LINEAR FORMS FROM THE GOWERS UNIFORMITY NORM

DAVID CONLON, JACOB FOX, AND YUFEI ZHAO

This is a companion note to [1] elaborating on the concluding remark in §7 under the heading *Gowers uniformity norms*. The purpose of this note is to sketch the argument showing that the relative Szemerédi theorem, Theorem 2.4 in [1], for $(r+1)$-term arithmetic progressions holds when the linear forms condition on $\nu\colon \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ is replaced by an alternate condition on the Gowers uniformity norm U^r:

$$\|\nu - 1\|_{U^r} = o(p^r), \quad \text{where } p := 1/\|\nu\|_\infty \leq 1. \tag{1}$$

Recall that the Gowers uniformity norm U^r is defined by

$$\|f\|_{U^r} = \mathbb{E}\Big[\prod_{\omega\in\{0,1\}^r} f(x_0 + \omega\cdot\mathbf{x})\Big| x_0, x_1, \dots, x_r \in \mathbb{Z}_N\Big]^{1/2^r}.$$

The application we have in mind is $\nu = p^{-1}1_S$ where $S \subseteq \mathbb{Z}_N$ satisfies $S \subseteq \mathbb{Z}_N$ and $p = |S|/N$.

We do not give all the details in this note and we also assume familiarity with [1]. We sketch how to modify the argument in [1] to show the result under the assumption (1). As noted in Footnote 5 on page 16 of [1], the only hypotheses needed for the proof of the counting lemma are the strong linear forms condition, as in Lemma 6.3, and also (34) in [1]. The Gowers uniformity hypotheses also implies the conclusion of Lemma 6.2, which gives the conclusion of Lemma 2.15, thereby allowing us to apply the weak regularity lemma, Theorem 2.16.

As in [1], we work in the hypergraph setting. Recall that for a finite set e, we write $V_e = \prod_{j\in e} V_j$, where each V_j is a finite set. We assume this notation for Definition 1 and Lemmas 2 and 3.

Definition 1 (Gowers uniformity norm)**.** For any function $g\colon V_e \to \mathbb{R}$, define

$$\|g\|_{U^e} := \mathbb{E}\Big[\prod_{\omega\in\{0,1\}^e} g(x_e^{(\omega)})\Big| x_e^{(0)}, x_e^{(1)} \in V_e\Big]^{1/2^{|e|}}.$$

There are two notions of Gowers uniformity norm: one for functions $\mathbb{Z}_N \to \mathbb{R}$ and one for functions $V_e \to \mathbb{R}$. Observe that the representation of $\nu\colon \mathbb{Z}_N \to \mathbb{R}_{\geq 0}$ by a weighted hypergraph ν in the proof of the relative Szemerédi theorem [1, §3] preserves the Gowers uniformity norm.

The following inequality is the Gowers-Cauchy-Schwarz inequality for hypergraphs. The proof is by r applications of the standard Cauchy-Schwarz inequality.

Lemma 2 (Gowers-Cauchy-Schwarz inequality)**.** *For any collection of functions $g_\omega\colon V_e \to \mathbb{R}$, $\omega \in \{0,1\}^e$, one has*

$$\Big|\mathbb{E}\Big[\prod_{\omega\in\{0,1\}^e} g_\omega(x_e^{(\omega)})\Big| x_e^{(0)}, x_e^{(1)} \in V_e\Big]\Big| \leq \prod_{\omega\in\{0,1\}^e} \|g_\omega\|_{U^e}.$$

