IMPROVED ESTIMATES FOR POLYNOMIAL ROTH TYPE
THEOREMS IN FINITE FIELDS

By

DONG DONG, XIAOCHUN LI AND WILL SAWIN

Abstract. We prove that, under certain conditions on the function pair \( \varphi_1 \) and \( \varphi_2 \), the bilinear average \( q^{-1} \sum_{y \in \mathbb{F}_q} f_1(x + \varphi_1(y))f_2(x + \varphi_2(y)) \) along the curve \( (\varphi_1, \varphi_2) \) satisfies a certain decay estimate. As a consequence, Roth type theorems hold in the setting of finite fields. In particular, if \( \varphi_1, \varphi_2 \in \mathbb{F}_q[X] \) with \( \varphi_1(0) = \varphi_2(0) = 0 \) are linearly independent polynomials, then for any \( \mathcal{A} \subseteq \mathbb{F}_q, |\mathcal{A}| = \delta q \) with \( \delta > cq^{-1/12} \), there are \( \gg \delta^3 q^2 \) triplets \( x, x+\varphi_1(y), x+\varphi_2(y) \in \mathcal{A} \). This extends a recent result of Bourgain and Chang who initiated this type of problems, and strengthens the bound in a result of Peluse, who generalized Bourgain and Chang’s work. The proof uses discrete Fourier analysis and algebraic geometry.

1 Introduction

Let \( \mathbb{F}_q \) be the finite field with \( q = p^m \) elements (\( p \) is a prime). For any

\[
\varphi_1, \varphi_2 : \mathbb{F}_q \to \mathbb{F}_q,
\]

we are interested in the bilinear average operator along the “curve” \( \Gamma = (\varphi_1, \varphi_2) \):

\[
\mathcal{A}_\Gamma(f_1, f_2)(x) := \frac{1}{q} \sum_{y \in \mathbb{F}_q} f_1(x + \varphi_1(y))f_2(x + \varphi_2(y)).
\]

To state our main result, we first set up some notations. Let \( e_p(x) := e^{2\pi i x^2/p} \), and \( \psi_p(x) := e_p(Tr(x)) \), where \( Tr(x) := x + x^p + \cdots + x^{p^{m-1}} \) is the trace function. For \( f : \mathbb{F}_q \to \mathbb{C} \), define

\[
\mathbb{E}[f] = \mathbb{E}_\chi[f] = \frac{1}{q} \sum_{x \in \mathbb{F}_q} f(x), \quad \hat{f}(z) = \frac{1}{q} \sum_{x \in \mathbb{F}_q} f(x) \psi_p(-xz),
\]

\[
\|f\|_r = \left( \frac{1}{q} \sum_{x \in \mathbb{F}_q} |f(x)|^r \right)^{1/r}, \quad \|f\|_r = \left( \sum_{x \in \mathbb{F}_q} |f(x)|^r \right)^{1/r}.
\]
With these notations, it is easy to verify that
\[ \|f\|_r \leq \|f\|_s \quad \text{if } s > r \quad (\text{a special case of the Hölder inequality}); \]
\[ \|\hat{f}\|_2 = \|f\|_2 \quad (\text{Parseval}); \]
\[ f(x) = \sum_{z \in \mathbb{F}_q} \hat{f}(z) \psi_q(xz) \quad (\text{Fourier inversion}). \]

We also need a notion of generalized diagonal sets.

**Definition.** A set \( D \subseteq \mathbb{F}_q \times \mathbb{F}_q \) is called \( B \)-**generalized diagonal** (or simply **generalized diagonal**) when the parameter \( B \) is unimportant) if for any \( x \in \mathbb{F}_q \), there are at most \( B \) \( y \)'s such that \( (x, y) \in D \) and for any \( y \in \mathbb{F}_q \) there are at most \( B \) \( x \)'s such that \( (x, y) \in D \).

For example, the diagonal set \( \{ (x, x) : x \in \mathbb{F}_q \} \) is a 1-generalized diagonal set. Throughout this paper, the parameter \( B \) will be independent of the finite field.

Our main theorem below provides a framework to obtain a decay estimate for the bilinear operator \( A_\Gamma \) associated with various function pairs \( (\varphi_1, \varphi_2) \). Throughout the paper, \( A \lesssim B \) denotes the statement that \( |A| \leq C|B| \) for some positive constant \( C \). The dependence of \( C \) will be clear in context.

The behavior of the bilinear average is closely related to the following exponential sum associated to \( \Gamma \):
\[
(1.2) \quad K_\Gamma(x, y) := \begin{cases} 
\frac{1}{q} \sum_{z \in \mathbb{F}_q} \psi_q(x \varphi_1(z) + y \varphi_2(z)), & y \neq 0; \\
0, & y = 0.
\end{cases}
\]

**Theorem 1.1.** Given any \( \varphi_1, \varphi_2 : \mathbb{F}_q \to \mathbb{F}_q \), let the kernel \( K_\Gamma \) be as in (1.2). We define for any \( h, y, y' \in \mathbb{F}_q \),
\[
(1.3) \quad I_\Gamma(y, y', h) := \sum_{x \in \mathbb{F}_q} K_\Gamma(x, y)K_\Gamma(x - h, y + h)K_\Gamma(x, y')K_\Gamma(x - h, y' + h).
\]

Suppose that the following three conditions hold:
1. There exist \( \theta \in (0, 1] \) and \( c_1 > 0 \) such that \( \frac{1}{p} \sum_{y \in \mathbb{F}_q} \psi_q(s \varphi_1(y)) \leq c_1 q^{-\theta} \) for any \( s \neq 0 \).
2. There exist \( \alpha \in \left( \frac{1}{4}, 1 \right) \) and \( c_2 > 0 \) such that \( K_\Gamma(x, y) \leq c_2 q^{-\alpha} \) for any \( x, y \in \mathbb{F}_q \).
3. There exist \( \beta > 1, B \geq 1 \) and \( c_3 > 0 \) such that for any \( h \in \mathbb{F}_q^* := \mathbb{F}_q \setminus \{0\} \) we can find a \( B \)-generalized diagonal set \( D_{\Gamma, h} \) so that \( I_\Gamma \leq c_3 q^{-\beta} \) for any \( (y, y') \notin D_{\Gamma, h} \).

