of $\mathsf{E}(B,V)$, so that $|D|t^2 \gtrsim \mathsf{E}(B,V)$.

Applying Lemma 3, we obtain

$$\mathsf{E}_f(C,X) \lesssim \frac{|A|^2}{|D|^2t^2}\max\left\{|D|^{3/2}|X|^{3/2}|V|^{3/2}, |V|^2|X|^2, |V|^2|D|^2, |X|^2|D|^2\right\}.$$

A case analysis as before yields the proof. □

5. Proofs of Theorem 3 and Theorem 4

Since both proofs are almost identical, we prove only Theorem 3. To prove Theorem 4, it suffices to replace Lemma 5 by Lemma 6.

We begin by proving an auxiliary energy decomposition result:

Proposition 3. *Let $A \subset \mathbb{F}_q$. There exist subsets $C \subseteq B \subseteq A$, with $|A| \ll |B| \ll (\log|A|)^2|C|$ such that*

$$\mathsf{E}^{\times}(C)\mathsf{E}(B) \ll \frac{|A||B|^3|C|^3(\log|A|)}{q} + |A|^2|B||C|(\log|A|)^2 q.$$

Proof. Let B, C, D be the sets given by Lemma 2 so that $\mathsf{E}(B) \ll (\log|A|)|D|t^2$ for some $t \geq 1$. We have

$$\mathsf{E}^{\times}(C) = \sum_{\lambda\in\mathbb{F}_q} r_{C\cdot C}(\lambda)^2 = \sum_{\lambda\in\mathbb{F}_q} r_{C\cdot C}(\lambda)^2 = \sum_{\lambda\in S_0} r_{C\cdot C}(\lambda)^2 + \sum_{j=1}^{J}\sum_{\lambda\in S_j} r_{C\cdot C}(\lambda)^2, \tag{26}$$

where for a parameter κ we define

$$S_0 = \left\{\lambda \in \mathbb{F}_q : r_{C\cdot C}(\lambda) < 2\frac{|B||C||D|}{\kappa q}\right\}$$

and for $j = 0, \dots J$ we set

$$S_j = \left\{\lambda \in \mathbb{F}_q : 2^j\frac{|B||C||D|}{\kappa q} \leq r_{C\cdot C}(\lambda) < 2^{j+1}\frac{|B||C||D|}{\kappa q}\right\}.$$

Since $\max\{r_{C\cdot C}(\lambda) : \lambda \in \mathbb{F}_q\} \leq |A|$, we have $J = \log_2|A| \ll \log|A|$.

We use the identity $\sum_{\lambda\in\mathbb{F}_q} r_{C\cdot C}(\lambda) = |C|^2$ to bound

$$\sum_{\lambda\in S_0} r_{C\cdot C}(\lambda)^2 \leq 2\frac{|B||C||D|}{\kappa q}\sum_{\lambda\in S_0} r_{C\cdot C}(\lambda) \ll \frac{|B||C|^3|D|}{\kappa q}.$$

For $j \in [1, J]$, we apply Lemma 5 with $\kappa = |D|t|A|^{-1}$, as justified by Lemma 2, and with

$$n = 2^j\frac{|B||C||D|}{\kappa q}$$

we get

$$\sum_{\lambda\in S_j} r_{C\cdot C}(\lambda)^2 < (2n)^2|S_j| \ll \frac{|B||C||D|q}{\kappa^2}.$$

Returning to (26) concludes the proof. □

We now proceed with the proof of Theorem 3.

Proof of Theorem 3. We apply the algorithm described in Section 3.3, using Proposition 3 as the black box. To conclude the proof of the theorem, we must show that $\mathsf{E}^{\times}(T) \leq |A|^3/M$ for a suitable choice of M.