To illustrate the proof of Lemma 2, we consider the case $|e| = 2$. We have

$$
\begin{aligned}
&\mathbb{E}[g_{00}(x,y)g_{01}(x,y')g_{10}(x',y)g_{11}(x',y')|x,x' \in V_1,\ y,y' \in V_2]^4 \\
&= \mathbb{E}[\mathbb{E}[g_{00}(x,y)g_{01}(x,y')|x \in V_1]\mathbb{E}[g_{10}(x',y)g_{11}(x',y')|x' \in V_2]|y,y' \in V_2]^4 \\
&\le \mathbb{E}[\mathbb{E}[g_{00}(x,y)g_{01}(x,y')|x \in V_1]^2|y,y' \in V_2]^2\mathbb{E}[\mathbb{E}[g_{10}(x',y)g_{11}(x',y')|x' \in V_2]^2|y,y' \in V_2]^2 \\
&= \mathbb{E}[g_{00}(x,y)g_{00}(x',y)g_{01}(x,y')g_{01}(x',y')|x,x' \in V_1,\ y,y' \in V_2]^2 \\
&\qquad\qquad \mathbb{E}[g_{10}(x,y)g_{10}(x',y)g_{11}(x,y')g_{11}(x',y')|x,x' \in V_1,\ y,y' \in V_2]^2 \\
&= \mathbb{E}[\mathbb{E}[g_{00}(x,y)g_{00}(x',y)|y \in V_2]\mathbb{E}[g_{01}(x,y')g_{01}(x',y')|y' \in V_2]|x,x' \in V_1]^2 \\
&\qquad\qquad \mathbb{E}[\mathbb{E}[g_{10}(x,y)g_{10}(x',y)|y \in V_2]\mathbb{E}[g_{11}(x,y')g_{11}(x',y')|y' \in V_2]|x,x' \in V_1]^2 \\
&\le \mathbb{E}[\mathbb{E}[g_{00}(x,y)g_{00}(x',y)|y \in V_2]^2|x,x' \in V_1]\mathbb{E}[\mathbb{E}[g_{01}(x,y')g_{01}(x',y')|y' \in V_2]^2|x,x' \in V_1] \\
&\qquad\qquad \mathbb{E}[\mathbb{E}[g_{10}(x,y)g_{10}(x',y)|y \in V_2]^2|x,x' \in V_1]\mathbb{E}[\mathbb{E}[g_{11}(x,y')g_{11}(x',y')|y' \in V_2]^2|x,x' \in V_1] \\
&= \mathbb{E}[g_{00}(x,y)g_{00}(x',y)g_{00}(x,y')g_{00}(x',y')|x,x' \in V_1,\ y,y' \in V_2] \\
&\qquad \mathbb{E}[g_{01}(x,y)g_{01}(x',y)g_{01}(x,y')g_{01}(x',y')|x,x' \in V_1,\ y,y' \in V_2] \\
&\qquad \mathbb{E}[g_{10}(x,y)g_{10}(x',y)g_{10}(x,y')g_{10}(x',y')|x,x' \in V_1,\ y,y' \in V_2] \\
&\qquad \mathbb{E}[g_{11}(x,y)g_{11}(x',y)g_{11}(x,y')g_{11}(x',y')|x,x' \in V_1,\ y,y' \in V_2] \\
&= (\|g_{00}\|_{U^2}\,\|g_{01}\|_{U^2}\,\|g_{10}\|_{U^2}\,\|g_{11}\|_{U^2})^4.
\end{aligned}
$$

Both inequalities above are due to the usual Cauchy-Schwarz inequality. The extension to the general case is straightforward.

The following lemma relates the Gowers uniformity norm condition to certain linear forms within V_e.

Lemma 3. *If $\nu_e \colon V_e \to \mathbb{R}_{\ge 0}$ satisfies $\|\nu_e - 1\|_{U^e} = o(1)$, then*

$$\mathbb{E}\Big[\prod_{\omega \in \{0,1\}^e} \nu_e(x_e^{(\omega)})^{n_\omega} \Big| x_e^{(0)}, x_e^{(1)} \in V_e\Big] = 1 + o(1) \tag{2}$$

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

Proof. Applying the Gowers-Cauchy-Schwarz inequality, Lemma 2, applied with $g_\omega(x_e) = (\nu_e(x_e) - 1)^{n_\omega}$, one gets

$$\mathbb{E}\Big[\prod_{\omega \in \{0,1\}^e} (\nu_e(x_e^{(\omega)}) - 1)^{n_\omega} \Big| x_e^{(0)}, x_e^{(1)} \in V_e\Big] = o(1) \tag{3}$$

for any choice of exponents $n_\omega \in \{0,1\}$, as long as they are not all zero. We can write the left-hand side of (2) as

$$\mathbb{E}\Big[\prod_{\omega \in \{0,1\}^e} ((\nu_e(x_e^{(\omega)}) - 1) + 1)^{n_\omega} \Big| x_e^{(0)}, x_e^{(1)} \in V_e\Big].$$

The result follows by expanding each parenthesis $((\nu_e(x_e^{(\omega)})^{n_\omega} - 1) + 1)$ and bounding each term (except for the constant term) using (3). ∎

For the rest of this note, we assume the following hypergraph system setup. Recall that this is the hypergraph system used in the proof of the relative Szemerédi theorem in [1].

Setup 4. Let $J = \{0,1,2,\dots,r\}$ and $H = \binom{J}{r}$. Write $e_j := J \setminus \{j\} \in H$ for every $j \in J$. Let $V = (J, (V_j)_{j \in J}, r, H)$ be a hypergraph system. Note that H is the complete r-uniform hypergraph on $r+1$ vertices.