Then the bilinear average defined by (1.1) obeys
\[
(1.4) \quad \|A_\Gamma(f_1, f_2) - \mathbb{E}[f_1] \mathbb{E}[f_2]\|_2 \leq C q^{-\gamma} \|f_1\|_2 \|f_2\|_2
\]
with \( \gamma = \min\{\theta, \alpha - \frac{1}{4}, \beta - \frac{1}{4} \} \) and the positive constant \( C \) depending only on \( c_1, c_2, c_3, B \).
**Remark 1.2.** Note that $\varphi_1$ and $\varphi_2$ play different roles in the three conditions in the above theorem. However, when checking these conditions, one can always swap $\varphi_1$ and $\varphi_2$ if necessary, as the operator $A_\Gamma$ is symmetric in $\varphi_1$ and $\varphi_2$ (and so is the conclusion (1.4)).

Motivated by the non-conventional ergodic averages considered by Bergelson [1] and Frantzikinakis and Kra [9], Bourgain and Chang [4] are the first to consider quantitative estimate of the form (1.4) when $F_q$ is a prime field. They established (1.4) with $\gamma = \frac{1}{10}$ for the quadratic monomial curve $\Gamma = (y, y^2)$, via an elegant way combining discrete Fourier analysis, explicit evaluation of quadratic Gauss sum and Bombieri’s estimate for character sums with rational function arguments [3].

Peluse [15, Theorem 2.2] generalized Bourgain and Chang’s result to the polynomial curve $(\varphi_1(y), \varphi_2(y))$ for any linearly independent polynomials $\varphi_1, \varphi_2$. Her result also applies over arbitrary finite fields of large characteristic (and not just $F_p$). However, she must take $\gamma = 1/16$. Her method is based on careful analysis of the dimension of varieties created by multiple applications of Cauchy–Schwarz, and an exponential sum bound due to Kowalski (itself a direct application of Deligne’s estimate [6]).

Our result improves the decay rate from $\frac{1}{10}$ to $\frac{1}{8}$ in Peluse’s bound. This also improves the decay rate from $\frac{1}{10}$ to $\frac{1}{8}$ in the cases handled by Bourgain and Chang.

**Remark 1.3.** In the special case $\Gamma = (y, y^2)$ (and $F_q = F_p$) treated by Bourgain–Chang, our approach does not invoke Bombieri’s estimate as in Bourgain and Chang’s method. Indeed, when $\Gamma = (y, y^2)$, $K_\Gamma(x, y)$ is a quadratic Gauss sum [11], which can be evaluated explicitly as

$$K_\Gamma(x, y) = c_y \frac{1}{\sqrt{p}} e_p \left(-\frac{x^2}{4y}\right)$$

for some $|c_y| \leq 1$. Condition (3) can therefore be verified by

$$|I_\Gamma| \leq \frac{1}{p^2} \left| \sum_x e_p \left(-\frac{x^2}{4y}\right) e_p \left(\frac{(x-h)^2}{4(y+h)}\right) e_p \left(-\frac{(x-h)^2}{4(y'+h)}\right) \right| \leq p^{-\frac{3}{4}}$$

for $y \neq y'$, using the square root cancellation of quadratic Gauss sum again. Hence $\beta = \frac{3}{4}$. It is easy to check that $\theta = \alpha = \frac{1}{2}$, and thus $\gamma = \frac{1}{8}$.

To extend to the polynomial curve $\Gamma = (y, P(y))$, the condition (3) can be verified by Deligne’s fundamental work on exponential sums over finite fields [6]. When extending to the bi-polynomial case, we need to use Katz’s generalisation [13] of Deligne’s theorem on exponential sums over smooth affine varieties. Details will be given in Section 3.

Theorem 1.1 immediately implies a quantitative Roth type theorem:
Corollary 1.4. Let \( \phi_1, \phi_2 : \mathbb{F}_q \to \mathbb{F}_q \) be functions satisfying conditions (1), (2) and (3) (with parameters \( \theta, \alpha \) and \( \beta \), respectively) of Theorem 1.1. Then for any \( A \subseteq \mathbb{F}_q, |A| = \delta q \) with \( \delta > cq^{-\frac{5}{8}} \gamma \), \( \gamma = \min\{\theta, \alpha - \frac{1}{4}, \frac{\beta}{4} - \frac{1}{4}\} \), there are \( \gtrsim \delta^3 q^2 \) triplets \( x, x + \phi_1(y), x + \phi_2(y) \in A \).

We include its short proof (which is the same as that of Corollary 1.2 in [4]) here for the reader’s convenience. Indeed, given \( A \subseteq \mathbb{F}_q \), set \( f = f_1 = f_2 \) to be the indicator function of the set \( A \). Then \( \mathbb{E} f = \delta \) and \( \|f\|_2 = \delta^2 \). By Cauchy–Schwarz inequality and (1.4),

\[
\frac{1}{q^2} \sum_{x,y} f(x)f(x + \phi_1(y))f(x + \phi_2(y))
= \mathbb{E}_x(fA_\Gamma(f,f))
= \mathbb{E}_x[(\mathbb{E}f)^2 + f(A_\Gamma(f,f) - (\mathbb{E}f)^2)] \geq (\mathbb{E}f)^3 - \|f(A_\Gamma(f,f) - (\mathbb{E}f)^2)\|_1
\geq (\mathbb{E}f)^3 - \|f\|_2 \|A_\Gamma(f,f) - (\mathbb{E}f)^2\|_2 \geq (\mathbb{E}f)^3 - cq^{-\gamma}\|f\|_2^3
= \delta^3 - cq^{-\gamma}\delta^2 \gtrsim \delta^3,
\]

from which the corollary follows.

One interesting case of Theorem 1.1 is when \( \phi_1 \) and \( \phi_2 \) are polynomials:

Theorem 1.5. Let \( \Gamma = (\phi_1, \phi_2) \) with \( \phi_1, \phi_2 \in \mathbb{F}_q[X], \phi_1(0) = \phi_2(0) = 0 \). Suppose that \( \phi_1, \phi_2 \) are linearly independent. When \( \mathbb{F}_q \neq \mathbb{F}_p \), i.e., \( m \neq 1 \), assume further that \( p \) is greater than the degrees of \( \phi_1 \) and \( \phi_2 \). Then the average function \( A_\Gamma \) satisfies

(1.5) \[ \|A_\Gamma(f_1, f_2) - \mathbb{E}[f_1] \mathbb{E}[f_2]\|_2 \lesssim q^{-1/8} \|f_1\|_2 \|f_2\|_2, \]

with the implied constant depending only on the degrees of \( \phi_1 \) and \( \phi_2 \).

As before, we can obtain the corresponding Roth type theorem in which the lower bound \( p^{- \frac{1}{3}} \) of \( \delta \) is slightly better than the bound \( p^{- \frac{1}{4}} \) obtained in Bourgain and Chang’s paper [4].

