RANDOM WALKS ON DIRECT PRODUCTS OF GROUPS

ALIREZA SALEHI GOLSEFIDY AND SRIVATSA SRINIVAS

Abstract. Suppose $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$ is generated by a symmetric set S of cardinality n where p is a prime number. Suppose the Cheeger constants of the Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$ with respect to $\pi_L(S)$ and $\pi_R(S)$ are at least c_0, where π_L and π_R are projections to the left and the right components of $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$, respectively. Then the Cheeger constant of the Cayley graph of $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$ with respect to S is at least c where c is a positive number which only depends on n and c_0. This gives an affirmative answer to a question of Lindenstrauss and Varjú.

1. Introduction and statement of main results

1.1. **Background and main results.** Let $X^{(i)}$, $i \in \mathbb{N}$, be independent identically distributed (i.i.d.) random variables into a finite group G, with distribution μ. An *l-step random walk on G with distribution μ* is the random variable $X_\ell = X^{(\ell)}X^{(\ell-1)}\cdots X^{(1)}$. Given independent random variables X, Y into G with distribution μ and ν, respectively, we see that the distribution of XY is the convolution

$$\mu * \nu(y) := \sum_{x\in G} \mu(x)\nu(x^{-1}y)$$

of μ and ν. Similarly $f * g$ can be defined for any $f, g \in L^2(G)$. Now the distribution of X_ℓ is

$$\mu^{*(\ell)} := \underbrace{\mu * \cdots * \mu}_{\ell}.$$

We say that a measure μ on G is *symmetric* if for every $x \in G$ we have that $\mu(x) = \mu(x^{-1})$. We note that μ induces an operator $T_\mu : L^2(G) \to L^2(G)$ given by

$$T_\mu(f) = \mu * f.$$

We also note that G acts on $L^2(G)$ in the following manner

$$(x \cdot f)(x') = f(x^{-1} \cdot x'),$$

and so, $L^2(G)$ is a $\mathbb{C}[G]$-module. Note that we can identify $L^2(G)$ with $\mathbb{C}[G]$ by sending f to $\sum_{x\in G} f(x)x$, and via this identification $f * g$ is sent to fg, the product of f and g in $\mathbb{C}[G]$. Let $L^2(G)^\circ$ be the space orthogonal to the

A.S.G. acknowledges support by the NSF grants 1602137, 1902090, 2302519.

constants in $L^2(G)$. Note that $L^2(G)^\circ$ is a $\mathbb{C}[G]$-submodule of $L^2(G)$. Hence T_μ sends $L^2(G)^\circ$ to itself. We note that for every $f \in L^2(G)$ we have

$$\|T_\mu(f)\|_2 = \Big\| \sum_{x\in G} \mu(x) x\cdot f \Big\|_2 \le \sum_{x\in G} \mu(x)\|x\cdot f\|_2 = \|f\|_2.$$

We define the *spectral gap* of μ to be

$$\lambda(\mu) := \|T_\mu|_{L^2(G)^\circ}\|_{\mathrm{op}},$$

and inspired by the definition of Lyapunov exponent, we let

$$\mathcal{L}(\mu) = -\log \lambda(\mu),$$

where log denotes the base 2 logarithm.

Notice that the chosen name, spectral gap, might be a bit misleading. Since μ is a symmetric measure, T_μ is a self-adjoint averaging operator; so the spectrum of T_μ consists of real numbers in $[-1,1]$. The absolute value of eigenvalues of T_μ give us $|G|$ numbers (with multiplicity) in the interval $[0,1]$. The second largest number in this list is $\lambda(\mu)$, and the gap between $\lambda(\mu)$ and 1 is what we would like to control; clearly, $\mathcal{L}(\mu)$ gives us a way to measure this gap. In the literature, it is said that a family $\{\mu_i\}_i$ of probability measures has *the spectral gap property* if $\sup_i \lambda(\mu_i) < 1$, and so despite this caveat, we still use $\lambda(\mu)$ to denote the spectral gap of μ.

