Electronic Colloquium on Computational Complexity, Report No. 142 (2021)

Mixing of 3-term progressions in Quasirandom Groups

Amey Bhangale * Prahladh Harsha † Sourya Roy *

Abstract

In this note, we show the mixing of three-term progressions (x, xg, xg^2) in every finite quasirandom group, fully answering a question of Gowers. More precisely, we show that for any D-quasirandom group G and any three sets $A_1, A_2, A_3 \subset G$, we have

$$\left| \Pr_{x,y\sim G}\left[x \in A_1, xy \in A_2, xy^2 \in A_3\right] - \prod_{i=1}^{3} \Pr_{x\sim G}\left[x \in A_i\right] \right| \leq \left(\frac{2}{\sqrt{D}}\right)^{1/4}.$$

Prior to this, Tao answered this question when the underlying quasirandom group is $\mathrm{SL}_d(\mathbb{F}_q)$. Subsequently, Peluse extended the result to all nonabelian finite *simple* groups. In this work, we show that a slight modification of Peluse's argument is sufficient to fully resolve Gower's quasirandom conjecture for 3-term progressions. Surprisingly, unlike the proofs of Tao and Peluse, our proof is elementary and only uses basic facts from nonabelian Fourier analysis.

1 Introduction

In this note, we revisit a conjecture by Gowers [Gow08] about mixing of three term arithmetic progressions in quasirandom finite groups. Gowers initiated the study of quasirandom groups while refuting a conjecture of Babai and Sós [BS85] regarding the size of the largest product-free set in a given finite group. A finite group is said to be D-quasirandom for a positive integer $D > 1$ if all its non-trivial irreducible representations are at least D-dimensional. The quasirandomness property of groups can be used to show that certain "objects" related to the group "mix" well. For instance, the quasirandomness of the group $\mathrm{PSL}_2(\mathbb{F}_q)$ can be used to give an alternate (and weaker) proof [DSV03] that the Ramanujan graphs of Lubotzky, Philips and Sarnak [LPS88] are expanders. Bourgain and Gambard [BG08] used quasirandomness to prove that certain other Cayley graphs were expanders.

Gowers proved that for any D-quasirandom group G and any three subsets $A, B, C \subset G$ satisfying $|A| \cdot |B| \cdot |C| \geq |G|^3/D$, there exist $a \in A, b \in B, c \in C$ such that $ab = c$. More generally, he proved that the number of such triples $(a, b, c) \in A \times B \times C$ such that $ab = c$ is at least $(1-\eta)|A| \cdot |B| \cdot |C|/|G|$ provided $|A| \cdot |B| \cdot |C| \geq |G|^3/\eta^2 D$. In other words the set of triples of the form (a, b, ab) mix well in a quasirandom group. Gowers' proof of this result was the inspiration and the first step towards the recent optimal inapproximability result for satisfiable kLIN over nonabelian groups [BK21]. After proving

*University of California, Riverside, CA, USA. Email : ameyb@ucr.edu, sourya.roy@email.ucr.edu
†Tata Institute of Fundamental Research, Mumbai, India. Email : prahladh@tifr.res.in

ISSN 1433-8092

the well-mixing of triples of the form (a, b, ab) in quasirandom groups, Gowers conjectured a similar statement for triples of the form (x, xg, xg^2). More precisely, he conjectured the following statement: Let G be a D-quasirandom group and $f_1, f_2, f_3 : G \to \mathbb{C}$ such that $\|f_i\|_\infty \leq 1$, then

$$\left| \mathop{\mathbb{E}}_{x,y\sim G}\left[f_1(x)f_2(xy)f_3(xy^2)\right] - \prod_{i=1,2,3} \mathop{\mathbb{E}}_{x\sim G}\left[f_i(x)\right] \right| = o_D(1), \tag{1}$$

where the expression $o_D(1)$ goes to zero as D increases. The conjecture can be naturally extended to k-term arithmetic progressions and product of k functions for $k > 3$. However, in this note we will focus on the three term case.