Corollary 1.6. Let \( \phi_1, \phi_2 \in \mathbb{F}_q[X], \phi_1(0) = \phi_2(0) = 0 \), be linearly independent. When \( \mathbb{F}_q \neq \mathbb{F}_p \), i.e., \( m \neq 1 \), assume further that \( p \) is greater than the degrees of \( \phi_1 \) and \( \phi_2 \). Then for any \( A \subseteq \mathbb{F}_q, |A| = \delta q \) with \( \delta > cq^{-\frac{1}{12}} \), there are \( \gtrsim \delta^3 q^2 \) triplets \( x, x + \phi_1(y), x + \phi_2(y) \in A \).

Remark 1.7. In Theorem 1.5 and Corollary 1.6, the condition \( \phi_1 \) and \( \phi_2 \) are linearly independent without constant terms” can be weakened to “\( \phi_1 - \phi_1(0) \) and \( \phi_2 - \phi_2(0) \) are linearly independent”. The assumption “\( p \) is greater than the degrees of \( \phi_1 \) and \( \phi_1 \)” can also be weakened to “the degree of \( x\phi_1 + y\phi_2 \) is prime to \( p \) for any \( x, y \in \mathbb{F}_q \)”. See Section 3 for the detailed arguments.
**Remark 1.8.** The size of the largest progression free set of $\mathbb{F}_p^m$ is of the order $(cp)^m$ for some $c < 1$ [7] (see [2] for the independence of $p$). Viewing $\mathbb{F}_q$ as $\mathbb{F}_p^m$, Corollary 1.6 provides an upper bound for a polynomial progression free set of $\mathbb{F}_p^m$, $(p^{\frac{1}{12}})^m$, which is much smaller than the order of the largest linear progression free set. This is another feature of nonlinear progressions.

**Remark 1.9.** Our results cover some rational function pairs $(\varphi_1, \varphi_2)$, if not all, with obvious modifications in the domain of the summations (i.e., exclude the poles). For instance, when $\varphi_1(y) = y$, $\varphi_2(y) = \frac{1}{y}$ and $\mathbb{F}_q = \mathbb{F}_p$ (this case is also considered in [4]), we get the same conclusion as in Theorem 1.5 and Corollary 1.6. Indeed, in this case $K(x, y) = \frac{1}{\sqrt{p}} Kl(xy)$, where

$$Kl(a) := \frac{1}{\sqrt{p}} \sum_{z \neq 0} e_p(az + \frac{1}{z})$$

is a Kloosterman sum, which is bounded by 2. The estimate (for “off-diagonal” $y, y'$)

$$I_\Gamma = \sum_{x \in \mathbb{F}_p} p^{-2} Kl(xy) Kl((x-h)(y+h)) Kl(xy') Kl((x-h)(y'+h)) \lesssim p^{-\frac{3}{2}}$$

follows from Corollary 3.3 in [8]. Therefore, the three conditions in Theorem 1.1 are satisfied. For a general rational function pair $(\varphi_1, \varphi_2)$, the conditions in Theorem 1.1, especially condition (3), are not easy to check.

**Remark 1.10.** Theorem 1.5, in the case $\varphi_1(x) = x$, implies that the polynomial $x + \varphi_2(y - x)$ is an almost strong asymmetric expander in the sense of Tao’s paper [18]. It is possible that this result could also be established using [18, Theorem 3], but we do not pursue this.

**Remark 1.11.** We refer the reader to [16] for an earlier study of the existence of patterns in subsets of finite fields.

We will prove Theorem 1.1 in Section 2. In Section 3 we will verify the three conditions (1), (2), and (3) for certain polynomial pairs and henceforth prove Theorem 1.5.

## 2 Proof of Theorem 1.1

We prove the main theorem in this section. We follow the spirit in the second author’s work on the bilinear Hilbert transform along curves in [14]. First, by using Fourier inversion for $f_1$ and $f_2$, it is clear that

$$A_\Gamma(f_1, f_2)(x) = \sum_{n_1, n_2} \hat{f}_1(n_1) \hat{f}_2(n_2) \psi_q((n_1 + n_2)x) E_q[\psi_q(n_1 \varphi_1(y) + n_2 \varphi_2(y))].$$
Changing variables \( n_2 = n, n_1 = s - n \), we then split the bilinear average \( A_\Gamma(f_1, f_2)(x) \) into three terms:

\[
A_\Gamma(f_1, f_2)(x) = J_1 + J_2 + J_3,
\]

where

\[
J_1 = \hat{f}_1(0) \hat{f}_2(0) = \mathbb{E}[f_1] \mathbb{E}[f_2],
\]

\[
J_2 = \hat{f}_2(0) \sum_{s \neq 0} \hat{f}_1(s) \mathbb{E}_y [\psi_q(s \varphi_1(y))] \psi_q(sx),
\]

\[
J_3 = \sum_{s} \left( \sum_{n \neq 0} \hat{f}_1(s - n) \hat{f}_2(n) \mathbb{E}_y [\psi_q((s - n) \varphi_1(y) + n \varphi_2(y))] \right) \psi_q(sx).
\]

By the assumption (1), when \( s \neq 0 \), we get

\[
(2.1) \quad \mathbb{E}_y [\psi_q(s \varphi_1(y))] = 1 \sum_{y} \psi_q(s \varphi_1(y)) \lesssim 1. \]

Therefore, using Parseval’s identity and the Cauchy–Schwarz inequality (to bound \( \hat{f}_2(0) \)), we see that

\[
\|A_\Gamma(f_1, f_2) - \mathbb{E}[f_1] \mathbb{E}[f_2]\|_2 \\
\leq \|\hat{J}_2\|_2 + \|\hat{J}_3\|_2 \\
\lesssim \frac{1}{q^\theta} \|f_1\|_2 \|f_2\|_2 + \left( \sum_{s} \left| \sum_{n} \hat{f}_1(s - n) \hat{f}_2(n) K_\Gamma(s - n, n) \right| \right)^{\frac{1}{2}},
\]

where \( K_\Gamma \) is given by (1.2).

