

# AN ARITHMETIC TRANSFERENCE PROOF OF A RELATIVE SZEMERÉDI THEOREM

YUFEI ZHAO

ABSTRACT. Recently, Conlon, Fox, and the author gave a new proof of a relative Szemerédi theorem, which was the main novel ingredient in the proof of the celebrated Green-Tao theorem that the primes contain arbitrarily long arithmetic progressions. Roughly speaking, a relative Szemerédi theorem says that if  $S$  is a set of integers satisfying certain conditions, and  $A$  is a subset of  $S$  with positive relative density, then  $A$  contains long arithmetic progressions, and our recent results show that  $S$  only needs to satisfy a so-called linear forms condition.

This note contains an alternative proof of the new relative Szemerédi theorem, where we directly transfer Szemerédi's theorem, instead of going through the hypergraph removal lemma. This approach provides a somewhat more direct route to establishing the result, and it gives better quantitative bounds.

The proof has three main ingredients: (1) a transference principle/dense model theorem of Green-Tao and Tao-Ziegler (with simplified proofs given later by Gowers, and independently, Reingold-Trevisan-Tulsiani-Vadhan) applied with a discrepancy/cut-type norm (instead of a Gowers uniformity norm as it was applied in earlier works), (2) a counting lemma established by Conlon, Fox, and the author, and (3) Szemerédi's theorem as a black box.

## 1. INTRODUCTION

The celebrated Green-Tao theorem [6] states that the primes contain arbitrarily long arithmetic progressions (AP). A key ingredient in their work is a relative Szemerédi theorem. Szemerédi's theorem [14] states that any subset of the integers with positive upper density contains arbitrarily long arithmetic progressions. A relative Szemerédi theorem is a result where the ground set is no longer  $\mathbb{Z}$  but some sparse pseudorandom subset (or more generally some measure).

Green and Tao proved a relative Szemerédi theorem provided that the ground set satisfies certain pseudorandomness conditions known as the linear forms condition and the correlation condition. They then constructed a majorizing measure to the primes, using ideas from the work of Goldston and Yıldırım [2] (subsequently simplified in [15]), so that this majorizing measure satisfies the desired pseudorandomness conditions.

Recently, Conlon, Fox, and the author [1] gave a new proof of Green and Tao's relative Szemerédi theorem, requiring simpler pseudorandomness hypotheses on the ground set. We showed that a weak version of Green and Tao's linear forms condition is sufficient. A precise definition of our linear forms condition will be given in §2.

In [1], we obtained a relative extension of the famous hypergraph removal lemma [4, 8, 11, 12, 17], from which we deduced our relative Szemerédi theorem via standard arguments. Such an approach was also taken by Tao [16] in his work on constellations in the Gaussian primes, but our approach in [1] requires less stringent pseudorandomness hypotheses.

In this note, we give an alternative approach to proving the relative Szemerédi theorem in [1]. Instead of going through the hypergraph removal lemma, we use Szemerédi's theorem directly as a black box. To transfer Szemerédi's theorem to the sparse setting, we apply the dense model theorem of Green-Tao [6] and Tao-Ziegler [18], which was subsequently simplified by Gowers [5], and independently Reingold, Trevisan, Tulsiani, and Vadhan [10]. This tool lets us model a subset of a

---

This work was done while the author was an intern at Microsoft Research New England.

sparse pseudorandom set of integers by a dense subset. The dense model is a good approximation of the original set with respect to a discrepancy-type norm (similar to the cut metric for graphs). This contrasts previous proofs the Green-Tao theorem [6, 5, 10] where the dense model theorem is applied with respect to the Gowers uniformity norm, which gives a stronger notion of approximation.

Another important ingredient in the proof is the relative counting lemma of [1], which implies that the dense model behaves similarly to the original set in the number of arithmetic progressions.

The arithmetic transference approach presented here establishes the new relative Szemerédi theorem in a more direct fashion. It also gives better quantitative bounds than [1]. Indeed, instead of going through the hypergraph removal lemma, which currently has an Ackermann-type dependence on the bounds (due to the application of the hypergraph regularity lemma), we can now use Szemerédi's theorem as a black box and automatically transfer the best quantitative bounds available (currently the state-of-art is [13] for 3-term APs, [7] for 4-term APs, and [3] for longer APs). This answers a question that was left open in [1]. The approach presented here, however, is less general compared to [1], since it does not provide a relative hypergraph removal lemma, nor does it give a more general sparse regularity approach to hypergraphs.