Note that $\mathcal{L}(\mu) > 0$ implies that the support of μ generates G. If X is a random variable with values in G and distribution μ, we let $\lambda(X) := \lambda(\mu)$ and $\mathcal{L}(X) := \mathcal{L}(\mu)$.

The spectral gap $\lambda(\mu)$ gives us a measurement of how fast the random walk is getting equidistributed in G (at least in L^2-norm). To formulate this, for every finite set A, we let μ_A be the probability counting measure on A. Then because, for every probability measure ν on G, the orthogonal projection of ν onto the constants $\mathbb{C}\mu_G \subset L^2(G)$ is μ_G, we have that

$$\begin{aligned}\|\mu^{*(\ell)} - \mu_G\|_2^2 = \|\mu^{*(\ell)} * (\mu_{\{1\}} - \mu_G)\|_2^2 &= \|T_\mu^\ell(\mu_{\{1\}} - \mu_G)\|_2^2 \\ &\le \lambda(\mu)^{2\ell}\|\mu_{\{1\}} - \mu_G\|_2^2 = 2^{-2\mathcal{L}(\mu)\ell}\|\mu_{\{1\}} - \mu_G\|_2^2.\end{aligned}$$

Because of this type of control on convergence to equidistribution, we are interested in finding a lower bound independent of $|G|$ for $\mathcal{L}(\mu)$.

In this work, we investigate random walks in $G\times G$. In a forthcoming joint work of the first author with Mallahi-Karai and Mohammadi [10], a result in the following framework is proved: if two compact groups G_1 and G_2 are *drastically different groups*, then for a probability measure μ on $G_1 \times G_2$ we have $\mathcal{L}(\mu) > 0$ if $\mathcal{L}(\pi_L[\mu]) > 0$ and $\mathcal{L}(\pi_R[\mu]) > 0$, where $\pi_L : G_1 \times G_2 \to G_1$ and $\pi_R : G_1 \times G_2 \to G_2$ are projection maps. One cannot expect a similar result when G_1 and G_2 have a non-trivial common (topological) quotient. For instance, consider the case $G_1 = G_2 = G$ and let μ be the probability Haar measure of the diagonally embedded $\Delta(G)$ of G in $G \times G$. Then clearly $\mathcal{L}(\pi_L[\mu])$ and $\mathcal{L}(\pi_R[\mu])$ are positive, but $\mathcal{L}(\mu) = 0$. In this example, however, the support of μ does not generate (a dense subgroup) of $G \times G$.

What if we add this extra algebraic condition? This subtlety is highlighted by Lindenstrauss and Varjú in [16, Open problem 1.4] in the form of the following question:

Question 1 (Lindenstruass-Varjú)**.** *Suppose S is a symmetric generating set of $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$. Is it possible to estimate the spectral gap of μ_S in terms of the spectral gaps of the projections to the direct factors and $|S|$?*

In this article we give an affirmative answer to this question.

Theorem 1. *Let μ be a symmetric measure on $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$ such that $\mathcal{L}(\pi_L[\mu]), \mathcal{L}(\pi_R[\mu]) \geq c_0 > 0$ and the minimum of μ on its support is α_0. If the support of μ generates $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$ and $p \gg_{c_0,\alpha_0} 1$, then $\mathcal{L}(\mu) \gg_{c_0,\alpha_0} 1$, where the implied constant is a positive number which only depends on c_0 and α_0.*

In our forthcoming article, we use the results of this work together with modular representation theory of $\mathrm{SL}_2(\mathbb{F}_q)$, Bourgain-Katz-Tao's sum-product result on finite fields, an investigation of certain unipotent group schemes, Gowers's theory of quasi-random groups, and Bourgain-Gamburd's method of gaining entropy to study random-walks on group extensions. In particular, we extend Theorem 1, and show that if $\mathbb{G}$ is a connected simply-connected perfect $\mathbb{Q}$-algebraic group and F is a finite field with large enough characteristic, then the spectral gap of a random-walk on $\mathbb{G}(F)$ can be bounded by the spectral gap of the projection of the random-walk on simple quotients of $\mathbb{G}(F)$. For the sake of clarity, highlighting our entropy inequality, and separating more algebraic tools, we have decided to have this special case as a separate article.