Set \( \gamma_0 = \min\{ \alpha - \frac{1}{4}, \beta - \frac{1}{4} \} \). Hence it remains to show

\[
(2.2) \quad \sum_{s} \left| \sum_{n} \hat{f}_1(s - n) \hat{f}_2(n) K_\Gamma(s - n, n) \right| \lesssim \frac{1}{q^{2\gamma_0}} \|f_1\|_2^2 \|f_2\|_2^2.
\]

Next we choose to employ a \( TT^* \) method (our method and Bourgain–Chang’s diverge from here). The left-hand side of (2.2) equals

\[
\sum_{s} \sum_{n_1, n_2} \hat{f}_1(s - n_1) \hat{f}_1(s - n_2) \hat{f}_2(n_1) \hat{f}_2(n_2) K_\Gamma(s - n_1, n_1) K_\Gamma(s - n_2, n_2),
\]

which, after changing variables \( n_1 = v, n_2 = v + h, s = u + v \), can be rewritten as

\[
(2.3) \quad \sum_h \left( \sum_{u, v} F_h(u) G_h(v) K_\Gamma(u, v) K_\Gamma(u - h, v + h) \right),
\]
where $$F_h(u) = \hat{f}_1(u)f_1(u - h);$$
$$G_h(v) = \hat{f}_2(v)f_2(v + h).$$

When $$h = 0$$, using condition (2), we see that the inner double sum in (2.3) is bounded by
$$q^{-2\alpha} \|F_0\|_\ell \|G_0\|_\ell = q^{-2\alpha} \|f_1\|_2^2 \|f_2\|_2^2,$$
which is better than $$q^{-2\gamma_0} \|f_1\|_2^2 \|f_2\|_2^2$$ as $$\alpha > \gamma_0$$. Therefore, it remains to handle the case when $$h$$ is nonzero. The tool is the following bilinear form estimate, which may be interesting in its own right.

**Proposition 2.1.** Fix $$h \neq 0$$. Let $$\varphi_1, \varphi_2 : F_q \to F_q$$ satisfy (2) and (3) (with parameters $$\alpha$$ and $$\beta$$, respectively) of Theorem 1.1. Let $$\gamma_0 = \min\{\alpha - \frac{1}{4}, \beta - \frac{1}{4}\}$$.

Then for any $$F, G : F_q \to \mathbb{C},$$
\[
\sum_{u,v} F(u)G(v)K_\Gamma(u, v)K_\Gamma(u - h, v + h) \lesssim \frac{1}{q^{2\gamma_0}} \|F\|_\ell \|G\|_\ell.
\]

Once this proposition is proved, one can use (2.4) and apply the Cauchy–Schwarz inequality a few times to (2.3) to get the desired estimate (2.2).

It is easy to see that Proposition 2.1 can be reduced to the following finite field version of the Hörmander principle (see Theorem 1.1 in [10] for its continuous counterpart):

**Lemma 2.2.** Fix $$h \neq 0$$. Let $$\varphi_1, \varphi_2 : F_q \to F_q$$ satisfy (2) and (3) (with parameters $$\alpha$$ and $$\beta$$, respectively) in Theorem 1.1. Let $$\gamma_0 = \min\{\alpha - \frac{1}{4}, \beta - \frac{1}{4}\}$$.

Define an operator
$$T(g)(x) = \sum_{y \in F_q} g(y)K_\Gamma(x, y)K_\Gamma(x - h, y + h).$$

Then for all $$g : F_q \to \mathbb{C},$$
\[
\|T(g)\|_\ell \lesssim \frac{1}{q^{2\gamma_0}} \|g\|_\ell.
\]

**Proof.** We have
\[
\|T(g)\|_\ell^2 = \sum_{x, y, y'} g(y)g(y')K_\Gamma(x, y)K_\Gamma(x - h, y + h)K_\Gamma(x - h, y' + h)
\]
\[
\leq \sum_{(y, y') \in D_{\Gamma, h}} |g(y)||g(y')||I_\Gamma| + \sum_{(y, y') \notin D_{\Gamma, h}} |g(y)||g(y')||I_\Gamma|,
\]
where $D_{\Gamma,h}$ is the generalized diagonal set in condition (3) and
\[ I_{\Gamma} = \sum_{x} K_{\Gamma}(x,y)K_{\Gamma}(x-h,y+h)K_{\Gamma}(x,y')K_{\Gamma}(x-h,y'+h). \]

Using the definition of a generalized diagonal set and the trivial estimate $I_{\Gamma,1} \lesssim q^{4\alpha}$ from (2), the first term in (2.6) is estimated by
\[ \sum_{(y,y') \in D_{\Gamma,h}} |g(y)||g(y')|I_{\Gamma,1} \lesssim \sum_{y} |g(y)|^2 q^{4\alpha} = \frac{1}{q^{4\alpha-1}} \|g\|_2^2. \]

For the second term in (2.6), we use the assumption $I_{\Gamma,1} \lesssim \frac{1}{q^\beta}$ for $(y, y') \notin D_{\Gamma,h}$ and the Cauchy–Schwarz inequality to get the estimate
\[ \sum_{(y,y') \notin D_{\Gamma,h}} |g(y)||g(y')|I_{\Gamma,1} \lesssim \frac{\sqrt{q}\sqrt{q}}{q^\beta} \|g\|_2^2 = \frac{1}{q^{\beta-1}} \|g\|_2^2. \]

Combining (2.7) and (2.8), we obtain
\[ \|T(g)\|_2^2 \lesssim \left( \frac{1}{q^{4\alpha-1}} + \frac{1}{q^{\beta-1}} \right) \|g\|_2^2 \lesssim \frac{1}{q^{4\gamma_0}} \|g\|_2^2, \]
which is exactly what we aimed for: (2.5).

\[ \square \]

3 Proof of Theorem 1.5

To prove Theorem 1.5, first note that we can assume without loss of generality that the two polynomials $\varphi_1$ and $\varphi_2$ have distinct leading terms. This is because we can rewrite (1.5) as
\[ |E_{x,y}f_1(x + \varphi_1(y))f_2(x + \varphi_2(y))f_3(x) - E[f_1]E[f_2]E[f_3]| \lesssim q^{-1/8} \|f_1\|_2 \|f_2\|_2 \|f_3\|_2, \]
and do a change of variable $x \mapsto x + \varphi_1(y)$ on the left-hand side of (3.1) if necessary (we are indebted to Sarah Peluse for pointing this out).

We will verify that for linearly independent polynomials $\varphi_1, \varphi_2 \in \mathbb{F}_q[X]$ with distinct leading terms, the conditions (1), (2) and (3) of Theorem 1.1 are satisfied with parameters $\theta = \frac{1}{4}$, $\alpha = \frac{1}{2}$ and $\beta = \frac{3}{2}$, resp., and thus prove Theorem 1.5 using Theorem 1.1.

Let $d_1$ and $d_2$ denote the degrees of $\varphi_1$ and $\varphi_2$, respectively. Without loss of generality, we assume that $d_1 \leq d_2$.