The main theorem is stated in §2. In §3, we apply a dense model theorem to find a dense approximation of the original set. In §4, we apply a counting lemma to show that the dense model has approximately the same number of  $k$ -term APs as the original set. Finally in §5, we put everything together and apply Szemerédi's theorem as a black box to conclude the proof.

## 2. DEFINITIONS AND RESULTS

**Notation.** *Dependence on  $N$ .* We consider functions  $\nu = \nu^{(N)}$ , where  $N$  (usually suppressed) is assumed to be some large integer. We write  $o(1)$  for a quantity that tends to zero as  $N \rightarrow \infty$  along some subset of  $\mathbb{Z}$ . If the rate at which the quantity tends to zero depends on some other parameters (e.g.,  $k, \delta$ ), then we put these parameters in the subscript (e.g.,  $o_{k,\delta}(1)$ ).

*Expectation.* We write  $\mathbb{E}[f(x_1, x_2, \dots)|P]$  for the expectation of  $f(x_1, x_2, \dots)$  when the variables are chosen uniformly out of all possibilities satisfying  $P$ .

We shall use, as a black box, the following weighted version of Szemerédi's theorem as formulated, for example, in [6, Prop. 2.3]. It may be helpful to think of  $f$  as the indicator function  $1_A$  of some set  $A \subseteq \mathbb{Z}_N$ . It will be easier for us to work in  $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$  as opposed to  $[N] := \{1, \dots, N\}$ , although these two settings are easily seen to be equivalent.

**Theorem 2.1** (Szemerédi's theorem, weighted version). *Let  $k \geq 3$  and  $0 < \delta \leq 1$  be fixed. Let  $f: \mathbb{Z}_N \rightarrow [0, 1]$  be a function satisfying  $\mathbb{E}[f] \geq \delta$ . Then*

$$\mathbb{E}[f(x)f(x+d)f(x+2d)\cdots f(x+(k-1)d)|x, d \in \mathbb{Z}_N] \geq c(k, \delta) - o_{k,\delta}(1) \quad (1)$$

for some constant  $c(k, \delta) > 0$  which does not depend on  $f$  or  $N$ .

Gowers' results [3] (along with a Varnavides-type [19] averaging argument) imply that Theorem 2.1 holds with  $c(k, \delta) = \exp(-\exp(\delta^{-c_k}))$  with  $c_k = 2^{2^{k+9}}$  (see [13] and [7] for the current best bounds for  $k = 3$  and 4 respectively).

A *relative Szemerédi theorem* is an extension of Theorem 2.1 of the following form. Instead of  $0 \leq f \leq 1$  as in Theorem 2.1, we now assume that  $0 \leq f(x) \leq \nu(x)$  for all  $x \in \mathbb{Z}_N$ , where  $\nu: \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$  is some function (also called a majorizing measure) that satisfies certain pseudorandomness conditions. Here the function  $\nu$  is normalized so that  $\mathbb{E}[\nu] = 1 + o(1)$ . For instance, one can think of  $\nu$  as  $\frac{N}{|S|}1_S$  for some pseudorandom subset  $S \subseteq \mathbb{Z}_N$ , and  $f$  as  $1_A\nu$  with some  $A \subseteq S$ . So in this case (1) says that  $A$  contains many  $k$ -term APs when  $N$  is sufficiently large.

As in [1], the pseudorandomness condition that we assume on  $\nu$  is the linear forms condition, as follows.

**Definition 2.2** (Linear forms condition). A nonnegative function  $\nu = \nu^{(N)} : \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$  is said to obey the  $k$ -linear forms condition if one has

$$\mathbb{E} \left[ \prod_{j=1}^k \prod_{\omega \in \{0,1\}^{[k] \setminus \{j\}}} \nu \left( \sum_{i=1}^k (i-j)x_i^{(\omega_i)} \right)^{n_{j,\omega}} \middle| x_1^{(0)}, x_1^{(1)}, \dots, x_k^{(0)}, x_k^{(1)} \in \mathbb{Z}_N \right] = 1 + o(1) \quad (2)$$

for any choice of exponents  $n_{j,\omega} \in \{0, 1\}$ .