For $i \in \mathcal{I}_T$, by (23), (24) and Lemma 1, we have

$$\begin{aligned}
\mathsf{E}^{\times}(T) = \mathsf{E}^{\times}(\sqcup_{i\in\mathcal{I}_T} C_i) &\leq \left(\sum_i \mathsf{E}^{\times}(C_i)^4\right)^{1/4} \\
&\lesssim \left(\sum_i \left(\frac{|A_i|^2|B_i||C_i|}{\mathsf{E}(B_i)}\left(\frac{|B_i||C_i|^2(\log|A|)}{q} + q(\log|A|)^2\right)\right)^4\right)^{1/4} \\
&\leq \left(\sum_i \left(\frac{M|A_i|^2|B_i||C_i|}{|A|^{3-1/4}|B_i|^{1/4}}\left(\frac{|B_i||C_i|^2(\log|A|)}{q} + q(\log|A|)^2\right)\right)^4\right)^{1/4} \\
&\leq M\left(\frac{|A|^{15/4}(\log|A|)}{q} + |A|^{3/4}(\log|A|)^2 q\right)\left(\sum_i |C_i|\right)^{1/4} \\
&\leq M\left(\frac{|A|^4(\log|A|)}{q} + |A|(\log|A|)^2 q\right).
\end{aligned}$$

Comparing this with the upper bound $\mathsf{E}(S) \leq |A|^3/M$, in (22), we recover the optimal choice for M as given in the statement of the theorem. □

6. Proof of Theorem 5

The following lemma is proved by Mirzaei [16].

Lemma 7. *Let $A, B \subset \mathbb{F}_p$, with $|A||B||A-B| \ll p^2$. For a non-degenerate, quadratic polynomial $f \in \mathbb{F}_p[x,y]$, we have*

$$\mathsf{E}_4(A,B) \ll \frac{|f(A,A)|^2|B|^3}{|A|}.$$

We extract the following from the arguments of Rudnev, Shakan and Shkredov [25, Equations 3.2 and 3.3].

Lemma 8. *Let $A \subset \mathbb{F}_p$, then*

$$|A|^{24} \leq \mathsf{E}_4(A)^2|A-A|^5\Delta^4\mathsf{E}_4(A,D),$$

for some $\Delta \geq 1$ and $D \subset A-(A-A)$ such that $r_{A-(A-A)}(d) \approx \Delta$ for all $d \in D$.

We recall a Plünnecke-Ruzsa type inequality appearing in [19].

Lemma 9. *Given finite, non-empty sets A and B in an abelian group, we have*

$$|kA - lA| \leq \frac{|A+B|^{k+l}}{|B|^{k+l-1}},$$

where kA is used to denote the k-fold sum set of A

By Proposition 2, there exist $C \subset B \subset A$, with $|C| \gtrsim |B| \gg |A|$ such that

$$\mathsf{E}(B, A-A) \lesssim |A|^{1/3}|f(C,C)|^{2/3}|A-A|^{5/3}. \tag{27}$$

To justify the application of Proposition 2, we note that if the relevant p-constraint fails, we get

$$\frac{|A-A|^3|B|^2|C|}{\mathsf{E}(B,A-A)} > p^2 > |A|^{104/23}.$$

Then using

$$\mathsf{E}(B,A-A) \geq \frac{|B|^2|A-A|^2}{|B+A-A|} \geq \frac{|B|^2|A-A|^2}{|A+A-A|} \gg \frac{|A|^4}{|A-A|},$$

we see the required bound holds in this case.

Next, applying Lemma 7 to Lemma 8, under the condition $|A||A+A-A||A+A-A-A| \ll p^2$, we obtain

$$\begin{aligned}|A|^{25} \ll & \mathsf{E}_4(B)^2|B-B|^5|f(B,B)|^2(|D|\Delta)^2(|D|\Delta^2)\\ & \ll \mathsf{E}_4(B)^2|B-B|^5|f(B,B)|^2(|B||B-B|)^2(\mathsf{E}(B,B-B)).\end{aligned}$$

A second application of Lemma 7 gives

$$|A|^{19} \lesssim |f(A,A)|^6|A-A|^7\mathsf{E}(B,A-A). \tag{28}$$

Applying (27) to this gives the result.