Conditions (1) and (2) hold with $\theta = \alpha = \frac{1}{2}$, using the following theorem of Carlitz and Uchiyama [5] (which is a generalization of Weil [19]).
**Theorem 3.1.** Let $\mathbb{F}_q$ be the finite field with $q = p^n$ elements. Let $P \in \mathbb{F}_q[X]$ be a polynomial of degree $d$. Then

\begin{equation}
\sum_{z \in \mathbb{F}_q} \psi_q(P(z)) \leq (d - 1)q^{\frac{1}{2}},
\end{equation}

whenever $P \neq C^p - C + b$ for some $C \in \mathbb{F}_q[X]$ and $b \in \mathbb{F}_q$. In particular, (3.2) holds when $d$ is prime to $p$.

Indeed, condition (1) with $\theta = \frac{1}{2}$ immediately follows from the above theorem as the degree of $s\phi_1$ ($s \neq 0$) is prime to $p$. Similarly, as the polynomial $x\phi_1 + y\phi_2$ has degree prime to $p$ for any $x, y \in \mathbb{F}_q$, condition (2) is verified with $\alpha = \frac{1}{2}$. Note that the multiplicative constants depend only on the degrees of the polynomials.

Now we focus on the verification of condition (3). We will from now on write for simplicity $K = K_1$ and $I = I_1$. Recall that for $y \neq 0$,

$$K(x, y) = \frac{1}{q} \sum_{z \in \mathbb{F}_q} \psi_q(x\phi_1(z) + y\phi_2(z)).$$

We assume $y, y + h, y'$ and $y' + h$ are all nonzero (otherwise $I = 0$). Plug in the definition of $K$, put the sum over $x$ innermost, and we see that

$$I = \sum_{x \in \mathbb{F}_q} K(x, y)K(x - h, y + h)K(x, y')K(x - h, y' + h)$$

$$= \frac{1}{q^3} \sum_{x, z_1, z_2, z_3, z_4} \psi_q(x\phi_1(z_1) + y\phi_2(z_1) - (x - h)\phi_1(z_2) - (y + h)\phi_2(z_2) - x\phi_1(z_3)$$

$$- y'\phi_2(z_3) + (x - h)\phi_1(z_4) + (y' + h)\phi_2(z_4))$$

$$= \frac{1}{q^3} \sum_{z_1, z_2, z_3, z_4} \psi_q(F(z_1, z_2, z_3, z_4)),$$

where

$$G(z_1, z_2, z_3, z_4) = \phi_1(z_1) - \phi_1(z_2) - \phi_1(z_3) + \phi_1(z_4),$$

$$F(z_1, z_2, z_3, z_4) = y\phi_2(z_1) + h\phi_1(z_2) - (y + h)\phi_2(z_2) - y'\phi_2(z_3)$$

$$- h\phi_1(z_4) + (y' + h)\phi_2(z_4).$$

It remains to get the estimate

\begin{equation}
\sum_{z_1, z_2, z_3, z_4} \psi_q(F(z_1, z_2, z_3, z_4)) \lesssim q^{\frac{3}{2}}.
\end{equation}
We need machinery of algebraic geometry to prove (3.3). To benefit readers who are not very familiar with algebraic geometry, we first prove (3.3) in a simpler case. We assume \( \varphi_1(z) = z \), and consequently \( \varphi_2 \) has degree at least 2 by the linearly independence assumption. In this case, the restriction \( G(z_1, z_2, z_3, z_4) = 0 \) can be dropped once \( z_4 \) is replaced with \( z_2 + z_3 - z_1 \). Therefore, (3.3) is reduced to

\[
\sum_{z_1, z_2, z_3} \psi_q(F(z_1, z_2, z_3, z_2 + z_3 - z_1)) \lesssim q^{2}. \tag{3.4}
\]

Such a multidimensional character sum is studied by Deligne in his resolution of Weil conjectures:

**Theorem 3.2** ([6, Theorem 8.4]). Let \( f \in \mathbb{F}_q[X_1, \ldots, X_n] \) be a polynomial of degree \( d \geq 1 \). Suppose that \( d \) is prime to \( p \), and the projective hypersurface defined by the highest degree homogeneous term \( f_d \) is smooth, i.e., the gradient of \( f_d \) is nonzero at any point in \( \{f_d = 0\} \setminus \{0\} \). Then

\[
\sum_{z_1, \ldots, z_n} \psi_q(f(z_1, \ldots, z_n)) \leq (d - 1)^n q^{2}. 
\]

For notational convenience, we write \( d = d_2 \), the degree of \( \varphi_2 \). Let \( bz^d \) denote the leading term of \( \varphi_2(z) \). Then the highest degree homogeneous term of \( F(z_1, z_2, z_3, z_2 + z_3 - z_1) \) is

\[
F_d(z_1, z_2, z_3) = byz_1^d - b(y + h)z_2^d - by'z_3^d + b(y' + h)(z_2 + z_3 - z_1)^d.
\]

We need to verify the smoothness of \( \{F_d = 0\} \). By straightforward calculations, \( \nabla F_d = 0 \) implies

\[
\begin{cases}
    z_1 = \left( \frac{y' + h}{y} \right)^{\frac{1}{d}} (z_2 + z_3 - z_1), \\
    z_2 = \left( \frac{y' + h}{y' + h} \right)^{\frac{1}{d}} (z_2 + z_3 - z_1), \\
    z_3 = \left( \frac{y' + h}{y} \right)^{\frac{1}{d}} (z_2 + z_3 - z_1).
\end{cases}
\]

Recall that we have assumed \( y, y + h, y' \) and \( y' + h \) are all nonzero and thus the divisions make sense. The notation \( x^{1/n} \) means an element \( a \) in \( \mathbb{F}_q \) such that \( a^n = x \) (such \( a \) is usually not unique—we take the equation to be satisfied if it holds for any choice of \( n \)th root). The above system has nonzero solutions only when

\[
\left( \frac{y' + h}{y + h} \right)^{\frac{1}{d}} + \left( \frac{y' + h}{y'} \right)^{\frac{1}{d}} - \left( \frac{y' + h}{y} \right)^{\frac{1}{d}} = 1. \tag{3.5}
\]

The function

\[
\prod_{\zeta_1, \zeta_2, \zeta_3 \in \mu_{d-1}} \left( \zeta_1 \left( \frac{y' + h}{y + h} \right)^{\frac{1}{d}} + \zeta_2 \left( \frac{y' + h}{y'} \right)^{\frac{1}{d}} - \zeta_3 \left( \frac{y' + h}{y} \right)^{\frac{1}{d}} - 1 \right), \tag{3.6}
\]
where \( \mu_{d-1} \) are the \((d-1)\)-st roots of unity in \( \overline{\mathbb{F}}_q \), is a polynomial function in \((\frac{\gamma+y}{y})^\frac{1}{d-1}\), \((\frac{\gamma+y}{y})^\frac{1}{d-1}\), \((\frac{\gamma+y}{y})^\frac{1}{d-1}\), which is invariant under multiplying any of them by a \((d-1)\)-st root of unity, hence is a polynomial function in \((\frac{\gamma+y}{y}) \), \((\frac{\gamma+y}{y}) \) and thus a rational function in \(y, y', h\) (in fact, a polynomial divided by some power of \(yy'(y+h)\)). It vanishes if and only if equation (3.5) is satisfied for some choice of \((d-1)\)-st roots. Let the set of \((y, y')\) where this rational function vanishes be \(D_{\Gamma,h}\).