**Example 2.3.** For  $k = 3$ , condition (2) says that

$$\mathbb{E}[\nu(y+2z)\nu(y'+2z)\nu(y+2z')\nu(y'+2z')\nu(-x+z)\nu(-x'+z)\nu(-x+z')\nu(-x'+z') \\ \nu(-2x-y)\nu(-2x'-y)\nu(-2x-y')\nu(-2x'-y') | x, x', y, y', z, z' \in \mathbb{Z}_N] = 1 + o(1)$$

and similar conditions hold if one or more of the twelve  $\nu$  factors in the expectation are erased.

The main result of this note is the following theorem.

**Theorem 2.4** (Relative Szemerédi theorem). *Let  $k \geq 3$  and  $0 < \delta \leq 1$  be fixed. Let  $\nu : \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$  satisfy the  $k$ -linear forms condition. Assume that  $N$  is sufficiently large and relatively prime to  $(k-1)!$ . Let  $f : \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$  satisfy  $0 \leq f(x) \leq \nu(x)$  for all  $x \in \mathbb{Z}_N$  and  $\mathbb{E}[f] \geq \delta$ . Then*

$$\mathbb{E}[f(x)f(x+d)f(x+2d) \cdots f(x+(k-1)d) | x, d \in \mathbb{Z}_N] \geq c(k, \delta) - o_{k,\delta}(1), \quad (3)$$

where  $c(k, \delta)$  is the same constant which appears in Theorem 2.1. The rate at which the  $o_{k,\delta}(1)$  term goes to zero depends not only on  $k$  and  $\delta$  but also the rate of convergence in the  $k$ -linear forms condition for  $\nu$ .

This theorem was proved in [1] without the additional conclusion that  $c(k, \delta)$  can be taken to be the same as in Theorem 2.1. Indeed, the proof in [1] uses the hypergraph removal lemma as a black box, so that the constants  $c(k, \delta)$  there are much worse, with an Ackermann-type dependence due to the use of hypergraph regularity. In [6], Green and Tao also transferred Szemerédi's theorem directly to obtain the same constants  $c(k, \delta)$  as in Theorem 2.1, but under stronger pseudorandomness hypotheses for  $\nu$ . So Theorem 2.4 combines the conclusions of the two relative Szemerédi theorems in [1] and [6].

### 3. DENSE MODEL THEOREM

In this section, we show that the  $f$  in Theorem 2.4 can be modeled by a function  $\tilde{f} : \mathbb{Z}_N \rightarrow [0, 1]$ . We state our results in terms of a finite abelian group  $G$  (written additively), but there is no loss in thinking  $G = \mathbb{Z}_N$ . For  $x = (x_1, \dots, x_r) \in G^r$ , and  $I \subseteq [r]$ , we write  $x_I = (x_i)_{i \in I}$ .

**Definition 3.1.** Let  $G$  be a finite abelian group,  $r$  be a positive integer,  $\psi : G^r \rightarrow G$  be a surjective homomorphism, and  $f, \tilde{f} : G \rightarrow \mathbb{R}_{\geq 0}$  be two functions. We say that  $(f, \tilde{f})$  is an  $(r, \epsilon)$ -discrepancy pair with respect to  $\psi$  if

$$\left| \mathbb{E} \left[ (f(\psi(x)) - \tilde{f}(\psi(x))) \prod_{i=1}^r u_i(x_{[r] \setminus \{i\}}) \middle| x \in G^r \right] \right| \leq \epsilon \quad (4)$$

for all collections of functions  $u_1, \dots, u_r : G^{r-1} \rightarrow [0, 1]$ .

**Example 3.2.** When  $r = 2$  and  $\psi(x, y) = x + y$ , (4) says

$$|\mathbb{E}[(f(x+y) - \tilde{f}(x+y))u_1(y)u_2(x) | x, y \in G]| \leq \epsilon.$$

In other words, this says that the two weighted graphs  $g, \tilde{g} : G \times G \rightarrow \mathbb{R}_{\geq 0}$  given by  $g(x, y) = f(x+y)$  and  $\tilde{g}(x, y) = \tilde{f}(x+y)$  satisfy  $\|g - \tilde{g}\|_{\square} \leq \epsilon$ , where  $\|\cdot\|_{\square}$  is the cut norm for bipartite graphs.

