Combinatorics, Probability and Computing (2017) 26, 944–951.
doi:10.1017/S0963548317000190

Corners Over Quasirandom Groups

PAVEL ZORIN-KRANICH

Mathematisches Institut, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany
(e-mail: pzorin@math.uni-bonn.de)
http://www.math.uni-bonn.de/people/pzorin/

Received 3 July 2015; revised 1 April 2017; first published online 6 June 2017

Let G be a finite D-quasirandom group and $A \subset G^k$ a δ-dense subset. Then the density of the set of side lengths g of corners

$$\{(a_1,\dots,a_k),(ga_1,a_2,\dots,a_k),\dots,(ga_1,\dots,ga_k)\} \subset A$$

converges to 1 as $D \to \infty$.

2010 *Mathematics subject classification*: Primary 05D10
Secondary 20D60

1. Notation and background

In this article we will be concerned with a version of the multidimensional Szemerédi theorem over quasirandom groups. In order to state our results and put them into historical perspective, we begin by introducing appropriate notation. Let G be a countable group and let T_i be the commuting G-actions on G^k given by

$$T_i^g(a_1,\dots,a_k) := (a_1,\dots,a_{i-1},ga_i,a_{i+1},\dots,a_k).$$

We write $T^g_{[j,i]} := T^g_j \cdots T^g_i$. A (*BMZ*) *corner* in G^k is a configuration of the form

$$C(g,\vec{a}) := \{\vec{a}, T^g_{[1,1]}\vec{a},\dots,T^g_{[1,k]}\vec{a}\}, \quad g \in G, \vec{a} \in G^k. \tag{1.1}$$

We let $\vec{a}$ denote the *base point* and g the *side length* of a corner. A corner is called *non-trivial* if its side length is distinct from 1_G.

BMZ corners are not the only natural configurations generalizing the corners that appear in the commutative situation $G=\mathbb{Z}$. However, they seem to be the best behaved ones. Resolving a conjecture of Bergelson, McCutcheon and Zhang [5], Austin [4] has recently proved that if G is amenable and $A \subset G^k$ has positive upper Banach density, then A contains (many) non-trivial BMZ corners. This extends several previous results. The case $G=\mathbb{Z}$ is the multidimensional Szemerédi theorem due to Furstenberg and

Table 1. Previous results

	$\mathbb{Z}$	$\mathbb{Z}^k$	G^k
$k=2$	Roth (1953) [12]	Ajtai and Szemerédi (1974) [1]	Bergelson *et al.* (1997) [5]
$k \geqslant 3$	Szemerédi (1975) [13]	Furstenberg and Katznelson (1978) [9]	Austin (2016) [4]

Katznelson, from which the original Szemerédi theorem on arithmetic progressions in $\mathbb{Z}$ can be deduced using the projection map $\mathbb{Z}^k \to \mathbb{Z}$, $\vec{a} \mapsto a_1 + \cdots + a_k$. The cases $k = 2$ of all these results were known prior to the general cases, as indicated in Table 1. A finitary version of the multidimensional Szemerédi theorem reads as follows.

Theorem 1.1 (Bergelson and Tao [7, Theorem 11]). *Let $\delta > 0$ and $k \in \mathbb{N}$. Then there exist $\epsilon > 0$ and $N \in \mathbb{N}$ such that, for every finite group G with $|G| > N$, every subset $A \subset G^k$ with $|A| > \delta|G|^k$ contains at least $\epsilon|G|^{k+1}$ BMZ corners.*

This theorem is an easy consequence of Gowers' hypergraph removal lemma [10], and we reproduce the proof here in order to motivate both the definition of the BMZ corners and the change of variables that will be used in the proof of Theorem 2.2.

