A SHORT PROOF OF THE MULTIDIMENSIONAL SZEMERÉDI THEOREM IN THE PRIMES

JACOB FOX AND YUFEI ZHAO

Abstract. Tao conjectured that every dense subset of $\mathcal{P}^d$, the d-tuples of primes, contains constellations of any given shape. This was very recently proved by Cook, Magyar, and Titichetrakun and independently by Tao and Ziegler. Here we give a simple proof using the Green-Tao theorem on linear equations in primes and the Furstenberg-Katznelson multidimensional Szemerédi theorem.

Let $\mathcal{P}_N$ denote the set of primes at most N, and let $[N] := \{1, 2, \ldots, N\}$. Tao [12] conjectured the following result as a natural extension of the Green-Tao theorem [7] on arithmetic progressions in the primes and the Furstenberg-Katznelson [6] multidimensional generalization of Szemerédi's theorem. Special cases of this conjecture were proven in [4] and [11]. The conjecture was very recently resolved by Cook, Magyar, and Titichetrakun [5] and independently by Tao and Ziegler [13].

Theorem 1. *Let d be a positive integer, $v_1, \ldots, v_k \in \mathbb{Z}^d$, and $\delta > 0$. Then, if N is sufficiently large, every subset A of $\mathcal{P}_N^d$ of cardinality $|A| \geq \delta |\mathcal{P}_N|^d$ contains a set of the form $a + tv_1, \ldots, a + tv_k$, where $a \in \mathbb{Z}^d$ and t is a positive integer.*

In this note we give a short alternative proof of the theorem, using the landmark result of Green and Tao [8] (which is conditional on results later proved in [9] and with Ziegler in [10]) on the asymptotics for the number of primes satisfying certain systems of linear equations, as well as the following multidimensional generalization of Szemerédi's theorem established by Furstenberg and Katznelson [6].

Theorem 2 (Multidimensional Szemerédi theorem [6])**.** *Let d be a positive integer, $v_1, \ldots, v_k \in \mathbb{Z}^d$, and $\delta > 0$. If N is sufficiently large, then every subset A of $[N]^d$ of cardinality $|A| \geq \delta N^d$ contains a set of the form $a + tv_1, \ldots, a + tv_k$, where $a \in \mathbb{Z}^d$ and t is a positive integer.*

To prove Theorem 1, we begin by fixing $d, v_1, \ldots, v_k, \delta$. Using Theorem 2, we can fix a large integer $m > 2d/\delta$ so that any subset of $[m]^d$ with at least $\delta m^d/2$ elements contains a set of the form $a + tv_1, \ldots, a + tv_k$, where $a \in \mathbb{Z}^d$ and t is a positive integer.

We next discuss a sketch of the proof idea. The Green-Tao theorem [7] (also see [3] for some recent simplifications) states that there are arbitrarily long arithmetic progressions in the primes. It follows that for N large, $\mathcal{P}_N^d$ contains homothetic copies of $[m]^d$. We use a Varnavides-type argument [14] and consider a random homothetic copy of the grid $[m]^d$ inside $\mathcal{P}_N^d$. In expectation, the set A should occupy at least a $\delta/2$ fraction of the random homothetic copy of $[m]^d$. This follows from a linearity of expectation argument. Indeed, the Green-Tao-Ziegler result [8, 9, 10] and a second moment argument imply that most points of $\mathcal{P}_N^d$ appear in about the expected number of such copies of the grid $[m]^d$. Once we find a homothetic copy of $[m]^d$ containing at least $\delta m^d/2$ elements of A, we obtain by Theorem 2 a subset of A of the form $a + tv_1, \ldots, a + tv_k$, as desired.

To make the above idea actually work, we first apply the W-trick as described below. This is done to avoid certain biases in the primes. We also only consider homothetic copies of $[m]^d$ with

The first author was supported by a Simons Fellowship, NSF grant DMS-1069197, by an Alfred P. Sloan Fellowship, and by an MIT NEC Corporation Fund Award, and the second author was supported by an internship at Microsoft Research New England.

common difference $r \le N/m^2$ in order to guarantee that almost all elements of $\mathcal{P}_N^d$ are in about the same number of such homothetic copies of $[m]^d$.