To check that \(D_{\Gamma,h}\) is a generalized diagonal set, it suffices to check that for any fixed \(y\) the rational function is not identically zero in \(y'\) and vice versa. To do this, note that the coefficient of \((\frac{\gamma+y}{y})^\frac{1}{d-1}\) is nonzero and every other term with nonzero coefficient has a lower pole order at \(y' = 0\), so the rational function in \(y'\) has a nontrivial pole and is nonzero. Similarly, the coefficient of \((\frac{\gamma+y}{y})^\frac{1}{d-1}\) is nonzero and every other term has a lower pole order at \(y = 0\) (because \(h \neq 0\) so \(\frac{\gamma+y}{y} \) does not have a pole there), so the rational function in \(y\) has a nontrivial pole and is nonzero. By Deligne’s Theorem, (3.4) holds for any \((y, y') \notin D_{\Gamma,h}\). This finishes the verification of condition (3) with \(\beta = \frac{3}{2}\), assuming \(\phi_1(z) = z\).

Now we turn to the general case. In [12], Katz generalizes Deligne’s theorem to exponential sums over smooth affine varieties, and in [13], to singular algebraic varieties. We need the following special case of [13, Theorem 4].

**Theorem 3.3.** Let \(F, G \in \overline{\mathbb{F}}_q[X_1, \ldots, X_4]\). Assume that the degree of \(F\) is indivisible by \(p\), the homogeneous leading term of \(G\) defines a smooth projective hypersurface, and the homogeneous leading terms of \(G\) and that of \(F\) together define a smooth co-dimension-2 variety in the projective space. Then (3.3) holds, i.e.,

\[
\sum_{G(z_1, z_2, z_3, z_4) = 0} \psi_q(F(z_1, z_2, z_3, z_4)) \leq q^2.
\]

The proof of this theorem will be given in the Appendix (the reader could skip its long proof and use it as a “black box” on an early reading of the paper).

Now we are ready to prove (3.3) using Theorem 3.3. The first two conditions in the theorem are easy to check. To check the third condition, we handle two cases separately: \(d_1 < d_2\) and \(d_1 = d_2\).

First assume \(d_1 < d_2\). Let \(a z_1^{d_1}\) and \(b z_2^{d_2}\) denote the leading term of \(\phi_1\) and \(\phi_2\), respectively. The homogeneous leading terms of \(G\) and \(F\) are

\[G_{d_1}(z_1, z_2, z_3, z_4) := a z_1^{d_1} - a z_2^{d_1} - a z_3^{d_1} + a z_4^{d_1},\]

and

\[F_{d_2}(z_1, z_2, z_3, z_4) := b y z_1^{d_2} - b(y+h) z_2^{d_2} - by' z_3^{d_2} + b(y+h) z_4^{d_2},\]
respectively. We need to show that the Jacobian matrix
\[
J = \begin{bmatrix}
\nabla G_{d_1} \\
\nabla F_{d_2}
\end{bmatrix}
= \begin{bmatrix}
d_1 a z_1^{d_1-1} & -d_1 a z_2^{d_1-1} & -d_1 a z_3^{d_1-1} & d_1 a z_4^{d_1-1} \\
d_2 b y z_1^{d_2-1} & -d_2 b (y + h) z_2^{d_2-1} & -d_2 b y' z_3^{d_2-1} & d_2 b (y' + h) z_4^{d_2-1}
\end{bmatrix}
\]
has full rank at any point in \( \{ G_{d_1} = F_{d_2} = 0 \} \setminus \{ 0 \} \). When \( J \) has rank less than 2, assuming \( z_1, z_2, z_3, z_4 \neq 0 \), we can solve for each \( z_i \) and plug in \( G_{d_1} = 0 \) to get the equation
\[
(3.7) \quad \left( \frac{1}{y} \right) \frac{d_1}{z_2^{d_1-1}} - \left( \frac{1}{y + h} \right) \frac{d_1}{z_2^{d_1-1}} - \left( \frac{1}{y'} \right) \frac{d_1}{z_2^{d_1-1}} + \left( \frac{1}{y' + h} \right) \frac{d_1}{z_2^{d_1-1}} = 0.
\]
If one or two of the four variables \( z_1, z_2, z_3, z_4 \) are zero, then a new equation can be obtained by deleting the corresponding term(s) in the above equation. By the same arguments as before, the solutions to (3.7) and its variants lie in a generalized diagonal set. So we can apply Theorem 3.3 for pairs \((y, y')\) outside this set.

Secondly, consider the case \( d_1 = d_2 = d \). The homogeneous leading terms of \( G \) and \( F \) are
\[
G_d(z_1, z_2, z_3, z_4) := a z_1^d - a z_2^d - a z_3^d + a z_4^d,
\]
and
\[
F_d(z_1, z_2, z_3, z_4) := b y z_1^d - (b(y + h) - ah) z_2^d - b y' z_3^d + (b(y' + h) - ah) z_4^d,
\]
respectively. The Jacobian matrix becomes
\[
J = \begin{bmatrix}
\nabla G_d \\
\nabla F_d
\end{bmatrix}
= \begin{bmatrix}
daz_1^{d-1} & -daz_2^{d-1} & -daz_3^{d-1} & daz_4^{d-1} \\
dbyz_1^{d-1} & -d(b(y + h) - ah) z_2^{d-1} & -dby' z_3^{d-1} & d(b(y' + h) - ah) z_4^{d-1}
\end{bmatrix}.
\]
When \( z_1 z_2 z_3 z_4 \neq 0 \), \( J \) has rank 1 only when
\[
(3.8) \quad by = b(y + h) - ah = by' = b(y' + h) - ah.
\]
One or two terms in the above equation can be dropped if the corresponding variable is zero. Since we assume that \( \phi_1 \) and \( \phi_2 \) have distinct leading terms, \( a \neq b \). It is then easy to see that the solutions to (3.8) and its variants form a generalized diagonal set. So Theorem 3.3 applies again for pairs \((y, y')\) outside this generalized diagonal set, and we are done.
Acknowledgments  The first author was supported by LTS grant DO 0052. He would like to thank Ping Xi, Dingxin Zhang and Sarah Peluse for many helpful discussions. The third author was supported by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zurich Foundation. The authors thank Ben Green and Kannan Soundararajan for bringing Peluse’s paper [15] to their attention.