Proof of Theorem 1.1. Here and later we use a subscript to denote omission of the ith coordinate in a vector, as follows:

$$\vec{x}_{(i)} = (x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$$

For $i = 0, \ldots, k$ consider the changes of variables

$$N_i^{[0,k]}(\vec{x}_{(i)}) := (x_0, x_0x_1, \ldots, (x_0 \cdots x_{i-1}), (x_{i+1} \cdots x_k)^{-1}, \ldots, x_k^{-1}).$$

They are related to each other as follows: if $x \in G^{k+1}$ and $g = x_0 \cdots x_k$, then

$$T^g_{[1,i]} N_0^{[0,k]}(x_{(0)}) = N_i^{[0,k]}(x_{(i)}).$$

In particular, corners are precisely the configurations

$$\{N_i^{[0,k]}(x_{(i)}), i = 0, \ldots, k\}, \quad x \in G^{k+1}.$$

Define a $(k+1)$-partite k-uniform hypergraph F with vertex sets $X_0 = \cdots = X_k = G$ by

$$\vec{x}_{(i)} \in F :\Longleftrightarrow N_i^{[0,k]}(\vec{x}_{(i)}) \in A.$$

Then a corner corresponds to a simplex in the hypergraph F. If there were fewer than $\epsilon|G|^{k+1}$ simplices in F, then by the hypergraph removal lemma [10, Theorem 10.1] the hypergraph F could be made simplex-free by removing fewer than $\delta|G|^k$ edges. But if we remove the element of A corresponding to each removed edge and repeat the construction of F, we obtain an even smaller hypergraph that still contains simplices (since each remaining member of A gives rise to a trivial corner), a contradiction. □

A similar argument works for $A \subset \Phi^k$, where $\Phi \subset G$ is a set with $|\Phi^{-1}\Phi| \leqslant C|\Phi|$, with constants depending on C. This proves a version of Theorem 1.1 over infinite amenable groups that admit a Følner sequence satisfying the Tempelman condition. This argument does not seem to extend easily to general Følner sequences.

2. Main result

The problem of finding arithmetic progressions, and later more general configurations, in dense subsets of amenable groups was transferred to ergodic theory by Furstenberg [8], who reformulated Szemerédi's theorem as a multiple recurrence theorem and gave it a new proof. An important special case of the multiple recurrence theorem occurs for weakly mixing actions, when its conclusion can be strengthened to the extent that corners with almost every possible side length can be found.

A (necessarily infinite) group is called *weakly mixing* if it has no non-trivial finite-dimensional unitary representations. For such groups many combinatorial results can be strengthened substantially; see *e.g.* Bergelson and Tao [7]. A quantitative notion of weak mixing was introduced by Gowers [11]. A group is called *D-quasirandom* if it has no non-trivial unitary representation of dimension less than D. Our result says that in dense subsets over quasirandom groups one can find corners of almost every side length.

Theorem 2.1. *Let $\delta > 0$ and $k \in \mathbb{N}$. Then, for every finite D-quasirandom group G and every subset $A \subset G^k$ with $|A| > \delta |G|^k$, we have*

$$|\{g \in G : |\{\vec{a} \in G^k : C(g,\vec{a}) \subset A\}| > \epsilon(\delta,k)|G|^k/2\}| > (1 - o_{D\to\infty;\delta,k}(1))|G|,$$

where $\epsilon(\delta,k)$ is the quantity from Theorem 1.1.

The case $k = 2$ was previously shown in [3, 6], and we refer to those articles for further discussion of why BMZ corners are natural.

Since the set $\{\vec{a} \in G^k : C(g,\vec{a}) \subset A\}$ has density at least ϵ on average (over g) by Theorem 1.1, it suffices to show that its density is usually close to the average. We formulate this in the language of dynamical systems as a multiple weak mixing property.