Remarks. This argument also produces a relative multidimensional Szemerédi theorem, where the complexity of the linear forms condition on the majorizing measure depends on $d, v_1, \dots, v_k$ and δ. It seems plausible that the dependence on δ is unnecessary; this was shown for the one-dimensional case in [3]. Our arguments share some similarities with those of Tao and Ziegler [13], who also use the results in [8, 9, 10]. However, the proof in [13] first establishes a relativized version of the Furstenberg correspondence principle and then proceeds in the ergodic theoretic setting, whereas we go directly to the multidimensional Szemerédi theorem. Cook, Magyar, and Titichetrakun [5] take a different approach and develop a relative hypergraph removal lemma from scratch, and they also require a linear forms condition whose complexity depend on δ.

Conditional on a certain polynomial extension of the Green-Tao-Ziegler result (c.f. the Bateman-Horn conjecture [1]), one can also combine this sampling argument with the polynomial extension of Szemerédi's theorem by Bergelson and Leibman [2] to obtain a polynomial extension of Theorem 1.

The hypothesis that $|A| \ge \delta \, |\mathcal{P}_N|^d$ implies that

$$\sum_{n_1,\dots,n_d \in [N]} 1_A(n_1,\dots,n_d)\Lambda'(n_1)\cdots\Lambda'(n_d) \ge (\delta - o(1))N^d, \tag{1}$$

where 1_A is the indicator function of A, and $o(1)$ denotes some quantity that goes to zero as $N \to \infty$, and $\Lambda'(p) = \log p$ for prime p and $\Lambda'(n) = 0$ for nonprime n.

Next we apply the W-trick [8, §5]. Fix some slowly growing function $w = w(N)$; the choice $w := \log\log\log N$ will do. Define $W := \prod_{p \le w} p$ to be the product of all primes at most w. For each $b \in [W]$ with $\gcd(b, W) = 1$, define

$$\Lambda'_{b,W}(n) := \frac{\phi(W)}{W}\Lambda'(Wn + b)$$

where $\phi(W) = \#\{b \in [W] : \gcd(b, W) = 1\}$ is the Euler totient function. Also define

$$1_{A_{b_1,\dots,b_d,W}}(n_1,\dots,n_d) := 1_A(Wn_1 + b_1, \dots, Wn_d + b_d).$$

By (1) and the pigeonhole principle, we can choose $b_1, \dots, b_d \in [W]$ all coprime to W so that

$$\sum_{1 \le n_1,\dots,n_d \le N/W} 1_{A_{b_1,\dots,b_d,W}}(n_1,\dots,n_d)\Lambda'_{b_1,W}(n)\Lambda'_{b_2,W}(n)\cdots\Lambda'_{b_d,W}(n) \ge (\delta - o(1))\left(\frac{N}{W}\right)^d, \tag{2}$$

We shall write

$$\widetilde{N} := \lfloor N/W \rfloor, \qquad R := \lfloor \widetilde{N}/m^2 \rfloor, \qquad \widetilde{A} := 1_{A_{b_1,\dots,b_d,W}} \qquad \text{and} \qquad \widetilde{\Lambda}_j := \Lambda'_{b_j,W}$$

(all depending on N). So (2) reads

$$\sum_{n_1,\dots,n_d \in [\widetilde{N}]} \widetilde{A}(n_1,\dots,n_d)\widetilde{\Lambda}_1(n_1)\widetilde{\Lambda}_2(n_2)\cdots\widetilde{\Lambda}_d(n_d) \ge (\delta - o(1))\widetilde{N}^d \tag{3}$$

The Green-Tao result [8] (along with [9, 10]) says that $\Lambda'_{b_j,W}$ acts pseudorandomly with average value about 1 in terms of counts of linear forms. The statement below is an easy corollary of [8, Thm. 5.1].

Theorem 3 (Pseudorandomness of the W-tricked primes)**.** *Fix a linear map $\Psi = (\psi_1, \dots, \psi_t) : \mathbb{Z}^d \to \mathbb{Z}^t$ (in particular $\Psi(0) = 0$) where no two ψ_i, ψ_j are linearly dependent. Let $K \subseteq [-\widetilde{N}, \widetilde{N}]^d$ be any convex body. Then, for any $b_1, \dots, b_t \in [W]$ all coprime to W, we have*

$$\sum_{n \in K \cap \mathbb{Z}^d} \prod_{j \in [t]} \Lambda'_{b_j,W}(\psi_j(n)) = \#\{n \in K \cap \mathbb{Z}^d : \psi_j(n) > 0 \ \forall j\} + o(\widetilde{N}^d).$$

where $o(\widetilde{N}^d) := o(1)\widetilde{N}^d$. *Note that the error term does not depend on* $b_1, \ldots, b_t$ *(although it does depend on* Ψ*).*

The next lemma shows that A in expectation contains a considerable fraction of a random homothetic copy of $[m]^d$ with common difference at most $R = \lfloor N/m^2 \rfloor$ in the W-tricked subgrid of $\mathcal{P}_N^d$.