For the specific case of 3-term progressions, Tao [Tao13] proved the conjecture for the group $\mathrm{SL}_d(\mathbb{F}_q)$ for bounded d using algebraic geometric machinery. In particular, he proved that the right-hand side expression in Eq. (1) can be bounded by $O(1/q^{1/8})$ when $d = 2$ and $O_d(1/q^{1/4})$ for larger d. Tao's approach relied on algebraic geometry and was not amenable to other quasirandom groups. Later, Peluse [Pel18] proved the conjecture for all nonabelian finite simple groups. She used basic facts from nonabelian Fourier analysis to prove that the right-hand side expression in Eq. (1) can be bounded by $\sum_{1\neq\rho\in\hat{G}} 1/d_\rho$ where $\hat{G}$ represents the set of irreducible unitary representation of G and d_ρ the dimension of the irreducible representation ρ. This latter quantity is the *Witten zeta function* ζ_G of the group G minus one and can be bounded for *simple* finite quasirandom groups using a result due to Liebeck and Shalev [LS04].

In this paper, we show that a slight variation of Peluse's argument can be used to prove the conjecture for *all quasirandom groups* with *better* error parameters. More surprisingly, the proof stays completely elementary and short. Specifically, we prove the following statement:

Theorem 1. *Let G be a D-quasirandom finite group, i.e, its all non-trivial irreducible representations are at least D-dimensional. Let $f_1, f_2, f_3 : G \to \mathbb{C}$ such that $\|f_i\|_\infty \leq 1$ then*

$$\left| \mathop{\mathbb{E}}_{x,y\sim G}\left[f_1(x)f_2(xy)f_3(xy^2)\right] - \prod_{i=1,2,3} \mathop{\mathbb{E}}_{x\sim G}\left[f_i(x)\right] \right| \leq \left(\frac{2}{\sqrt{D}}\right)^{\frac{1}{4}}.$$

2 Preliminaries

We begin by recalling some basic representation theory and nonabelian Fourier analysis. See the monograph by Diaconis [Dia98, Chapter 2] for a more detailed treatment (with proofs).

We will be working with a finite group G and complex-valued functions $f\colon G \to \mathbb{C}$ on G. All expectations will be with respect to the uniform distribution on G. The *convolution* between two function $f, h\colon G \to \mathbb{C}$, denoted by $f * h$, is defined as follows:

$$(f * h)(x) := \mathop{\mathbb{E}}_{y}[f(xy^{-1})h(y)].$$

For any $p \geq 1$, the p-norm of any function $f\colon G \to \mathbb{C}$ is defined as

$$\|f\|_p^p := \mathop{\mathbb{E}}_{x}[|f(x)|^p].$$

For any element $g \in G$, the *conjugacy class of* g, denoted by $C(g)$, refers to the set $\{x^{-1}gx | x \in G\}$. Observe that the conjugacy classes form a partition of the group G. A function $f: G \to \mathbb{C}$ is said to be a *class function* if it is constant on conjugacy classes.

For any $b \in G$ we use $\Delta_b f(x) := f(x) \cdot f(xb)$. For any set $S \subset G$, $\mu_S: G \to \mathbb{R}$ denotes the scaled density function $\frac{|G|}{|S|}\mathbb{1}_S$. The scaling ensures that $\mathbb{E}_x[\mu_S(x)] = 1$.

Given a complex vector space V, we denote the vector space of linear operators on V by $\mathrm{End}(V)$. This space is endowed with the following inner product and norm (usually referred to as the *Hilbert-Schmidt* norm):

$$\text{For } A, B \in \mathrm{End}(V), \quad \langle A, B\rangle_{\mathrm{HS}} := \mathrm{Trace}(A^* B) \quad \text{and} \quad \|A\|^2_{\mathrm{HS}} := \langle A, A\rangle_{\mathrm{HS}} = \mathrm{Trace}(A^* A).$$

This norm is known to be submultiplicative (i.e, $\|AB\|_{\mathrm{HS}} \leq \|A\|_{\mathrm{HS}} \cdot \|B\|_{\mathrm{HS}}$).