When  $r = 3$  and  $\psi(x, y, z) = x + y + z$ , (4) says

$$|\mathbb{E}[(f(x+y+z) - \tilde{f}(x+y+z))u_1(y, z)u_2(x, z)u_3(x, y) | x, y, z \in G]| \leq \epsilon.$$

The following key lemma says that any  $0 \leq f \leq \nu$  can be approximated by a  $0 \leq \tilde{f} \leq 1$  in the above sense.

**Lemma 3.3.** *For every  $\epsilon > 0$  there is an  $\epsilon' = \exp(-\epsilon^{-O(1)})$  such that the following holds:*

*Let  $G$  be a finite abelian group,  $r$  be a positive integer, and  $\psi: G^r \rightarrow G$  be a surjective homomorphism. Let  $f, \nu: G \rightarrow \mathbb{R}_{\geq 0}$  be such that  $0 \leq f \leq \nu$ ,  $\mathbb{E}[f] \leq 1$ , and  $(\nu, 1)$  is an  $(r, \epsilon')$ -discrepancy pair with respect to  $\psi$ . Then there exists a function  $\tilde{f}: G \rightarrow [0, 1]$  so that  $\mathbb{E}[\tilde{f}] = \mathbb{E}[f]$  and  $(f, \tilde{f})$  is an  $(r, \epsilon)$ -discrepancy pair with respect to  $\psi$ .*

The proof of Lemma 3.3 uses the dense model theorem of Green-Tao [6] and Tao-Ziegler [18], which was later simplified in [5] and [10]. The expository note [9] has a nice and short write-up of the proof of the dense model theorem, and we quote the statement from there.

Let  $X$  be a finite set. For any two functions  $f, g: X \rightarrow \mathbb{R}$ , we write  $\langle f, g \rangle = \mathbb{E}[f(x)g(x)|x \in X]$ . For  $\mathcal{F}$  a collection of functions  $\varphi: X \rightarrow [-1, 1]$ , we write  $\mathcal{F}^k$  to mean the collections of all functions of the form  $\prod_{i=1}^{k'} \varphi_i$ , where  $\varphi_i \in \mathcal{F}$  and  $k' \leq k$ . In particular, if  $\mathcal{F}$  is closed under multiplication, then  $\mathcal{F}^k = \mathcal{F}$ .

**Lemma 3.4** (Green-Tao-Ziegler dense model theorem). *For every  $\epsilon > 0$ , there is a  $k = (1/\epsilon)^{O(1)}$  and an  $\epsilon' = \exp(-(1/\epsilon)^{O(1)})$  such that the following holds:*

*Suppose that  $\mathcal{F}$  is a collection of functions  $\varphi: X \rightarrow [-1, 1]$  on a finite set  $X$ ,  $\nu: X \rightarrow \mathbb{R}_{\geq 0}$  satisfies*

$$|\langle \nu - 1, \varphi \rangle| \leq \epsilon' \text{ for all } \varphi \in \mathcal{F}^k,$$

*and  $f: X \rightarrow \mathbb{R}_{\geq 0}$  satisfies  $f \leq \nu$  and  $\mathbb{E}[f] \leq 1$ . Then there is a function  $\tilde{f}: X \rightarrow [0, 1]$  such that  $\mathbb{E}[\tilde{f}] = \mathbb{E}[f]$ , and*

$$|\langle f - \tilde{f}, \varphi \rangle| \leq \epsilon \text{ for all } \varphi \in \mathcal{F}.$$

We shall use Lemma 3.4 with  $\mathcal{F}$  closed under multiplication, so that  $k$  plays no role. This is an important point in our simplification over previous approaches using the dense model theorem.

*Proof of Lemma 3.3.* For any collection of functions  $u_1, \dots, u_r: G^{r-1} \rightarrow \mathbb{R}$ , define a generalized convolution  $(u_1, \dots, u_r)_\psi^*: G \rightarrow \mathbb{R}$  by

$$(u_1, \dots, u_r)_\psi^*(x) = \mathbb{E} \left[ \prod_{i=1}^r u_i(y_{[r] \setminus \{i\}}) \middle| y \in G^r, \psi(y) = x \right].$$