4 Appendix

We explain in detail how to realize Theorem 3.3 as a special case of Katz’s theorem. We will try to explain this derivation for mathematicians who are not experts in algebraic geometry. (However, Katz’s proof requires much more advanced algebraic geometry than we can go into here.)

We first restate part of Katz’s theorem. Then we will explain Katz’s notation and how it applies to our case.

Theorem (Katz, [13, Theorem 4]). Let $N$ and $d$ be natural numbers, let $k$ be a finite field in which $d$ is invertible, let $\psi : k \to \mathbb{C}^\times$ be an additive character. Let $X$ be a closed subscheme of $\mathbb{P}^N$ of dimension $d$. Let $L$ be a section of $H^0(X, \mathcal{O}(1))$ and $H$ a section of $H^0(X, \mathcal{O}(D))$. Let $V, f, \epsilon, \delta$ be defined as in [13, pp. 878–879]. If assumptions (H1)' and (H2) of [13, p. 878] hold, and $\epsilon \leq \delta$, then

$$\left| \sum_{x \in V(k)} \psi(f(x)) \right| \leq C \times (\#k)^{(n+1+\delta)/2}$$

where $C$ is a constant depending only on $N, d$, and the number and degree of the equations defining $X$.

We will choose our data so that $k = \mathbb{F}_q$.

$$V(k) = \{ z_1, z_2, z_3, z_4 \in \mathbb{F}_p \mid G(z_1, z_2, z_3, z_4) = 0 \},$$
$$\psi(f(x)) = \psi_q(F(z_1, z_2, z_3, z_4)) \quad \text{for} \quad x = (z_1, z_2, z_3, z_4) \in V(k), \quad n = 3,$$

and $\epsilon = \delta = -1$. Furthermore, $C$ will depend only on the degree of $F$ and $G$.

Examining Katz’s bound, and plugging in these statements, it is clear that if we can in fact choose our data in this way, while verifying Katz’s conditions, we obtain exactly our stated bound.

In what remains, we will first explain all of Katz’s notation that is needed to choose $(X, L, H)$ so that

$$V(k) = \{ z_1, z_2, z_3, z_4 \in \mathbb{F}_p \mid G(z_1, z_2, z_3, z_4) = 0 \} \quad \text{and} \quad \psi(f(x)) = \psi_q(F(z_1, z_2, z_3, z_4)),$$
and second we will verify (H1)’ and (H2) and calculate $\epsilon$, $\delta$, explaining more of Katz’s notation along the way.

We take $N = 4$, so $\mathbb{P}^N = \mathbb{P}^4$, which we view as the projective space with coordinates $z_1, z_2, z_3, z_4, z_5$. We let $\tilde{G}$ be the homogenization of $G$, where we add additional powers of $z_5$ to all the non-leading terms of $G$ to make every term have equal degree. Let $X$ be the vanishing scheme of $\tilde{G}$. We must choose $L$ as an element of $H^0(X, \mathcal{O}(1))$, which is the space of linear functions in the variables $z_1, z_2, z_3, z_4, z_5$, and we choose $L = z_5$. Now Katz defines $V$ to be the locus in $X$ where $L$ is nonzero. Let us check that

$$V(\mathbb{F}_q) = \{ z_1, z_2, z_3, z_4 \in \mathbb{F}_q \mid G(z_1, z_2, z_3, z_4) = 0 \}.$$

To do this, recall that $\mathbb{P}^4(\mathbb{F}_q)$ is the set of quintuples $(z_1, z_2, z_3, z_4, z_5) \in \mathbb{F}_q$, not all zero, up to multiplication by nonzero scalars. By the definition of $X$, $X(\mathbb{F}_q)$ is the subset of $\mathbb{P}^4(\mathbb{F}_q)$ consisting of tuples $(z_1, z_2, z_3, z_4, z_5)$ with $\tilde{G}(z_1, z_2, z_3, z_4, z_5) = 0$. Hence $V(\mathbb{F}_q)$ is the set of tuples $(z_1, z_2, z_3, z_4, z_5)$, with $z_5$ nonzero, up to scalar multiplication, that solve the equation $\tilde{G}(z_1, z_2, z_3, z_4, z_5) = 0$. For each such tuple there exists a unique scalar multiplication that sends $z_5$ to 1, so we can express it equally as the set of tuples $(z_1, z_2, z_3, z_4)$ with $\tilde{G}(z_1, z_2, z_3, z_4, 1) = 0$. By construction,

$$\tilde{G}(z_1, z_2, z_3, z_4, 1) = G(z_1, z_2, z_3, z_4),$$

so

$$V(\mathbb{F}_q) = \{ z_1, z_2, z_3, z_4 \in \mathbb{F}_q \mid G(z_1, z_2, z_3, z_4) = 0 \},$$

as desired.

Next, because

$$\psi_q : \mathbb{F}_q \to \mathbb{C}^\times$$

is an additive character, we set $\psi = \psi_q$. We then need to choose $H$, a homogeneous form of degree $d$ in the variables $z_1, z_2, z_3, z_4, z_5$, so that $f(x) = F(z_1, z_2, z_3, z_4)$. Katz defines $f$ as $H/L^d$. We take $d$ to be the degree of $F$ and $H$ to be the homogenization $\tilde{F}$ of $F$, just as we did with $G$. Because we are using the bijection between 4-tuples and 5-tuples that sends $(z_1, z_2, z_3, z_4)$ to $(z_1, z_2, z_3, z_4, 1)$, we need to check that $f(z_1, z_2, z_3, z_4, 1) = F(z_1, z_2, z_3, z_4)$. This follows because

$$f(z_1, z_2, z_3, z_4, 1) = \frac{\tilde{F}(z_1, z_2, z_3, z_4, 1)}{L(z_1, z_2, z_3, z_4, 1)^d} = \frac{\tilde{F}(z_1, z_2, z_3, z_4, 1)}{1^d} = F(z_1, z_2, z_3, z_4).$$

We have therefore shown how to specialize the left side of Katz’s bound to the left side of our own bound. It remains to check Katz’s assumptions and also the assumptions we made in applying Katz’s bound. These are as follows:
(1) \(d\) is invertible in \(k\).
(2) Katz’s assumption \((H1)^\prime\) holds.
(3) Katz’s assumption \((H2)\) holds.
(4) \(\delta = -1\).
(5) \(\epsilon = -1\).
(6) \(n = 3\).
(7) \(C\) depends only on the degree of \(F\) and \(G\).