For a weighted hypergraph ν on V, we write $\|\nu\|_\infty$ to mean the maximum value taken by any ν_e, $e \in H$. Throughout we assume that $\|\nu\|_\infty \geq 1$.

The next two lemmas show that the inputs to the proof of the counting lemma in [1] (see Footnote 5 on page 16) remain valid when we assume that

$$\|\nu_e - 1\|_{U^e} = o(\|\nu\|_\infty^{-r}) \quad \text{for all } e \in H.$$

Lemma 5 (Strong linear forms)**.** *Assume Setup 4. Let ν be a weighted hypergraph on V satisfying*

$$\|\nu_e - 1\|_{U^e} = o(1) \textit{ for all } e \in H \setminus \{e_0\}.$$

For each $\iota \in \{0,1\}$ and $e \in H \setminus \{e_0\}$, let $g_e^{(\iota)} \colon V_e \to \mathbb{R}_{\geq 0}$ be a function so that either $g_e^{(\iota)} \leq 1$ or $g_e^{(\iota)} \leq \nu_e$ holds. Then

$$\Big|\mathbb{E}\Big[(\nu_{e_0}(x_{e_0}) - 1) \prod_{\iota \in \{0,1\}} \Big(\prod_{e \in H \setminus \{e_0\}} g_e^{(\iota)}(x_0^{(\iota)}, x_{e \setminus \{0\}})\Big) \Big| x_0^{(0)}, x_0^{(1)} \in V_0,\ x_{e_0} \in V_{e_0}\Big]\Big|$$
$$\leq (1 + o(1)) \|\nu_{e_0} - 1\|_{U^{e_0}} \|\nu\|_\infty^r. \quad (4)$$

Proof. For each $\iota = 0,1$ and $e \in H \setminus \{e_0\}$, let $\bar{g}_e^{(\iota)}$ be either 1 or ν_e so that $g_e^{(\iota)} \leq \bar{g}_e^{(\iota)}$ holds. For $\emptyset \subseteq d \subseteq e_0$, define

$$X_d := \prod_{\omega \in \{0,1\}^d} (\nu_{e_0}(x_{e_0 \setminus d}, x_d^{(\omega)}) - 1),$$
$$Y_d := \prod_{\iota \in \{0,1\}} \prod_{\substack{e \in H \setminus \{e_0\} \\ e \supseteq d}} \prod_{\omega \in \{0,1\}^d} g_e^{(\iota)}(x_0^{(\iota)}, x_d^{(\omega)}, x_{e \setminus (d \cup \{0\})}),$$

and

$$Q_d := \mathbb{E}\big[X_d Y_d \big| x_{d \cup \{0\}}^{(0)}, x_{d \cup \{0\}}^{(1)} \in V_{d \cup \{0\}},\ x_{e_0 \setminus d} \in V_{e_0 \setminus d}\big].$$

We observe that $|Q_\emptyset|$ is equal to the left-hand side of (4) and

$$Q_{e_0} = \mathbb{E}\Big[\prod_{\omega \in \{0,1\}^{e_0}} (\nu_{e_0}(x_{e_0}^{(\omega)}) - 1) \Big| x_J^{(0)}, x_J^{(1)} \in V_J\Big] = \|\nu_{e_0} - 1\|_{U^{e_0}}^{2^r}.$$

We claim that if $j \in e_0 \setminus d$ then

$$|Q_d|^{1/2^{|d|}} \leq (1 + o(1)) Q_{d \cup \{j\}}^{1/2^{|d|+1}} \|\nu\|_\infty, \quad (5)$$

from which it would follow by induction that

$$|\text{LHS of (4)}| = |Q_\emptyset| \leq (1 + o(1)) Q_{e_0}^{1/2^r} \|\nu\|_\infty^r = (1 + o(1)) \|\nu_{e_0} - 1\|_{U^{e_0}} \|\nu\|_\infty^r$$

as desired. Now we prove (5). Let $Y_d = Y_d^{\ni j} Y_d^{\not\ni j}$ where $Y_d^{\ni j}$ consists of all the factors in Y_d that contain x_j in the argument, and $Y_d^{\not\ni j}$ consists of all other factors. Let $\overline{Y}_d^{\not\ni j}$ denote $Y_d^{\not\ni j}$ with all $g^{(\iota)}$ replaced by $\overline{g}^{(\iota)}$. Using the Cauchy-Schwarz inequality and $Y_d^{\not\ni j} \leq \overline{Y}_d^{\not\ni j}$ one has[1]