Representations and Characters: A *representation* $\rho: G \to \mathrm{End}(V)$ is a homomorphism from G to the set of linear operators on V for some finite-dimensional vector space V over $\mathbb{C}$, i.e., for all $x, y \in G$, we have $\rho(xy) = \rho(x)\rho(y)$. The dimension of the representation ρ, denoted by d_ρ, is the dimension of the underlying $\mathbb{C}$-vector space V. The *character* of a representation ρ, denoted by $\chi_\rho: G \to \mathbb{C}$, is defined as $\chi_\rho(x) := \mathrm{Trace}(\rho(x))$.

The representation $1: G \to \mathbb{C}$ satisfying $1(x) = 1$ for all $x \in G$ is the *trivial* representation. A representation $\rho: G \to \mathrm{End}(V)$ is said to *reducible* if there exists a non-trivial subpsace $W \subset V$ such that for all $x \in G$, we have $\rho(x)W \subset W$. A representation is said to be *irreducible* otherwise. The set of all irreducible representations of G (upto equivalences) is denoted by $\hat{G}$.

For every representation $\rho: G \to \mathrm{End}(V)$, there exists an inner product $\langle , \rangle_V$ over V such that every $\rho(x)$ is unitary (i.e, $\langle \rho(x)u, \rho(x)v\rangle_V = \langle u, v\rangle_V$ for all $u, v \in V$ and $x \in G$). Hence, we might wlog. assume that all the representations we are considering are unitary.

The following are some well-known facts about representations and characters.

Proposition 2. *1. The group G is abelian iff $d_\rho = 1$ for every irreducible representation ρ in $\hat{G}$.*

2. For any finite group G, $\sum_{\rho \in \hat{G}} d_\rho^2 = |G|$.

3. [orthogonality of characters] For any $\rho, \rho' \in \hat{G}$ we have: $\mathbb{E}_x\left[\chi_\rho(x)\overline{\chi_{\rho'}(x)}\right] = \mathbb{1}[\rho = \rho']$.

Definition 3 (quasirandom groups)**.** *A nonabelian group G is said to be D-*quasirandom *for some positive integer $D > 1$ if all its non-trivial irreducible representations ρ satisfy $d_\rho \geq D$.*

Nonabelian Fourier analysis: Given a function $f: G \to \mathbb{C}$ and an irreducible representation $\rho \in \hat{G}$, the Fourier transform is defined as follows:

$$\hat{f}(\rho) := \mathop{\mathbb{E}}_x[f(x)\rho(x)].$$

The following proposition summarizes the basic properties of Fourier transform that we will need.

Proposition 4. *For any $f, h: G \to \mathbb{C}$, we have the following*

1. *[Fourier transform of trivial representation]*
$$\hat{f}(1) = \mathop{\mathbb{E}}_{x}[f(x)].$$

2. *[Convolution]*
$$\widehat{f * h}(\rho) = \hat{f}(\rho) \cdot \hat{h}(\rho).$$

3. *[Fourier inversion formula]*
$$f(x) = \sum_{\rho \in \hat{G}} d_\rho \cdot \langle \hat{f}(\rho), \rho(x) \rangle_{\mathrm{HS}}.$$

4. *[Parseval's identity]*
$$\|f\|_2^2 = \sum_{\rho \in \hat{G}} d_\rho \cdot \|\hat{f}(\rho)\|_{\mathrm{HS}}^2.$$

5. *[Fourier transfrom of class functions] For any class function $f : G \to \mathbb{C}$, the Fourier transform satisfies*
$$\hat{f}(\rho) = c \cdot I_{d_\rho}$$
for some constant $c = c(f, \rho) \in \mathbb{C}$. In other words, the Fourier transform is a scaling of the Identity operator I_{d_ρ}.

The following claim (also used by Peluse [Pel18]) observes that the scaled density function $\mu_{gC(g)}$ has a very simple Fourier transform since it is a translate of the class function $\mu_{C(g)}$

Claim 5. *For any $g \in G$ and $\rho \in \hat{G}$ we have:*
$$\hat{\mu}_{gC(g)}(\rho) = \frac{\chi_\rho(g)}{d_\rho} \cdot \rho(g)$$
where $C(g)$ refers to the conjugacy class of g. Moreover, $\|\hat{\mu}_{gC(g)}\|_{\mathrm{HS}}^2 = \frac{|\chi_\rho(g)|^2}{d_\rho}$