1.2. **Applications.** Studying spectral gaps of random walks on compact groups is an extremely interesting subject with many applications. For instance, it has been used to give an *explicit construction of expander graphs*, or recently it has been used in *affine sieve* and *sieve in groups*. We refer the reader to the beautiful surveys by Lubotzky [18] and Kowalski [14] for details and many more applications. Here first we present a combinatorial interpretation of Theorem 1 as it is formulated in the abstract, and then give one application which is new in nature and it is a consequence of our main result.

1.2.1. *Cheeger constant and expander graphs.* One of the interesting applications of the study of spectral gaps of random walks on a family of finite groups is its connection with the explicit construction of *expander graphs*. For a graph with finitely many vertices $\mathcal{G} = (V, E)$, the *Cheeger constant* of $\mathcal{G}$ is defined to be

$$e(\mathcal{G}) = \inf_{A \subset V, |A| < |V|/2} |\partial(A)|/|A|$$

where $\partial(A)$ is the set of vertices that are not in A but connected to a vertex in A via an edge. This constant measures how connected the graph $\mathcal{G}$ is.

A family $(\mathcal{G}_i)_{i\in\mathcal{I}}$ of finite graphs is called an *expander family* if there exist positive numbers k and c such that for all $i \in \mathcal{I}$, the maximum degree of a vertex in $\mathcal{G}_i$ is at most k and $e(\mathcal{G}_i) > c$. Expander families have an interesting history and have found applications in various areas of computer science and mathematics (see [12] and [15]). An expander family gives us a family of *sparse* (not many edges attached to a vertex) yet *highly connected* (expansion constant bounded below) graphs.

A result of Dodziuk [5] and Alon [1] (see [17, Proposition 4.2.4]) gives an isoperimetric inequality for regular graphs. In particular, this result implies that if $\mathcal{G}$ is the Cayley graph of a finite group with respect to a symmetric generating set S, then $\mathcal{L}(\mu_S)$ has a positive lower bound in terms of $|S|$ and the Cheeger constant of $\mathcal{G}$, and conversely the Cheeger constant of $\mathcal{G}$ has a positive lower bound in terms of $\mathcal{L}(\mu_S)$ and $|S|$. Hence Theorem 1 implies the following.

Theorem 2. *Suppose S is a symmetric generating set of $\mathrm{SL}_2(\mathbb{F}_p)\times\mathrm{SL}_2(\mathbb{F}_p)$ where p is a prime number. Let π_L and π_R be the projections to the left and the right components of $\mathrm{SL}_2(\mathbb{F}_p)\times\mathrm{SL}_2(\mathbb{F}_p)$, respectively. Suppose $|\pi_L(S)| = |\pi_R(S)| = |S|$. Let $\mathcal{G}_L$ and $\mathcal{G}_R$ be the Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$ with respect to $\pi_L(S)$ and $\pi_R(S)$, respectively. Suppose the Cheeger constant $e(\mathcal{G}_L)$ and $e(\mathcal{G}_R)$ are at least c_0. Then the Cheeger constant of the Cayley grpah of $\mathrm{SL}_2(\mathbb{F}_p)\times\mathrm{SL}_2(\mathbb{F}_p)$ with respect to S is at least a positive number which depends only on $|S|$ and c_0.*

Let's remark that the technical condition on the cardinality of $\pi_L(S)$ and $\pi_R(S)$ is not needed. The authors in their forthcoming article on random walks on perfect groups will show that if μ and μ' are two probability measures on a finite group that have the same support and the minimum of μ and μ' on their support is at least α_0, then $\mathcal{L}(\mu) \ll_{\alpha_0} \mathcal{L}(\mu') \ll_{\alpha_0} \mathcal{L}(\mu)$ (this is probably well-known to experts). This result allows us to remove the mentioned technical condition.