To check the required p-constraint, note that Lemma 9 implies

$$|A||A+A-A||A+A-A-A| \leq \frac{|A-A|^7}{|A|^4}.$$

Hence, if the condition of Lemma 7 (as applied to $\mathsf{E}_4(B,D)$) fails, we have $|A-A| \gg |A|^{28/23}$, which gives the required result. A similar analysis is necessary for the application of Lemma 7 to $\mathsf{E}_4(B)$; the ensuing required p-constraint is more forgiving and is already satisfied.

7. Proof of Theorems 6, 7 and 8

We state some auxiliary exponential sum estimates, derived from basic applications of Hölder's inequality. See also [14, Equation 3.7].

Lemma 10. *Let $X, Y \subseteq \mathbb{F}_q$. We have*

$$\left|\sum_{x\in X}\sum_{y\in Y}\psi(xy)\right|^4 \leq q|X|^3\mathsf{E}(Y). \tag{29}$$

and

$$\left|\sum_{x\in X}\sum_{y\in Y}\psi(xy)\right|^8 \leq q|X|^4|Y|^4\mathsf{E}(X)\mathsf{E}(Y). \tag{30}$$

In addition, we will use the following result of Konyagin and Shkredov [13, Lemma 4].

Lemma 11. *Let $X_1 \subseteq X \subseteq \mathbb{F}_q$. Then*

$$\frac{1}{q}\sum_{y\in\mathbb{F}_q}\left|\sum_{x\in X}\psi(xy)\right| \geq \frac{|X_1|^2}{|X|^{1/2}\mathsf{E}(X_1)^{1/2}}.$$

Let us record corollaries of Theorems 1 and 4, obtained through the same scheme as [31, Lemma 4].

Lemma 12. *Let $f \in \mathbb{F}_p[x]$ denote a quadratic polynomial and let $T \subseteq \mathbb{F}_p$, with $|T| \leq p^{5/8}$ and the property that $f(T) \subseteq T$. Then there exists a subset $U \subseteq T$, with $|U| \gg |T|$ such that $\mathsf{E}(U) \lesssim |T|^{3-1/5}$.*

Proof. Under the assumption $|T| \leq p^{5/8}$, Theorem 1 implies that $T = B \sqcup C$ such that

$$\max\{\mathsf{E}(B), \mathsf{E}(f(C))\} \lesssim |T|^{3-1/5}.$$

Now, either $|B| \approx |T|$ or $|C| \approx |T|$. If the former is true, we may take $U = B$. If the latter is true, we take $U = f(C)$. Then $f(C) \subseteq f(T) \subseteq T$ by our assumption and clearly $|f(C)| \approx |T|$ as required. □

We also require the following analogue of the above, based on Theorem 4. The proof is essentially the same as that of [31, Lemma 4].

Lemma 13. *Let $f \in \mathbb{F}_q(x)$ denote a rational function of degree d, satisfying* (5) *and let $T \subseteq \mathbb{F}_q$ with the property that $f(T) \subseteq T$. Then there exists a subset $U \subseteq T$, with $|U| \geq |T|/(d+1)$ such that $\mathsf{E}(U) \lesssim |T|M_T^{-1}$, where M_T is given by* (14).

Proofs of Theorems 6 and 7. To prove Theorem 6, we use Lemma 13, to identify a subset $U \subseteq T$, with $|U| \geq |T|/(d+1)$. Then applying the bound on the additive energy of U, given by Lemma 13, to (29), we obtain the required result.