Proof. We begin by observing that
$$\begin{aligned} \hat{\mu}_{gC(g)}(\rho) &= \mathop{\mathbb{E}}_{x}\left[\mu_{gC(g)}(x) \cdot \rho(x)\right] \\ &= \mathop{\mathbb{E}}_{x}\left[\mu_{gC(g)}(gx) \cdot \rho(gx)\right] \\ &= \mathop{\mathbb{E}}_{x}\left[\mu_{gC(g)}(gx) \cdot \rho(g) \cdot \rho(x)\right] \\ &= \rho(g) \cdot \mathop{\mathbb{E}}_{x}\left[\mu_{C(g)}(x) \cdot \rho(x)\right] \\ &= \rho(g) \cdot \hat{\mu}_{C(g)}(\rho). \end{aligned}$$

On the other hand, as $\mu_{C(g)}$ is a class function, we have $\hat{\mu}_{C(g)}(\rho) = c \cdot I_{d_\rho}$ for some constant $c \in \mathbb{C}$. The constant c can be determined by taking trace on either side of $c\dot{I}_{d_\rho} = \hat{\mu}_{C(g)} = \mathbb{E}_x[\mu_{C(g)}(x) \cdot \rho(x)]$ and noting that $\mathrm{Trace}(\rho(x)) = \chi_\rho(g)$ as follows:
$$c \cdot d_\rho = \mathop{\mathbb{E}}_{x}\left[\mu_{C(g)}(x) \cdot \chi_\rho(x)\right] = \mathop{\mathbb{E}}_{x}\left[\mu_{C(g)}(x)\right] \cdot \chi_\rho(g) = \chi_\rho(g).$$

Hence, $c = \frac{\chi_\rho(g)}{d_\rho}$ and $\hat{\mu}_{gC(g)} = \frac{\chi_\rho(g)}{d_\rho} \cdot \rho(g)$. Lastly we have,

$$\begin{aligned}
\|\hat{\mu}_{gC(g)}\|_{\mathrm{HS}}^2 &= \left\| \frac{\chi_\rho(g)}{d_\rho} \cdot \rho(g) \right\|_{\mathrm{HS}}^2 \\
&= \frac{|\chi_\rho(g)|^2}{d_\rho^2} \cdot \mathrm{Trace}\left(\rho(g)^* \cdot \rho(g)\right) \\
&= \frac{|\chi_\rho(g)|^2}{d_\rho^2} \cdot d_\rho && \text{(By unitariness of } \rho(g)\text{)} \\
&= \frac{|\chi_\rho(g)|^2}{d_\rho}. && \square
\end{aligned}$$

The key property of D-quasirandom groups that we will be using is the following inequality due to Babai, Nikolov and Pyber, the proof of which we provide for the sake of completeness.

Lemma 6 ([BNP08])**.** *If G is a D-quasirandom group and $f_1, f_2 \colon G \to \mathbb{C}$ such that either f_1 or f_2 is mean zero then*

$$\|f_1 * f_2\|_2 \leq \frac{1}{\sqrt{D}} \cdot \|f_1\|_2 \cdot \|f_2\|_2.$$

Proof.

$$\begin{aligned}
\|f_1 * f_2\|^2 &= \sum_{\rho \in \hat{G}} d_\rho \|\widehat{f_1 * f_2}(\rho)\|_{\mathrm{HS}}^2 \\
&= \sum_{\rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho) \cdot \hat{f}_2(\rho)\|_{\mathrm{HS}}^2 \\
&\leq \sum_{\rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho)\|_{\mathrm{HS}}^2 \cdot \|\hat{f}_2(\rho)\|_{\mathrm{HS}}^2 && \text{(By submultiplicativity of norm)} \\
&= \sum_{1 \neq \rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho)\|_{\mathrm{HS}}^2 \cdot \|\hat{f}_2(\rho)\|_{\mathrm{HS}}^2 && \text{(By mean zeroness)} \\
&\leq \frac{1}{D} \cdot \sum_{1 \neq \rho \in \hat{G}} d_\rho^2 \|\hat{f}_1(\rho)\|_{\mathrm{HS}}^2 \cdot \|\hat{f}_2(\rho)\|_{\mathrm{HS}}^2 && \text{(By } D\text{-quasirandomness)} \\
&\leq \frac{1}{D} \left(\sum_{1 \neq \rho \in \hat{G}} d_\rho \|\hat{f}_1(\rho)\|_{\mathrm{HS}}^2 \right) \cdot \left(\sum_{1 \neq \rho \in \hat{G}} d_\rho \|\hat{f}_2(\rho)\|_{\mathrm{HS}}^2 \right) \\
&\leq \frac{1}{D} \cdot \|f_1\|_2^2 \cdot \|f_2\|_2^2 . && \square
\end{aligned}$$

The following is a simple corollary of Lemma 6.