The first condition follows because \(d\) is co prime to \(p\), hence invertible in \(\mathbb{F}_q\).

Katz’s assumption \((H1)^\prime\) is that \(X\) is Cohen–Macaulay and equidimensional of dimension \(n \geq 1\). Because \(H\) is the hypersurface defined by a single nontrivial equation \(\tilde{G} = 0\) in \(\mathbb{P}^4\), \(H\) has dimension 3. Thus \(H\) is a complete intersection, hence a locally complete intersection, thus by [17, Lemma 00SB] is Cohen–Macaulay. This verifies assumptions (2) and (6).

Katz defines \(C\) as an explicit function of his numerical data, which consists of \(N\), the number \(r\) of equations needed to define \(X\), the degrees of those equations, and \(d\). In our case \(N = 4\), \(r = 1\), the degree of the unique equation needed to define \(X\) is the degree of \(G\), and \(d\) is the degree of \(F\). Hence \(C\) is some explicit function of those degrees (assumption (7)).

Katz defines \(\epsilon\) as the dimension of the singular locus of the scheme-theoretic intersection \(X \cap L\). For us \(L\) is the closed subset of \(\mathbb{P}^4\) where \(z_5 = 0\). (Katz abuses notation slightly to use \(L\) also to refer to the vanishing locus of \(L\).) So \(X \cap L\) is the closed subset where \(z_5 = 0\) and \(\tilde{G} = 0\). Because \(z_5 = 0\), we can ignore \(z_5\) and work in \(\mathbb{P}^3\) with coordinates \(z_1, z_2, z_3, z_4\). When we do this, because all non-leading monomials of \(G\) were multiplied by a positive power of \(z_5\) in \(\tilde{G}\), all non-leading monomials become 0 and we are left with just the zero-locus. So \(X \cap L\) is the vanishing locus of the leading term of \(G\) in \(\mathbb{P}^3\), which we assumed in the statement of the theorem is a nonsingular hypersurface, so its singular locus is empty, to which by convention Katz assigns dimension \(-1\), verifying \(\epsilon = -1\) (assumption (5)).

Katz defines \(\delta\) as the dimension of the singular locus of the scheme-theoretic intersection \(X \cap L \cap H\), and \((H2)\) is his assumption that this has dimension \(n - 2\). This is the joint vanishing locus of \(\tilde{G}, z_5, \) and \(\tilde{F}\) in \(\mathbb{P}^4\), which for the same reason as before is the vanishing locus of the leading terms of \(F\) and \(G\) in \(\mathbb{P}^3\). Because we assumed this is a smooth subscheme of codimension 2, it has dimension \(3 - 2 = n - 2\), verifying condition \((H2)\), and its singular locus is empty and has dimension \(-1\), verifying \(\delta = -1\) (assumptions (3) and (4)).
REFERENCES

[1] V. Bergelson, Ergodic Ramsey theory—an update, in Ergodic Theory of $\mathbb{Z}^d$ Actions, Cambridge University Press, Cambridge, 1996, pp. 1–61.
[2] J. Blasiak, T. Church, H. Cohn, J. Grochow, E. Naslund, W. Sawin and C. Umans, On cap sets and the group-theoretic approach to matrix multiplication, Discrete Anal. (2017), paper no. 3.
[3] E. Bombieri, On exponential sums in finite fields, Amer. J. Math. 88 (1966), 71–105.
[4] J. Bourgain and M. C. Chang, Nonlinear Roth type theorems in finite fields, Israel J. Math. 221 (2017), 853–867.
[5] L. Carlitz and S. Uchiyama, Bounds for exponential sums, Duke Math. J. 24 (1957), 37–41.
[6] P. Deligne, La conjecture de Weil. I, Inst. Hautes Études Sci. Publ. Math. 43 (1974), 273–307.
[7] J. Ellenberg and D. Gijswijt, On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression, Ann. of Math. (2) 185 (2017), 339–343.
[8] É. Fouvry, E. Kowalski and P. Michel, A study in sums of products, Philos. Trans. Roy. Soc. A 373 (2015), Article ID 20140309.
[9] N. Frantzikinakis and B. Kra, Polynomial averages converge to the product of integrals, Israel J. Math. 148 (2005), 267–276.
[10] L. Hörmander, Oscillatory integrals and multipliers on $L^p$, Ark. Mat. 11 (1973), 1–11.
[11] H. Iwaniec and E. Kowalski, Analytic Number Theory, American Mathematical Society, Providence, RI, 2004.
[12] N. Katz, Sommes exponentielles, Société Mathématique de France, Paris, 1980.
[13] N. Katz, Estimates for “singular” exponential sums, Internat. Math. Res. Notices (1999), 875–899.
[14] X. Li, Bilinear Hilbert transforms along curves I: The monomial case, Anal. PDE 6 (2013), 197–220.
[15] S. Peluse, Three-term polynomial progressions in subsets of finite fields, Israel J. Math. 228 (2018), 379–405.
[16] I. D. Shkredov On monochromatic solutions of some nonlinear equations in $\mathbb{Z}/p\mathbb{Z}$, Mat. Zametki 88 (2010), 625–634; translation in Math. Notes 88 (2010), 603–611.
[17] Stacks project authors, The stacks project, https://stacks.math.columbia.edu (2018)
[18] T. Tao, Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets, Contr. to Disc. Math. 10 (2014), 22–98.
[19] A. Weil, On the Riemann hypothesis in function fields, Proc. Nat. Acad. Sci. USA. 27 (1941), 345–347.

Dong Dong
CENTER FOR SCIENTIFIC COMPUTATION AND MATHEMATICAL MODELING
UNIVERSITY OF MARYLAND
COLLEGE PARK, MD 20742, USA
e-mail: ddong12@cscammm.umd.edu

Xiaochun Li
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
URBANA, IL 61801, USA
e-mail: xcli@math.uiuc.edu
Will Sawin
DEPARTMENT OF MATHEMATICS
COLUMBIA UNIVERSITY
NEW YORK, NY 10027, USA
email: sawin@math.columbia.edu

(Received September 21, 2017 and in revised form November 11, 2018)