Lemma 4. *If* $\widetilde{A}$ *satisfies* (3), *then*

$$\sum_{\substack{n_1,\ldots,n_d \in [\widetilde{N}] \\ r \in [R]}} \left(\sum_{i_1,\ldots,i_d \in [m]} \widetilde{A}(n_1 + i_1 r, \ldots, n_d + i_d r) \right) \prod_{j \in [d]} \prod_{i \in [m]} \widetilde{\Lambda}_j(n_j + ir)$$
$$\geq (\delta m^d - dm^{d-1} - o(1)) R \widetilde{N}^d. \quad (4)$$

Proof of Theorem 1 (assuming Lemma 4) . By Theorem 3 we have

$$\sum_{\substack{n_1,\ldots,n_d \in [\widetilde{N}] \\ r \in [R]}} \prod_{j \in [d]} \prod_{i \in [m]} \widetilde{\Lambda}_j(n_j + ir) = (1 + o(1)) R \widetilde{N}^d,$$

So by (4), for sufficiently large N, there exists some choice of $n_1, \ldots, n_d \in [\widetilde{N}]$ and $r \in [R]$ so that

$$\sum_{i_1,\ldots,i_d \in [m]} \widetilde{A}(n_1 + i_1 r, \ldots, n_d + i_d r) \geq \frac{1}{2} \delta m^d.$$

This means that a certain dilation of the grid $[m]^d$ contains at least $\delta m^d/2$ elements of A, from which it follows by the choice of m that it must contain a set of the form $a + tv_1, \ldots, a + tv_k$. ■

Lemma 4 follows from the next lemma by summing over all choices of $i_1, \ldots, i_d \in [m]$.

Lemma 5. *Suppose* $\widetilde{A}$ *satisfies* (3). *Fix* $i_1, \ldots, i_d \in [m]$. *Then we have*

$$\sum_{\substack{n_1,\ldots,n_d \in [\widetilde{N}] \\ r \in [R]}} \widetilde{A}(n_1 + i_1 r, \ldots, n_d + i_d r) \prod_{j \in [d]} \prod_{i \in [m]} \widetilde{\Lambda}_j(n_j + ir) \geq \left(\delta - \frac{d}{m} - o(1) \right) R \widetilde{N}^d. \quad (5)$$

Proof. By a change of variables $n_j' = n_j + i_j r$ for each j, we write the LHS of (5) as

$$\sum_{r \in [R]} \sum_{\substack{n_1',\ldots,n_d' \in \mathbb{Z} \\ n_j' - i_j r \in [\widetilde{N}] \ \forall j}} \widetilde{A}(n_1', \ldots, n_d') \prod_{j \in [d]} \prod_{i \in [m]} \widetilde{\Lambda}_j(n_j' + (i - i_j) r). \quad (6)$$

Note that (6) is at least

$$\sum_{r \in [R]} \sum_{\widetilde{N}/m < n_1',\ldots,n_d' \leq \widetilde{N}} \widetilde{A}(n_1', \ldots, n_d') \prod_{j \in [d]} \prod_{i \in [m]} \widetilde{\Lambda}_j(n_j' + (i - i_j) r). \quad (7)$$

By (3) and Theorem 3 we have

$$\sum_{\widetilde{N}/m < n_1,\ldots,n_d \leq \widetilde{N}} \widetilde{A}(n_1, \ldots, n_d) \widetilde{\Lambda}_1(n_1) \widetilde{\Lambda}_2(n_2) \cdots \widetilde{\Lambda}_d(n_d) \geq \left(\delta - \frac{d}{m} - o(1) \right) \widetilde{N}^d \quad (8)$$