Corollary 7. *If G is D-quasirandom; $f \colon G \to \mathbb{C}$ has zero mean and $\|f\|_\infty \leq 1$ then*

$$\mathop{\mathbb{E}}_b \left[\left| \mathop{\mathbb{E}}_x \Delta_b f(x) \right| \right] \leq \frac{1}{\sqrt{D}}.$$

Proof. Let $f'(x) := f(x^{-1})$. We have,

$$\begin{aligned}
\mathop{\mathbb{E}}_{b}\left[\left|\mathop{\mathbb{E}}_{x}\Delta_b f(x)\right|\right] &= \mathop{\mathbb{E}}_{b}\left[\left|\mathop{\mathbb{E}}_{x} f(x)f(xb)\right|\right] \\
&= \mathop{\mathbb{E}}_{b}\left[\left|\mathop{\mathbb{E}}_{x} f'(x^{-1})f(xb)\right|\right] \\
&= \mathop{\mathbb{E}}_{b}\left[\left|f' * f(b)\right|\right] \\
&\le \mathop{\mathbb{E}}_{b}\left[\left|f' * f(b)\right|^2\right]^{1/2} && \text{(By Cauchy-Schwarz inequality)} \\
&= \|f' * f\|_2 \\
&\le \frac{1}{\sqrt{D}} \cdot \|f'\|_2 \cdot \|f\|_2 && \text{(By Lemma 6)} \\
&\le \frac{1}{\sqrt{D}}. && \text{(Since } \|f\|_2 \le \|f\|_\infty \le 1\text{).}
\end{aligned}$$

□

3 Proof of Theorem 1

The following proposition is where we deviate from Peluse's proof [Pel18]. We give an elementary proof for *every* quasirandom group while Peluse proved the same result for *simple* finite groups using the result of Liebeck and Shalev [LS04] to bound the Witten zeta function ζ_G for *simple* finite groups.

Proposition 8. *Let G be a D-quasirandom group. Let $f\colon G \to \mathbb{C}$ such that $\|f\|_\infty \le 1$, $\mathbb{E}[f] = 0$ and f_b is the mean zero component of the function $\Delta_b f$ (i.e., $f_b(x) = \Delta_b f(x) - \mathbb{E}_x[\Delta_b f(x)]$). Then*

$$\mathop{\mathbb{E}}_{g,b}\left[\left|\mathop{\mathbb{E}}_{x}\left[\Delta_b f(x) \cdot (f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})})(x)\right]\right|\right] \le \frac{1}{\sqrt{D}}\,.$$

Proof. Let us denote the expression on the L.H.S. as Γ. We use simple manipulations and previously stated facts to simplify the expression.

$$\begin{aligned}
\Gamma^2 &\le \mathop{\mathbb{E}}_{g,b}\left[\|\Delta_b f\|_2 \cdot \|(f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_2\right]^2 && \text{(By Cauchy-Schwarz inequality)} \\
&\le \mathop{\mathbb{E}}_{g,b}\left[\|f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_2\right]^2 && \text{(Since } \|\Delta_b f\|_2 \le 1\text{)} \\
&\le \mathop{\mathbb{E}}_{g,b}\left[\|f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_2^2\right] && \text{(By Cauchy Schwarz inequality)} \\
&= \mathop{\mathbb{E}}_{g,b}\left[\sum_{1\ne\rho\in\hat{G}} d_\rho \cdot \|\hat{f}_{g^{-1}bg}(\rho) \cdot \hat{\mu}_{g^{-1}C(g^{-1})}(\rho)\|_{\mathrm{HS}}^2\right] && \text{(By Parseval's identity \& } \hat{f}_{g^{-1}bg}(1) = 0\text{)} \\
&\le \mathop{\mathbb{E}}_{g,b}\left[\sum_{1\ne\rho\in\hat{G}} d_\rho \cdot \|\hat{f}_{gbg^{-1}}(\rho)\|_{\mathrm{HS}}^2 \cdot \|\hat{\mu}_{g^{-1}C(g^{-1})}(\rho)\|_{\mathrm{HS}}^2\right] && \text{(By submultiplicativity of norm)}
\end{aligned}$$