Theorem 2.2. *Let G be a compact D-quasirandom group, $k \geqslant 0$, and $f_i : G^k \to [-1,1]$, $i \in [0,k]$. Consider the multicorrelation sequence*

$$c_g := \int_{G^k} \prod_{i=0}^{k} f_i T^g_{[1,i]}.$$

Then

$$\int_g |c_g - \textstyle\int_h c_h| = o_{D\to\infty;k}(1).$$

In other words, the multicorrelation sequence converges to its average *in density* as $D \to \infty$. Here and later, compact groups are equipped with the normalized Haar measure and $fT = f \circ T$ denotes the composition of functions f and T.

https://doi.org/10.1017/S0963548317000190 Published online by Cambridge University Press

Proof of Theorem 2.1 assuming Theorem 2.2. Let $f_0 = \cdots = f_k = 1_A$. Then the multicorrelation sequence c_g counts the elements $\vec{a} \in G^k$ such that $C(g,\vec{a}) \subset A$. On the other hand, by Theorem 1.1 we have $\int_g c_g > \epsilon(\delta,k)$ provided $|G| > D$ is large enough. □

3. Tools

In this and the next section G always denotes a compact group with normalized Haar measure. Quasirandomness will be used in the following form.

Lemma 3.1 (Austin [2, Corollary 3]). *Let V be a (real or complex) Hilbert space equipped with an (orthogonal or unitary) right action of a compact D-quasirandom group G, and let P be the projection onto the invariant subspace. Then for every $u, v \in V$ we have*

$$\int_G |\langle u, vg\rangle - \langle Pu, Pv\rangle|^2 \mathrm{d}g \leqslant D^{-1}\|u\|^2\|v\|^2.$$

This result was stated for left actions by Austin [2]; the version above follows by considering either the adjoint action or the opposite group.

We use the following version of the van der Corput lemma.

Lemma 3.2 (Austin [3, Lemma 1]). *Let V be a (real or complex) Hilbert space and $u : G \to V$ a measurable function. Then for every $v \in V$ with $\|v\| \leqslant 1$ we have*

$$\int |\langle v, u(g)\rangle| \mathrm{d}g \leqslant \sqrt{\iint |\langle u(g), u(h)\rangle| \mathrm{d}g\mathrm{d}h}.$$

For a function $F : G^k \to \mathbb{R}$, $k \geqslant 1$, the k-variable Gowers box norm is defined by

$$\|F\|_{\square^k}^{2^k} = \int \prod_{\vec{\epsilon}\in\{0,1\}^{[1,k]}} F(\vec{x}_{(i)}^{\epsilon}) \mathrm{d}(x_{j,\epsilon})_{j\in[1,k],\epsilon\in\{0,1\}},$$

where $\vec{x}_j^{\epsilon} = x_{j,\epsilon_j}$ and $[1,k] = \{1,\ldots,k\}$. See, for example, Tao [16] for a discussion of the basic properties of these norms.

Recall a version of the (weak) weighted hypergraph regularity lemma [14, Lemma 2.9]. This particular version can be found in Tao [15, Corollary 6.8] for $k = 2$, and the proof for general k is similar.

Lemma 3.3 (weak regularity lemma). *For every $k \in \mathbb{N}$ and $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that every measurable function $F : G^k \to [-1,1]$ can be written as $F = F_s + F_u$, where*

(1) *F_s is measurable with respect to $\vee_{j=1}^k B_j$, where each B_j is a σ-algebra on G^k generated by at most M atoms that does not depend on the jth coordinate,*
(2) *$\|F_u\|_{\square^k} \leqslant \epsilon$, and*
(3) *$|F_s|, |F_u| \leqslant 2$.*

https://doi.org/10.1017/S0963548317000190 Published online by Cambridge University Press

4. Multiple weak mixing

Theorem 2.2 is proved by induction on k via the following steps.

(1) Prove a Gowers box norm estimate for the average in question.
(2) Apply the hypergraph regularity lemma to split one of the functions into a structured and a quasirandom part.
(3) Estimate the quasirandom part using step (1) and the structured part using the inductive hypothesis.

Step (1) in this plan is given by the following estimate.