(the difference between the left-hand side sums of (3) and (8) consists of terms with $(n_1, \ldots, n_d)$ in some box of the form $[\widetilde{N}]^{j-1} \times [\widetilde{N}/m] \times [\widetilde{N}]^{d-j}$, which can be upper bounded by using $\widetilde{A} \leq 1$, applying Theorem 3, and then taking the union bound over $j \in [d]$). It remains to show that

$$(7) - R \cdot (\text{LHS of } (8)) = o(\widetilde{N}^{d+1}).$$

We have

$$
\begin{aligned}
&(7) - R \cdot (\text{LHS of } (8)) \\
&= \sum_{\substack{\widetilde{N}/m < n_1', \dots, n_d' \le \widetilde{N} \\ r \in [R]}} \widetilde{A}(n_1', \dots, n_d') \left(\prod_{j \in [d]} \prod_{i \in [m]} \widetilde{\Lambda}_j(n_j' + (i - i_j)r) - \prod_{j \in [d]} \widetilde{\Lambda}_j(n_j) \right) \\
&= \sum_{\widetilde{N}/m < n_1', \dots, n_d' \le \widetilde{N}} \widetilde{A}(n_1', \dots, n_d') \left(\prod_{j \in [d]} \widetilde{\Lambda}_j(n_j') \right) \left(\sum_{r \in [R]} \left(\prod_{j \in [d]} \prod_{i \in [m] \setminus \{i_j\}} \widetilde{\Lambda}_j(n_j' + (i - i_j)r) \ - 1 \right) \right).
\end{aligned}
$$

By the Cauchy-Schwarz inequality and $0 \le \widetilde{A} \le 1$, the above expression can be bounded in absolute value by $\sqrt{ST}$, where

$$
\begin{aligned}
S &= \sum_{\widetilde{N}/m < n_1', \dots, n_d' \le \widetilde{N}} \left(\prod_{j \in [d]} \widetilde{\Lambda}_j(n_j') \right) \left(\sum_{r \in [R]} \left(\prod_{j \in [d]} \prod_{i \in [m] \setminus \{i_j\}} \widetilde{\Lambda}_j(n_j' + (i - i_j)r) \ - 1 \right) \right)^2 \\
T &= \sum_{\widetilde{N}/m < n_1', \dots, n_d' \le \widetilde{N}} \prod_{j \in [d]} \widetilde{\Lambda}_j(n_j').
\end{aligned}
$$

By Theorem 3 we have $S = o(\widetilde{N}^{d+2})$ (to see this, expand the square, apply Theorem 3, and observe the cancellations) and $T = O(\widetilde{N}^d)$, so that $\sqrt{ST} = o(\widetilde{N}^{d+1})$, as desired. ■

Acknowledgment. We thank David Conlon, Ben Green, and Terry Tao for helpful comments.

References

[1] P. T. Bateman and R. A. Horn, *A heuristic asymptotic formula concerning the distribution of prime numbers*, Math. Comp. **16** (1962), 363–367.

[2] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerdi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753.

[3] D. Conlon, J. Fox, and Y. Zhao, *A relative Szemerédi theorem*, arXiv:1305.5440.

[4] B. Cook, A. Magyar, *Constellations in* $\mathbb{P}^d$, Int. Math. Res. Not. IMRN **2012** (2012), no. 12, 2794–2816.

[5] B. Cook, A. Magyar, and T. Titichetrakun, *A multidimensional Szemerédi theorem in the primes*, arXiv:1306.3025.

[6] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291.

[7] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. of Math. **167** (2008), no. 2, 481–547.

[8] ——, *Linear equations in primes*, Ann. of Math. **171** (2010), no. 3, 1753–1850.

[9] ——, *The Möbius function is strongly orthogonal to nilsequences*, Ann. of Math. **175** (2012), no. 2, 541–566.

[10] B. Green, T. Tao, and T. Ziegler, *An inverse theorem for the Gowers* $U^{s+1}[N]$*-norm*, Ann. of Math. **176** (2012), no. 2, 1231–1372.

[11] A. Magyar, and T. Titichetrakun, *Corners in dense subsets of* $\mathbb{P}^d$, arXiv:1306.3026.

[12] T. Tao, *The Gaussian primes contain arbitrarily shaped constellations*, J. Anal. Math. **99** (2006), 109–176.

[13] T. Tao and T. Ziegler, *A multi-dimensional Szemerédi theorem for the primes via a correspondence principle*, arXiv:1306.2886.

[14] 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: `fox@math.mit.edu`

Department of Mathematics, MIT, Cambridge, MA 02139-4307
E-mail address: `yufeiz@math.mit.edu`