Then the left-hand side of (4) can be written as  $|\langle f - \tilde{f}, (u_1, \dots, u_r)_\psi^* \rangle|$ . Let  $\mathcal{F}$  be the set of functions which can be obtained by convex combinations of functions of the form  $(u_1, \dots, u_r)_\psi^*$ , varying over all combinations of functions  $u_1, \dots, u_r: G^{r-1} \rightarrow [0, 1]$  (but  $\psi$  is fixed). Then  $(f, \tilde{f})$  being an  $(r, \epsilon')$ -discrepancy pair with respect to  $\psi$  is equivalent to  $|\langle f - \tilde{f}, \varphi \rangle| \leq \epsilon$  for all  $\varphi \in \mathcal{F}$ . The desired claim would then follow from Lemma 3.4 and the triangle inequality provided we can show that  $\mathcal{F}$  is closed under multiplication. It suffices to show that for  $u_1, \dots, u_r, u'_1, \dots, u'_r: G^{r-1} \rightarrow [0, 1]$ , the product of  $(u_1, \dots, u_r)_\psi^*$  and  $(u'_1, \dots, u'_r)_\psi^*$  still lies in  $\mathcal{F}$ . Indeed, we have

$$\begin{aligned} (u_1, \dots, u_r)_\psi^*(x)(u'_1, \dots, u'_r)_\psi^*(x) &= \mathbb{E} \left[ \prod_{i=1}^r u_i(y_{[r] \setminus \{i\}}) u'_i(y'_{[r] \setminus \{i\}}) \middle| y, y' \in G^r, \psi(y) = \psi(y') = x \right] \\ &= \mathbb{E} \left[ \prod_{i=1}^r u_i(y_{[r] \setminus \{i\}}) u'_i(y_{[r] \setminus \{i\}} + z_{[r] \setminus \{i\}}) \middle| y, z \in G^r, \psi(y) = x, \psi(z) = 0 \right] \\ &= \mathbb{E}[(v_{1, z_{[r] \setminus \{1\}}}, v_{2, z_{[r] \setminus \{2\}}}, \dots, v_{r, z_{[r] \setminus \{r\}}})_\psi^*(x) | z \in G^r, \psi(z) = 0] \end{aligned}$$

where  $v_{i, z_{[r] \setminus \{i\}}}: G^{r-1} \rightarrow [0, 1]$  is defined by  $v_{i, z_{[r] \setminus \{i\}}}(y_{[r] \setminus \{i\}}) = u_i(y_{[r] \setminus \{i\}}) u'_i(y_{[r] \setminus \{i\}} + z_{[r] \setminus \{i\}})$ . This shows that the product of two such generalized convolutions is a convex combination of generalized convolutions, so that  $\mathcal{F}$  is closed under multiplication.  $\blacksquare$

## 4. COUNTING LEMMA

Next we show that if  $(f, \tilde{f})$  is a  $(k-1, \epsilon)$ -discrepancy pair, with  $f \leq \nu$  and  $\tilde{f} \leq 1$ , then  $f$  and  $\tilde{f}$  have similar number of (weighted)  $k$ -term APs. This is a special case of the counting lemma for sparse hypergraphs from [1], whose self-contained proof takes up about 4 pages [1, Sec. 6].

**Lemma 4.1** (*k*-AP counting lemma). *For every  $k \geq 3$  and  $\gamma > 0$ , there exists an  $\epsilon > 0$  so that the following holds.*

Let  $\nu, f, \tilde{f}: \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$  be functions. Suppose that  $\nu$  satisfies the  $k$ -linear forms condition and  $N$  is sufficiently large. Suppose also that  $0 \leq f \leq \nu$ ,  $0 \leq \tilde{f} \leq 1$ , and  $(f, \tilde{f})$  is a  $(k-1, \epsilon)$ -discrepancy pair with respect to each of  $\psi_1, \dots, \psi_k$ , where  $\psi_j: \mathbb{Z}_N^{k-1} \rightarrow \mathbb{Z}_N$  is defined by

$$\psi_j(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k) := \sum_{i \in [k] \setminus \{j\}} (i - j)x_i.$$

Then

$$\left| \mathbb{E} \left[ \prod_{i=0}^{k-1} f(a + id) \middle| a, d \in \mathbb{Z}_N \right] - \mathbb{E} \left[ \prod_{i=0}^{k-1} \tilde{f}(a + id) \middle| a, d \in \mathbb{Z}_N \right] \right| \leq \gamma. \quad (5)$$