$$Q_d^2 = \mathbb{E}[\mathbb{E}[X_d Y_d^{\ni j} | x_j \in V_j] Y_d^{\not\ni j}]^2 \leq \mathbb{E}[\mathbb{E}[X_d Y_d^{\ni j} | x_j \in V_j]^2]\ \mathbb{E}[(Y_d^{\not\ni j})^2]$$
$$\leq \mathbb{E}[\mathbb{E}[X_d Y_d^{\ni j} | x_j \in V_j]^2]\ \mathbb{E}[(\overline{Y}_d^{\not\ni j})^2] = Q_{d \cup \{j\}}\ \mathbb{E}[(\overline{Y}_d^{\not\ni j})^2], \quad (6)$$

[1] The key difference between this argument and the proof of Lemma 6.3 in [1] is that here we use the Cauchy-Schwarz inequality to bound by $\mathbb{E}[\mathbb{E}[X_d Y_d^{\ni j} | x_j \in V_j]^2]\ \mathbb{E}[(\overline{Y}_d^{\not\ni j})^2]$, which contains an undesirable square $(\overline{Y}_d^{\not\ni j})^2$, whereas in [1] we bound by $\mathbb{E}[\mathbb{E}[X_d Y_d^{\ni j} | x_j \in V_j]^2 \overline{Y}_d^{\not\ni j}]\ \mathbb{E}[\overline{Y}_d^{\not\ni j}]$ so that there is no loss in terms of $\|\nu\|_\infty$.

where the outer expectations are taken over all free variables. Note that

$$\overline{Y}_d^{\not\ni j} = \prod_{\iota\in\{0,1\}} \prod_{\omega\in\{0,1\}^d} \bar{g}_{e_j}^{(\iota)}(x_0^{(\iota)}, x_d^{(\omega)}, x_{e\setminus(d\cup\{0\})})$$

is the product of at most $2^{|d|+1}$ factors of the norm ν_{e_j}. So

$$(\overline{Y}_d^{\not\ni j})^2 \le \overline{Y}_d^{\not\ni j} \sup Y_d^{\not\ni j} \le \overline{Y}_d^{\not\ni j} \left\|\nu\right\|_\infty^{2^{|d|+1}}.$$

Since $\left\|\nu_{J\setminus\{j\}} - 1\right\|_{U^{J\setminus\{j\}}} = o(1)$, Lemma 3 implies that $\mathbb{E}[\overline{Y}_d^{\not\ni j}] = 1 + o(1)$. Thus

$$\mathbb{E}[(\overline{Y}_d^{\not\ni j})^2] \le (1 + o(1)) \left\|\nu\right\|_\infty^{2^{|d|+1}}.$$

So (6) implies (5), as desired. ■

Remark. A straightforward modification of the proof shows that if $g_e \colon V_e \to \mathbb{R}_{\ge 0}$ is a function so that $g_e \le \nu_e$ or $g_e \le 1$ for every $e \in H \setminus \{e_0\}$, then

$$\left|\mathbb{E}\Big[(\nu_{e_0}(x_{e_0}) - 1) \prod_{e\in H\setminus\{e_0\}} g_e(x_e) \Big| x_J \in V_J\Big]\right| \le (1 + o(1)) \left\|\nu_{e_0} - 1\right\|_{U^{e_0}} \left\|\nu\right\|_\infty^{r/2}.$$

Indeed, in the proof, the corresponding $\overline{Y}_d^{\not\ni j}$ now has at most only $2^{|d|}$ factors, so that $(\overline{Y}_d^{\not\ni j})^2$ can be bounded by $\overline{Y}_d^{\not\ni j} \left\|\nu\right\|_\infty^{2^{|d|}}$, thereby saving a factor of 2 in the exponent of $\left\|\nu\right\|_\infty$. This implies that if $S \subseteq \mathbb{Z}_N$, $\nu = p^{-1}1_S$, and $\left\|\nu - 1\right\|_{U^r} = o(p^{r/2})$ then S contains approximately the correct count of $(r+1)$-term arithmetic progressions. This was mentioned in the concluding remarks of [1].