Proposition 4.1. *Let G be a compact D-quasirandom group and $k \geqslant 1$. Then for every tuple of functions $f_i : G^k \to [-1,1]$, $i \in [0,k]$, we have*

$$\int_G \Big| \textstyle\int_{G^k} \displaystyle\prod_{i=0}^{k} f_i T^g_{[1,i]} \Big| \mathrm{d}g \leqslant \min_i \| f_i N_i^{[0,k]} \|_{\square^k} + C_k D^{-2^{-k}}.$$

Proof. By induction on k. For $k=1$ the box norm is just the absolute value of the integral, so writing

$$\int_G \Big| \textstyle\int_{G^1} \displaystyle\prod_{i=0}^{1} f_i T^g_{[1,i]} \Big| \mathrm{d}g \leqslant \int_G \big| \textstyle\int_{G^1} f_0 \cdot f_1 T^g_1 - \int f_0 \int f_1 \big| \mathrm{d}g + \big|\int f_0\big| \big|\int f_1\big|$$

we can estimate the second term by the minimum of the box norms. In the first term we apply Jensen's inequality and Lemma 3.1 with the Hilbert space $L^2(G)$ and the unitary right G-action $(g,f) \mapsto f T^g_1$. Since the invariant subspace of this action consists only of the constant functions, the projection onto this subspace amounts to integration over G.

Suppose now that $k > 1$ and the claim is known for $k-1$. Applying $T^{g^{-1}}_{[1,k]}$ to the function in the inner integral and reversing the order of the indices $0,\dots,k$, we see that the bound with f_0 follows from the bound with f_k, so it suffices to establish bounds with $f_1,\dots,f_k$.

Applying Lemma 3.2 with $X = G$, $V = L^2(G^k)$, $v = f_0$, and $u(g) = \prod_{i=1}^k f_i T^g_{[1,i]}$, we estimate the square of the left-hand side of the conclusion by

$$\int_h \int_g \Big| \textstyle\int_{G^k} \displaystyle\prod_{i=1}^{k} f_i T^g_{[1,i]} \cdot f_i T^h_{[1,i]} \Big| = \int_h \int_g \Big| \textstyle\int_{G^k} \displaystyle\prod_{i=1}^{k} (f_i \cdot f_i T^h_{[1,i]}) T^g_{[2,i]} \Big|.$$

In the last step we have made the change of variables $(g,h) \mapsto (g,hg)$ on G^2 and used the fact that T^g_1 is a measure-preserving transformation. Pulling one of the integrals out of the absolute value, we obtain the estimate

$$\int_h \int_g \int_{a_1} \Big| \textstyle\int_{a_2,\dots,a_k} \displaystyle\prod_{i=1}^{k} (f_i \cdot f_i T^h_{[1,i]}) T^g_{[2,i]} (a_1,\dots,a_k) \Big|.$$

https://doi.org/10.1017/S0963548317000190 Published online by Cambridge University Press

Applying the inductive hypothesis for each fixed pair (h, a_1), for any $i \in [1, k]$ we obtain the estimate

$$\int_h \int_{a_1} \|(f_i \cdot f_i T^h_{[1,i]})(a_1, N_i^{[1,k]} \cdot)\|_{\square^{k-1}} + C_{k-1} D^{-2^{-k+1}},$$

where $N_i^{[1,k]}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k)$ is defined similarly to $N_i^{[0,k]}$.