Let us explain why Lemma 4.1 is a special case of [1, Thm. 2.17]. We use the hypergraph notation from [1, Sec. 2]. Let  $V = (J, (V_j)_{j \in J}, k-1, H)$  be a hypergraph system, where  $J = [k]$ ,  $V_j = \mathbb{Z}_N$  for every  $j \in J$ , and  $H = \binom{J}{k-1}$  (corresponding to a simplex). Let  $(\nu_e)_{e \in H}$ ,  $(g_e)_{e \in H}$ , and  $(\tilde{g}_e)_{e \in H}$  be weighted hypergraphs on  $V$  defined by

$$\begin{aligned} \nu_{[k] \setminus \{j\}}(x_{[k] \setminus \{j\}}) &= \nu(\psi_j(x_{[k] \setminus \{j\}})) \\ g_{[k] \setminus \{j\}}(x_{[k] \setminus \{j\}}) &= f(\psi_j(x_{[k] \setminus \{j\}})) \\ \tilde{g}_{[k] \setminus \{j\}}(x_{[k] \setminus \{j\}}) &= \tilde{f}(\psi_j(x_{[k] \setminus \{j\}})) \end{aligned}$$

for  $j \in [k]$  and  $x_{[k] \setminus \{j\}} \in V_{[k] \setminus \{j\}} = \mathbb{Z}_N^{k-1}$ . Then the weighted hypergraph  $(\nu_e)_{e \in H}$  satisfies the  $H$ -linear forms condition [1, Def. 2.8] (which is equivalent to  $\nu: \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$  satisfying the  $k$ -linear forms condition). That  $(f, \tilde{f})$  is a  $(k-1, \epsilon)$ -discrepancy pair with respect to  $\psi_j$  is equivalent to  $(g_{[k] \setminus \{j\}}, \tilde{g}_{[k] \setminus \{j\}})$  being an  $\epsilon$ -discrepancy pair as weighted hypergraphs [1, Def. 2.13]. Note that

$$\mathbb{E} \left[ \prod_{i=0}^{k-1} f(a + id) \middle| x, d \in \mathbb{Z}_N \right] = \mathbb{E} \left[ \prod_{e \in H} g_e(x_e) \middle| x \in V_J \right]$$

(to see this, let  $a = \psi_1(x_2, \dots, x_k)$  and  $d = -(x_1 + \dots + x_k)$ ) and similarly with  $\tilde{f}$  and  $\tilde{g}_e$ . Then the relative hypergraph counting lemma [1, Thm. 2.17] reduces to Lemma 4.1.

## 5. PROOF OF THE RELATIVE SZEMERÉDI THEOREM

*Proof of Theorem 2.4.* We begin with the following simple observation, that for any  $g, g': \mathbb{Z}_N \rightarrow \mathbb{R}_{\geq 0}$ , if  $(g, g')$  is a  $(k-1, \epsilon)$ -discrepancy pair with respect to one  $\psi_j$  from Lemma 4.1, then it is so with respect to all  $\psi_j$ . This is simply because  $1, 2, \dots, k-1$  all have multiplicative inverses in  $\mathbb{Z}_N$ , as  $N$  is coprime to  $(k-1)!$ , and a scaling of variables in (4) allows one to convert one linear form  $\psi_j$  to another  $\psi_{j'}$ .

The linear forms condition on  $\nu$  implies that  $(\nu, 1)$  is a  $(k-1, o(1))$ -discrepancy pair with respect to  $\psi_1$  from Lemma 4.1. Indeed, we have the following inequality

$$\left| \mathbb{E} \left[ (\nu(\psi_1(x)) - 1) \prod_{i=1}^r u_i(x_{[r] \setminus \{i\}}) \middle| x \in G^r \right] \right| \leq \mathbb{E} \left[ \prod_{\omega \in \{0,1\}^r} (\nu(\psi_1(x^{(\omega)})) - 1) \middle| x^{(0)}, x^{(1)} \in G^r \right]^{1/2^r} \quad (6)$$

which is proved by a sequence of Cauchy-Schwarz inequalities, similar to [1, Lem. 6.2]. The right-hand side of (6) is  $o(1)$  by the linear forms condition (expand the product so that each term is  $\pm 1 + o(1)$  by (2), and everything cancels accordingly).