$$= \mathop{\mathbb{E}}_{g,b}\left[\sum_{1\neq\rho\in\hat{G}} \|\hat{f}_{g^{-1}bg}(\rho)\|_{\mathrm{HS}}^2 \cdot |\chi_\rho(g)|^2\right] \qquad \text{(By Claim 5)}$$
$$= \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{g}\left[|\chi_\rho(g)|^2 \cdot \mathop{\mathbb{E}}_{b}\left[\left\|\hat{f}_{gbg^{-1}}(\rho)\right\|_{\mathrm{HS}}^2\right]\right].$$

Now using the fact that gbg^{-1} is uniformly distributed in G for a fixed g and a uniformly random b in G, we can simply the above expression as follows.

$$\begin{aligned}
\Gamma^2 &\leq \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{g}\left[|\chi_\rho(g)|^2 \cdot \mathop{\mathbb{E}}_{b}\left[\left\|\hat{f}_b(\rho)\right\|_{\mathrm{HS}}^2\right]\right] \\
&= \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{b}\left[\left\|\hat{f}_b(\rho)\right\|_{\mathrm{HS}}^2\right] \cdot \mathop{\mathbb{E}}_{g}\left[|\chi_\rho(g)|^2\right] \\
&= \sum_{1\neq\rho\in\hat{G}} \mathop{\mathbb{E}}_{b}\left[\left\|\hat{f}_b(\rho)\right\|_{\mathrm{HS}}^2\right] && \text{(By orthogonality of } \chi_\rho\text{)} \\
&= \mathop{\mathbb{E}}_{b}\left[\sum_{1\neq\rho\in\hat{G}} \left\|\hat{f}_b(\rho)\right\|_{\mathrm{HS}}^2\right].
\end{aligned}$$

Finally, we use the fact that all the terms in the summation are non-negative and the group G is a D-quasirandom group.

$$\begin{aligned}
\Gamma^2 &\leq \frac{1}{D} \cdot \mathop{\mathbb{E}}_{b}\left[\sum_{1\neq\rho\in\hat{G}} d_\rho \cdot \left\|\hat{f}_b(\rho)\right\|_{\mathrm{HS}}^2\right] \\
&= \frac{1}{D} \cdot \mathop{\mathbb{E}}_{b}\left[\|f_b\|_2^2\right] && \text{(By Parseval's identity)} \\
&\leq \frac{1}{D}, && \text{(Because } \|f_b\|_2^2 \leq 1\text{).}
\end{aligned}$$

The proof of this lemma is similar to the proof of the BNP inequality (Lemma 6). The key difference being that we have a complete characterization of the Fourier transform of $\mu_{gC(g)}$ from Claim 5 which we use to give a sharper bound. □

We are now ready to prove the main Theorem 1. This part of the proof is similar to the corresponding expression that appears in the paper of Peluse [Pel18], which is in turn inspired by Tao's adaptation of Gowers' repeated Cauchy-Schwarzing trick to the nonebelian setting. We, however, present the entire proof for the sake of completeness.