Lemma 6. *Assume Setup 4. Let ν be a weighted hypergraph on V satisfying*

$$\left\|\nu_e - 1\right\|_{U^e} = o(\left\|\nu\right\|_\infty^{-r+1}) \text{ for all } e \in H.$$

Define $\nu'_{e_0} \colon V_{e_0} \to \mathbb{R}_{\ge 0}$ by

$$\nu'_{e_0}(x_{e_0}) := \mathbb{E}\Big[\prod_{e\in H\setminus\{e_0\}} \nu_e(x_e) \Big| x_0 \in V_0\Big].$$

Then

$$\mathbb{E}[(\nu'_{e_0} - 1)^2] = o(1). \tag{7}$$

Expanding (7) we see that it suffices to prove the following lemma.

Lemma 7. *Assume Setup 4. Let ν be a weighted hypergraph on V satisfying*

$$\left\|\nu_e - 1\right\|_{U^e} = o(\left\|\nu\right\|_\infty^{-r+1}) \text{ for all } e \in H.$$

We have

$$\mathbb{E}\Big[\prod_{e\in H\setminus\{e_0\}} \prod_{\iota\in\{0,1\}} \nu_e(x_0^{(\iota)}, x_{e\setminus\{0\}})^{n_{e,\iota}} \Big| x_0^{(0)}, x_0^{(1)} \in V_0,\ x_{J\setminus\{0\}} \in V_{J\setminus\{0\}}\Big] = 1 + o(1)$$

for any choices of exponents $n_{e,\iota} \in \{0,1\}$.

Proof (sketch). It suffices to show, by induction on $\sum_{e,\iota} n_{e,\iota}$, that for any $j \in J \setminus \{0\}$,

$$\mathbb{E}\Big[(\nu_{e_j}(x_0^{(0)}, x_{e_j\setminus\{0\}}) - 1) \prod_{\substack{e\in H\setminus\{e_0\},\ \iota\in\{0,1\} \\ (e,\iota)\ne(e_j,0)}} \nu_e(x_0^{(\iota)}, x_{e\setminus\{0\}})^{n_{e,\iota}} \Big| x_0^{(0)}, x_0^{(1)} \in V_0,\ x_{e_0} \in V_{e_0}\Big] = o(1). \tag{8}$$

We apply the Cauchy-Schwarz inequality to bound (8), as in the proof of Lemma 5, doubling (one at a time) each vertex in $e_j \setminus \{0\}$. At each application of the Cauchy-Schwarz inequality (similar to (6)), we obtain a main factor along with a secondary factor that can be upper bounded in a way that contributes a factor of $(1+o(1))\left\|\nu\right\|_{\infty}$ to the bound of (8). After $r-1$ applications of the Cauchy-Schwarz inequality, we bound the magnitude of (8) by

$$(1+o(1))\left\|\nu\right\|_{\infty}^{r-1}\mathbb{E}\Big[\prod_{\omega\in\{0,1\}^{e_j\setminus\{0\}}}(\nu_{e_j}(x_0^{(0)},x_{e_j\setminus\{0\}}^{(\omega)})-1)\prod_{\omega\in\{0,1\}^{e_j\setminus\{0\}}}\nu_{e_j}(x_0^{(1)},x_{e_j\setminus\{0\}}^{(\omega)})^{n_{e,\iota}}\Big|x_{e_j}^{(0)},x_{e_j}^{(1)}\in V_{e_j}\Big]^{1/2^{r-1}}$$

Applying the Cauchy-Schwarz inequality one more time, we can bound the second factor by

$$\mathbb{E}\Big[\prod_{\omega\in\{0,1\}^{e_j}}(\nu_{e_j}(x_{e_j}^{(\omega)})-1)\Big|x_{e_j}^{(0)},x_{e_j}^{(1)}\in V_{e_j}\Big]^{1/2^r}\mathbb{E}\Big[\prod_{\omega\in\{0,1\}^{e_j}}\nu_{e_j}(x_{e_j}^{(\omega)})^{n_{e,\iota}}\Big|x_{e_j}^{(0)},x_{e_j}^{(1)}\in V_{e_j}\Big]^{1/2^r},$$

where the first factor is $\left\|\nu_{e_j}-1\right\|_{U^{e_j}}$ and the second factor is $1+o(1)$ by Lemma 3. It follows that the magnitude of (8) is bounded by $(1+o(1))\left\|\nu\right\|_{\infty}^{r-1}\left\|\nu_{e_j}-1\right\|_{U^{e_j}}=o(1)$. ■

References

[1] D. Conlon, J. Fox, and Y. Zhao. A relative Szemerédi theorem. Preprint.

Mathematical Institute, Oxford OX1 3LB, United Kingdom
E-mail address: `david.conlon@maths.ox.ac.uk`

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`