Since  $(\nu, 1)$  is a  $(k-1, o(1))$ -discrepancy pair with respect to  $\psi_1$ , Lemma 3.3 implies that there exists  $\tilde{f}: G \rightarrow [0, 1]$  so that  $\mathbb{E}[\tilde{f}] = \mathbb{E}[f] \geq \delta$  (if  $\mathbb{E}[f] > 1$ , then replace  $f$  by  $\delta f / \mathbb{E}[f]$ ) and  $(f, \tilde{f})$  is a  $(k-1, o(1))$ -discrepancy pair with respect to  $\psi_1$ , and hence with respect to all  $\psi_j$ ,  $1 \leq j \leq k$ . So

$$\begin{aligned} \mathbb{E} \left[ \prod_{i=0}^{k-1} f(x+id) \middle| x, d \in \mathbb{Z}_N \right] &\geq \mathbb{E} \left[ \prod_{i=0}^{k-1} \tilde{f}(x+id) \middle| x, d \in \mathbb{Z}_N \right] - o(1) \\ &\geq c(k, \delta) - o_{k, \delta}(1), \end{aligned}$$

where the first inequality is by Lemma 4.1 and the second inequality is by Theorem 2.1.  $\blacksquare$

**Acknowledgments.** The author would like to thank Jacob Fox and David Conlon for careful readings of the manuscript.

#### REFERENCES

- [1] D. Conlon, J. Fox, and Y. Zhao, *A relative Szemerédi theorem.*, arXiv:1305.5440.
- [2] D. A. Goldston and C. Y. Yıldırım, *Higher correlations of divisor sums related to primes. I. Triple correlations*, *Integers* **3** (2003), A5, 66.
- [3] W. T. Gowers, *A new proof of Szemerédi’s theorem*, *Geom. Funct. Anal.* **11** (2001), no. 3, 465–588.
- [4] ———, *Hypergraph regularity and the multidimensional Szemerédi theorem*, *Ann. of Math.* **166** (2007), no. 3, 897–946.
- [5] ———, *Decompositions, approximate structure, transference, and the Hahn-Banach theorem*, *Bull. Lond. Math. Soc.* **42** (2010), no. 4, 573–606.
- [6] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, *Ann. of Math.* **167** (2008), no. 2, 481–547.
- [7] ———, *New bounds for Szemerédi’s theorem. II. A new bound for  $r_4(N)$* , *Analytic number theory*, Cambridge Univ. Press, Cambridge, 2009, pp. 180–204.
- [8] B. Nagle, V. Rödl, and M. Schacht, *The counting lemma for regular  $k$ -uniform hypergraphs*, *Random Structures Algorithms* **28** (2006), no. 2, 113–179.
- [9] O. Reingold, L. Trevisan, M. Tulsiani, and S. Vadhan, *New proofs of the Green-Tao-Ziegler dense model theorem: An exposition*, arXiv:0806.0381.
- [10] ———, *Dense subsets of pseudorandom sets*, 49th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society, 2008, pp. 76–85.
- [11] V. Rödl and J. Skokan, *Regularity lemma for  $k$ -uniform hypergraphs*, *Random Structures Algorithms* **25** (2004), no. 1, 1–42.
- [12] ———, *Applications of the regularity lemma for uniform hypergraphs*, *Random Structures Algorithms* **28** (2006), no. 2, 180–194.
- [13] T. Sanders, *On Roth’s theorem on progressions*, *Ann. of Math. (2)* **174** (2011), no. 1, 619–636.
- [14] E. Szemerédi, *On sets of integers containing no  $k$  elements in arithmetic progression*, *Acta Arith.* **27** (1975), 199–245.
- [15] T. Tao, *A remark on Goldston-Yıldırım correlation estimates*, available at <http://www.math.ucla.edu/~tao/preprints/Expository/gy-corr.dvi>.
- [16] ———, *The Gaussian primes contain arbitrarily shaped constellations*, *J. Anal. Math.* **99** (2006), 109–176.
- [17] ———, *A variant of the hypergraph removal lemma*, *J. Combin. Theory Ser. A* **113** (2006), no. 7, 1257–1280.
- [18] T. Tao and T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, *Acta Math.* **201** (2008), no. 2, 213–305.
- [19] P. Varnavides, *On certain sets of positive density*, *J. London Math. Soc.* **34** (1959), 358–360.

DEPARTMENT OF MATHEMATICS, MIT, CAMBRIDGE, MA 02139-4307  
*E-mail address:* yufeiz@math.mit.edu