Proof of Theorem 1. Let us denote the L.H.S. of the expression by Θ_{f_1,f_2,f_3}. Without loss of generality we assume $\mathbb{E}[f_3] = 0$. Now we have,

$$\begin{aligned}
\Theta_{f_1,f_2,f_3}^4 &= \left|\mathop{\mathbb{E}}_{x,y}\left[f_1(x)f_2(xy)f_3(xy^2)\right]\right|^4 \\
&= \left|\mathop{\mathbb{E}}_{x,z}\left[f_1(xz^{-1})f_2(x)f_3(xz)\right]\right|^4 && \text{(Change of variables: } x \leftarrow xy, z \leftarrow y\text{)}
\end{aligned}$$

$$
\begin{aligned}
&\leq \Big| \mathop{\mathbb{E}}_{x,z_1,z_2} \left[f_1(xz_1^{-1})f_1(xz_2^{-1})f_3(xz_1)f_3(xz_2)\right]\Big|^2 \\
&\qquad\qquad\qquad\qquad\qquad \text{(Cauchy-Schwarz over } x;\ \|f_2\|_\infty = 1 \text{ and expansion)} \\
&= \Big| \mathop{\mathbb{E}}_{y,z,a} \left[f_1(y)f_1(ya)f_3(yz^2)f_3(yza^{-1}z)\right]\Big|^2 \\
&\qquad\qquad\qquad\qquad\qquad \text{(Change of variables: } y \leftarrow xz_1^{-1}, z \leftarrow z_1, a \leftarrow z_1z_2^{-1}) \\
&= \Big| \mathop{\mathbb{E}}_{y,z,a} \left[\Delta_a f_1(y) \cdot \Delta_{z^{-1}a^{-1}z}\ f_3(yz^2)\right]\Big|^2 \\
&\leq \Big| \mathop{\mathbb{E}}_{y,a,z_1,z_2} \left[\Delta_{z_1^{-1}a^{-1}z_1}\ f_3(yz_1^2) \cdot \Delta_{z_2^{-1}a^{-1}z_2}\ f_3(yz_2^2)\right]\Big|, \\
&\qquad\qquad\qquad\qquad\qquad \text{(Cauchy-Schwarz over } y, a;\ \|f_1\|_\infty \leq 1 \text{).}
\end{aligned}
$$

Now, using the following change of variables, $z \leftarrow z_1,\ x \leftarrow yz_1^2,\ b \leftarrow z_1^{-1}a^{-1}z_1,\ g \leftarrow z_1^{-1}z_2$, we get

$$
\begin{aligned}
\Theta^4_{f_1,f_2,f_3} &\leq \Big| \mathop{\mathbb{E}}_{x,b,z,g} \left[\Delta_b\ f_3(x) \cdot \Delta_{g^{-1}bg}\ f_3(xz^{-1}gzg)\right]\Big| \\
&= \Big| \mathop{\mathbb{E}}_{x,b,g} \Big[\Delta_b\ f_3(x) \cdot \mathop{\mathbb{E}}_{z}[\Delta_{g^{-1}bg}\ f_3(xz^{-1}gzg)]\Big]\Big| \\
&= \Big| \mathop{\mathbb{E}}_{x,b,g} \Big[\Delta_b\ f_3(x) \cdot \mathop{\mathbb{E}}_{a}[\Delta_{g^{-1}bg}\ f_3(xa^{-1}) \cdot \frac{|G|}{|C(g^{-1})|} 1_{g^{-1}C(g^{-1})}(a)]\Big]\Big| \\
&= \Big| \mathop{\mathbb{E}}_{x,b,g} \Big[\Delta_b\ f_3(x) \cdot \mathop{\mathbb{E}}_{a}[\Delta_{g^{-1}bg}\ f_3(xa^{-1}) \cdot \mu_{g^{-1}C(g^{-1})}(a)]\Big]\Big| \\
&= \Big| \mathop{\mathbb{E}}_{x,b,g} \Big[\Delta_b\ f_3(x) \cdot \Delta_{g^{-1}bg}\ f_3 * \mu_{g^{-1}C(g^{-1})}(x)\Big]\Big|.
\end{aligned}
$$

We now separate the function $\Delta_{g^{-1}bg}\ f_3$ from its the mean zero part as follows: Let $\Delta_{g^{-1}bg}\ f_3 = f'_{g^{-1}bg} + f_{g^{-1}bg}$ where $f'_{g^{-1}bg} = \mathbb{E}_x[\Delta_{g^{-1}bg}\ f_3(x)]$ and $f_{g^{-1}bg}(x) = \Delta_{g^{-1}bg}\ f_3(x) - f'_{g^{-1}bg}$.