Theorem 7 is proved similarly. To obtain (15), we use Lemma 12 in place of Lemma 13. To prove (16), firstly through the same process, we identify a subset $V \subseteq S$, with $|V| \gg |S|$. Then we simply apply the bounds on the additive energy of U and V to (30). □

Proof of Theorem 8. To prove (17), we apply Proposition 1 to the set A and polynomial $g(x,y) = f(x) + f(y)$. Since f is quadratic, the polynomial g is non-degenerate. Note that $\mathsf{E}(f(A)) \ll \mathsf{E}_g(A)$. This ensures existence of sets $B, C \subseteq A$, with $|C| \gtrsim |B| \gg |A|$ such that

$$\mathsf{E}(B)^3\mathsf{E}(f(C))^2 \lesssim |A|^{14}.$$

Further note that $\mathsf{E}(B) \geq |B|^4|B+B|^{-1} \geq |B|^4|A+A|^{-1} \gg |A|^3$.

We use Lemma 11 with $X = f(A)$ and $X_1 = f(C)$ to get the required result. To check the required p-constraint of Proposition 1, note that $|A|^6(\mathsf{E}(B))^{-1} \ll |A|^3$ and so our application is justified if $|A|^3 \ll p^2$.

To prove (18), we note that since $f(T) \subseteq T$, we may apply Lemma 12, which ensures existence of $U \subseteq T$ with $|U| \approx |T|$ and such that $\mathsf{E}(U) \lesssim |T|^{3-1/5}$. We apply Lemma 11, with $X = T$ and $X_1 = U$, obtaining the required result. □

Acknowledgements

We thank Ilya Shkredov, Igor Shparlinksi and Arne Winterhof for their helpful comments and suggestions. The second author was supported by the Austrian Science Fund FWF grants P 30405 and P 34180.

References

[1] A. Balog and T.D. Wooley, *A low–energy decomposition theorem*, *Quart. J. Math.*, **68** (2017), 207–226.
[2] J. Bourgain and M. Z. Garaev, *Sumsets of reciprocals in prime fields and multilinear Kloosterman sums*, *Izv. Ross. Akad. Nauk Ser. Mat.*, **78(4)** (2014), 19–72; translation in *Izv. Mat.*, **78** (2014), 656–707.
[3] J. Bourgain, A. Glibichuk and S. V. Konyagin, *Estimates for the number of sums and products and for exponential sums over subgroups in fields of prime order*, *J. Lond. Math. Soc.*, **73** (2006), 380–398.
[4] B. Bukh and J. Tsimerman, *Sum-product estimates for rational functions*, *Proc. Lond. Math. Soc.*, **104** (2012), 1–26.
[5] M-C. Chang, *Expansion of quadratic maps in prime fields*, *Proc. Amer. Math. Soc.*, **142** (2013), 85–92.
[6] J. Cilleruelo, M. Garaev, A. Ostafe and I. E. Shparlinski, *On the concentration of points of polynomial maps and applications*, *Math. Z.*, **272** (2012), 825–837.
[7] M. Garaev, *The sum-product estimate for large subsets of prime fields*, *Proc. Amer. Math. Soc.*, **136** (2008), 2735–2739.
[8] V. C. Garcia, *The finite Littlewood problem in* $\mathbb{F}_p$, *Ramanujan J.*, **47** (2018), 1–14.
[9] J. Gutierrez and I. E. Shparlinski, *Expansion of orbits of some dynamical systems over finite fields*, *Bull. Austral. Math. Soc.*, **82** (2010), 232–239.
[10] N.Hegyvári, *Some remarks on multilinear exponential sums with an application*, *J. Number Theory*, **132** (2012), 94–102.
[11] D. Koh, M. Mirzaei, T. Pham and C-Y. Shen, *Exponential sum estimates over prime fields*, *Int. J. Number Theory*, 16 (2020), 291–308.
[12] D. Koh, H. Nassajian Mojarrad, T. Pham and C. Valculescu, *Four-variable expanders over the prime fields*, *Proc. Amer. Math. Soc.*, 146 (2018), 5025–5034.
[13] S. V. Konyagin and I. D. Shkredov, *A quantitative version of the Beurling-Helson theorem*, *Funct. Anal. Its Appl.*, **49** (2015), 110–121.
[14] S. V. Konyagin and I. E. Shparlinski, *Character sums with exponential functions and their applications*, *Cambridge Univ. Press*, 1999.
[15] S. Macourt, *Decomposition of subsets of finite fields*, *Funct. Approx. Comment. Math.*, **61** (2019), 243–255.
[16] M. Mirzaei, *A note on conditional expanders over prime fields*, *Discrete Mathematics*, to appear.
[17] A. Mohammadi and S. Stevens, *Attaining the exponent* 16/13 *for the sum-product problem in finite fields*, preprint, `arXiv:2103.08252 [math.CO]`.
[18] A. Ostafe, *Polynomial values in affine subspaces of finite fields*, *JAMA*, 138 (2019), 49–81.
[19] G. Petridis, *New proofs of Plünnecke-type estimates for product sets in groups*, *Combinatorica*, **32** (2012), 721–733.
[20] T. Pham, L.A. Vinh and F. de Zeeuw, *Three-variable expanding polynomials and higher-dimensional distinct distances*, *Combinatorica*, **39** (2017), 411 – 426.
[21] O. Roche-Newton and L. Li, *An improved sum-product estimate for general finite fields*, *SIAM J. Discrete Math.*, **25** (2011), 1285–1296.
[22] O. Roche-Newton and I. E. Shparlinski, *Polynomial values in subfields and affine subspaces of finite*, *Quart. J. Math.*, **66** (2015), 693–706.
[23] O. Roche-Newton, I. E. Shparlinski and A. Winterhof, *Analogues of the Balog-Wooley decomposition for subsets of finite fields and character sums with convolutions*, *Ann. Comb.*, **23** (2019), 183–205.
[24] M. Rudnev, *On the number of incidences between points and planes in three dimensions*, *Combinatorica*, (2018) **38**, 219–238.