The contribution of the second summand is admissible, so we only have to consider the first summand. Raising it to the power 2^{k-1} and applying Jensen's inequality, we obtain the bound

$$\int_{h,a_1} \|(f_i \cdot f_i T^h_{[1,i]})(a_1, N_i^{[1,k]} \cdot)\|^{2^{k-1}}_{\square^{k-1}}.$$

Expanding the definition of the box norm and observing that

$$T^h_{[1,i]}(a_1, N_i^{[1,k]} \vec{x}_{(i)}) = N_i^{[0,k]}(ha_1, a_1^{-1} x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k),$$

we can write the above expression in the form

$$\int_{h,a_1} \int \prod_{\vec{\epsilon} \in \{0,1\}^{[1,k] \setminus \{i\}}} f_i N_i^{[0,k]}(a_1, a_1^{-1} x_{1,\epsilon_1}, x_{2,\epsilon_2}, \ldots, x_{k,\epsilon_k})$$
$$\times f_i N_i^{[0,k]}(ha_1, a_1^{-1} x_{1,\epsilon_1}, x_{2,\epsilon_2}, \ldots, x_{k,\epsilon_k}) \mathrm{d}(x_{j,\epsilon})_{j \in [1,k] \setminus \{i\}, \epsilon \in \{0,1\}}.$$

With the change of variables $(a_1, x_{1,0}, x_{1,1}) \mapsto (a_1, a_1 x_{1,0}, a_1 x_{1,1})$, this becomes

$$\int_{h,a_1} \int \prod_{\epsilon} f_i N_i^{[0,k]}(a_1, \vec{x}^{\epsilon}_{(i)}) \cdot \prod_{\epsilon} f_i N_i^{[0,k]}(ha_1, \vec{x}^{\epsilon}_{(i)}) \mathrm{d}(x_{j,\epsilon})_{j \in [1,k] \setminus \{i\}, \epsilon \in \{0,1\}}.$$

We interpret the integral in all variables but h as an inner product in $L^2(G^{2k-1})$ and the appearance of h in the first argument of the second product as a right unitary action of G on this space. Applying Lemma 3.1, we obtain an admissible error term and the bound

$$\int_{a_1} \int P\left(\prod_{\epsilon} f_i N_i^{[0,k]}\right)^2 (a_1, \vec{x}^{\epsilon}_{(i)}) \mathrm{d}(x_{j,\epsilon})_{j \in [1,k] \setminus \{i\}, \epsilon \in \{0,1\}},$$

where P denotes the projection onto the invariant subspace. But this projection is simply integration in the variable a_1, so this can be written as

$$\int \left(\int_{a_1} \prod_{\epsilon} f_i N_i^{[0,k]}(a_1, \vec{x}^{\epsilon}_{(i)})\right)^2 \mathrm{d}(x_{j,\epsilon})_{j \in [1,k] \setminus \{i\}, \epsilon \in \{0,1\}}.$$

Relabelling $a_1 = x_{0,0}$ in the first factor of the square and $a_1 = x_{0,1}$ in the second factor, we see that this coincides with

$$\|f_i N_i^{[0,k]}\|^{2^k}_{\square^k},$$

and the conclusion follows. □

https://doi.org/10.1017/S0963548317000190 Published online by Cambridge University Press

Proof of Theorem 2.2. By induction on k. The base case $k = 0$ is very easy.

Now let $k \geqslant 1$, and suppose that the result holds for $k-1$. Let $\epsilon > 0$ be arbitrary and apply the weak regularity lemma to the function $f_k N_k^{[0,k]}$, so that

$$f_k N_k^{[0,k]} = \sum_{l \in L} \prod_{i=0}^{k-1} F_{i,l} + F_u,$$

where all functions on the right-hand side are uniformly bounded, the functions $F_{i,l}$ do not depend on the ith and the kth coordinates, the index set L has size $O_\epsilon(1)$, and $\|F_u\|_{\square^k} \leqslant \epsilon$. Split the multicorrelation sequence accordingly as

$$c_g = \sum_{l \in L} c_{l,g} + c_{u,g}.$$

The inverse of the change of variables $N_i^{[0,k]}$ is given by

$$(N_i^{[0,k]})^{-1}(\vec{a}) = (a_1, a_1^{-1}a_2, \dots, a_{i-1}^{-1}a_i, a_{i+1}^{-1}a_{i+2}, \dots, a_{k-1}^{-1}a_k, a_k^{-1}),$$

and it can be verified that we have

$$((N_k^{[0,k]})^{-1} T_{[1,k]}^g)_{(i)} = ((N_k^{[0,k]})^{-1} T_{[1,i]}^g)_{(i)}.$$