$$
\begin{aligned}
\Theta^4_{f_1,f_2,f_3} &\leq \Big| \mathop{\mathbb{E}}_{x,b,g} \Big[\Delta_b\ f_3(x) \cdot (f_{g^{-1}bg} + f'_{g^{-1}bg}) * \mu_{g^{-1}C(g^{-1})}(x)\Big]\Big| \\
&\leq \mathop{\mathbb{E}}_{b,g} \Big[\Big| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x) \cdot f_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}(x)\right]\Big|\Big] \\
&\qquad\qquad + \mathop{\mathbb{E}}_{b,g} \Big[\Big| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x) \cdot f'_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}(x)\right]\Big|\Big] \\
&\leq \frac{1}{\sqrt{D}} + \mathop{\mathbb{E}}_{b,g} \Big[\Big| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x)\right]\Big| \cdot \|f'_{g^{-1}bg} * \mu_{g^{-1}C(g^{-1})}\|_\infty\Big] \\
&\qquad\qquad\qquad\qquad\qquad \text{(Using Proposition 8 to bound the first expectation)} \\
&= \frac{1}{\sqrt{D}} + \mathop{\mathbb{E}}_{b,g} \Big[\Big| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x)\right]\Big| \cdot |f'_{g^{-1}bg}|\Big] \\
&\leq \frac{1}{\sqrt{D}} + \mathop{\mathbb{E}}_{b} \Big[\Big| \mathop{\mathbb{E}}_{x} \left[\Delta_b\ f_3(x)\right]\Big|\Big] \qquad\qquad \text{(Using } |f'_{g^{-1}bg}| \leq 1) \\
&\leq \frac{2}{\sqrt{D}}, \qquad\qquad \text{(By Corollary 7 and } \|f_3\|_\infty \leq 1).
\end{aligned}
$$

□

References

[BG08] JEAN BOURGAIN and ALEX GAMBURD. *Uniform expansion bounds for Cayley graphs of $SL_2(\mathbb{F}_p)$*. Ann. of Math., 167:625–642, 2008. 1

[BK21] AMEY BHANGALE and SUBHASH KHOT. *Optimal inapproximability of satisfiable k-LIN over non-abelian groups*. In *Proc. 53rd ACM Symp. on Theory of Computing (STOC)*, pages 1615–1628. 2021. arXiv:2009.02815, eccc:2020/TR20-130. 1

[BNP08] LÁSZLÓ BABAI, NIKOLAY NIKOLOV, and LÁSZLÓ PYBER. *Product growth and mixing in finite groups*. In *Proc. 19th Annual ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 248–257. 2008. 5

[BS85] LÁSZLÓ BABAI and VERA T. SÓS. *Sidon sets in groups and induced subgraphs of Cayley graphs*. Eur. J. Comb., 6(2):101–114, 1985. 1

[Dia98] PERSI DIACONIS. *Group representations in probability and statistics*, volume 11 of *IMS Lecture Notes Monogr. Ser.* Institute of Mathematical Statistics, 1998. 2

[DSV03] GIULIANA DAVIDOFF, PETER SARNAK, and ALAIN VALETTE. *Elementary Number Theory, Group Theory and Ramanujan Graphs*. London Mathematical Society Student Texts. Cambridge University Press, 2003. 1

[Gow08] WILLIAM TIMOTHY GOWERS. *Quasirandom groups*. Comb. Probab. Comput., 17(3):363–387, 2008. arXiv:0710.3877. 1

[LPS88] ALEXANDER LUBOTZKY, RALPH PHILLIPS, and PETER SARNAK. *Ramanujan graphs*. Combinatorica, 8(3):261–277, 1988. 1

[LS04] MARTIN W LIEBECK and ANER SHALEV. *Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks*. Journal of Algebra, 276(2):552–601, 2004. 2, 6

[Pel18] SARAH PELUSE. *Mixing for three-term progressions in finite simple groups*. Math. Proc. Cambridge Philos. Soc., 165(2):279–286, 2018. arXiv:1612.07385. 2, 4, 6, 7

[Tao13] TERRENCE TAO. *Mixing for progressions in nonabelian groups*. Forum of Mathematics, Sigma, 1:e2, 2013. arXiv:1212.2586. 2

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il