[25] M. Rudnev, G. Shakan and I. Shkredov, *Stronger sum-product inequalities for small sets*, *Proc. Amer. Math. Soc.*, **148** (2020), 1467–1479.
[26] M. Rudnev, I. Shkredov and S. Stevens, *On the energy variant of the sum-product conjecture*, *Rev. Mat. Iberoam*, **36(1)** (2020), 207–232.
[27] I. D. Shkredov, *A remark on sets with small Wiener norm*, in: Raigorodskii A., Rassias M. (eds) *Trigonometric sums and their applications*, Springer, Cham, 2020.
[28] I. D. Shkredov and I. E. Shparlinski, *Double character sums with intervals and arbitrary sets*, *Proc. Steklov Inst. Math.*, **303** (2018), 239–258.
[29] J. Solymosi, *Incidences and spectra of graphs*, *Bolyai Soc. Math. Stud.*, **19** (2008), 499–513.
[30] S. Stevens and A. Warren, *On sumsets of convex functions*, preprint, `arXiv:2102.05446 [math.CO]`.
[31] C. Swaenepoel and A. Winterhof, *Additive double character sums over structured sets and applications*, *Acta Arith.*, to appear.
[32] L. Vinh, *Szemerédi-Trotter type theorems and sum-product estimates in finite fields*, *Eur. J. Combin.*, **32** (2011), 1177–1181.
[33] I. M. Vinogradov, *An introduction to the theory of numbers*, *Pergamon Press*, London and New York, 1955.
[34] V. Vu, *Sum-product estimates via directed expanders*, *Math. Res. Lett.*, **15** (2008), 375–388.
[35] B. Xue, *Asymmetric estimates and the sum-product problems*, *Acta Arith.*, to appear.

A.M.: School of Mathematics, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran.

Email address: `a.mohammadi@ipm.ir`

S.S.: Johannn Radon Institute for Computational and Applied Mathematics (RICAM), Linz, Austria

Email address: `sophie.stevens@oeaw.ac.at`