Thus the actions $T_{[1,k]}$ and $T_{[1,i]}$ coincide on the functions $F_{i,l}$. This is a common theme in the hypergraph regularity approach to multiple ergodic averages in the work of Austin (although it takes much less effort to exploit this phenomenon in our compact group setting than in the setting of infinite amenable groups). Since the maps $f \mapsto f T_{[1,i]}^g$ are algebra homomorphisms, it follows that

$$c_{l,g} = \int_{G^k} \prod_{i=0}^{k-1} (f_i \cdot F_{i,l}(N_k^{[0,k]})^{-1}) T_{[1,i]}^g.$$

This is an average (in the last coordinate of G^k) of multicorrelation sequences of length $k-1$, so its total variation is bounded by $o_{D\to\infty;k-1}(1)$ by the inductive hypothesis. On the other hand, we have

$$\int_g |c_{u,g}| \leqslant \epsilon + o_{D\to\infty;k}(1)$$

by Proposition 4.1. This shows that the total variation of the multicorrelation sequence c_g can be estimated by

$$|L| o_{D\to\infty;k-1}(1) + \epsilon + o_{D\to\infty;k}(1) = \epsilon + o_{D\to\infty;k,\epsilon}(1).$$

Since $\epsilon > 0$ was arbitrary, this concludes the proof. □

Acknowledgement

The anonymous referee's constructive criticism has led to a greatly improved presentation.

https://doi.org/10.1017/S0963548317000190 Published online by Cambridge University Press

References

[1] Ajtai, M. and Szemerédi, E. (1974) Sets of lattice points that form no squares. *Studia Sci. Math. Hungar.* **9** 9–11.

[2] Austin, T. (2015) Quantitative equidistribution for certain quadruples in quasi-random groups. *Combin. Probab. Comput.* **24** 376–381. With erratum.

[3] Austin, T. (2016) Ajtai–Szemerédi theorems over quasirandom groups. In *Recent Trends in Combinatorics*, Vol. 159 of The IMA Volumes in Mathematics and its Applications, Springer, pp. 453–484.

[4] Austin, T. (2016) Non-conventional ergodic averages for several commuting actions of an amenable group. *J. Analyse Math.* **130** 243–274. arXiv:1309.4315

[5] Bergelson, V., McCutcheon, R. and Zhang, Q. (1997) A Roth theorem for amenable groups. *Amer. J. Math.* **119** 1173–1211.

[6] Bergelson, V., Robertson, D. and Zorin-Kranich, P. (2017) Triangles in Cartesian squares of quasirandom groups. *Combin. Probab. Comput.* **26** 161–182.

[7] Bergelson, V. and Tao, T. (2014) Multiple recurrence in quasirandom groups. *Geom. Funct. Anal.* **24** 1–48.

[8] Furstenberg, H. (1977) Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.* **31** 204–256.

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

[10] Gowers, W. T. (2007) Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math.* (2) **166** 897–946.

[11] Gowers, W. T. (2008) Quasirandom groups. *Combin. Probab. Comput.* **17** 363–387.

[12] Roth, K. F. (1953) On certain sets of integers. *J. London Math. Soc.* **28** 104–109.

[13] Szemerédi, E. (1975) On sets of integers containing no k elements in arithmetic progression. *Acta Arith.* **27** 199–245.

[14] Tao, T. (2006) A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A* **113** 1257–1280.

[15] Tao, T. (2007) The ergodic and combinatorial approaches to Szemerédi's theorem. In *Additive Combinatorics*, Vol. 43 of CRM Proc. Lecture Notes, AMS, pp. 145–193.

[16] Tao, T. (2012) *Higher Order Fourier Analysis*, Vol. 142 of Graduate Studies in Mathematics, AMS.