1.2.2. *Representation and character varieties.* Let $F_2 = \langle a, b\rangle$ be the free group freely generated by a and b. The SL_2-representation variety of F_2 is given by the functor

$$\mathrm{Rep}_2(A) := \mathrm{Hom}(F_2, \mathrm{SL}_2(A))$$

from the category of unital commutative rings to the category of sets. It is clear that $\mathrm{Rep}_2(A)$ can be identified with two copies of $\mathrm{SL}_2(A)$. The group SL_2 acts on Rep_2 by conjugation. The geometric quotient of Rep_2 by SL_2 is a variety and by works of Fricke it has an integral model which is denoted by Ch_2. In particular, for every $\rho \in \mathrm{Rep}_2(A)$, we get a point $[\rho] \in \mathrm{Ch}_2(A)$, and we have $[\rho_1] = [\rho_2]$ for $\rho_1, \rho_2 \in \mathrm{Rep}_2(A)$ if there is $x \in \mathrm{SL}_2(A)$ such that $\rho_2(y) = x\rho_1(y)x^{-1}$ for all $y \in F_2$ (notice that the converse of this statement is not correct, but that is not related to our application). For every element $w \in F_2$ and $[\rho] \in \mathrm{Ch}_2(A)$, we let

$$t_w([\rho]) := \mathrm{Tr}(\rho(w)).$$

By works of Fricke, we can view t_w as a regular function on Ch_2. For every positive δ, we let

$$\mathrm{Rep}_2(\mathbb{F}_p)_\delta := \{\rho \in \mathrm{Rep}_2(\mathbb{F}_p) \mid \mathcal{L}(\mu_{\{\rho(a)^{\pm 1},\rho(b)^{\pm 1}\}}) \geq \delta\}.$$

Note that $\bigcup_{\delta>0} \mathrm{Rep}_2(\mathbb{F}_p)_\delta = \{\rho \in \mathrm{Rep}_2(\mathbb{F}_p) \mid \rho(F_2) = \mathrm{SL}_2(\mathbb{F}_p)\}$. We can show that many t_w's can distinguish two distinct points $[\rho_1]$ and $[\rho_2]$ of $\mathrm{Ch}_2(\mathbb{F}_p)$ if $\rho_1, \rho_2 \in \mathrm{Rep}_2(\mathbb{F}_p)_\delta$.

Corollary 3. *Suppose δ is a positive number and $\rho_1, \rho_2 \in \mathrm{Rep}_2(\mathbb{F}_p)_\delta$. Suppose a, b freely generate a free group F_2. Then there is $0 < \beta := \beta(\delta) < 1$ such that for every positive integer ℓ the following statements hold.*

(1) If $\rho_1 \neq \rho_2$, then

$$\mu^{*(\ell)}_{\{a^{\pm 1},b^{\pm 1}\}}(\{w \in F_2 \mid \rho_1(w) = \rho_2(w)\}) \leq p^{-1} + \beta^\ell |\mathrm{SL}_2(\mathbb{F}_p)|.$$

(2) If $[\rho_1] \neq [\rho_2]$, then either

$$\mu^{*(\ell)}_{\{a^{\pm 1},b^{\pm 1}\}}(\{w \in F_2 \mid t_w([\rho_1]) = t_w([\rho_2])\}) \leq 5p^{-1} + \beta^\ell |\mathrm{SL}_2(\mathbb{F}_p)|,$$

or there is an automorphism $\widehat{\phi}$ of $\mathrm{SL}_2(\mathbb{F}_p)$ such that for every $w \in F_2$, $\rho_2(w) = \pm\widehat{\phi}(\rho_1(w))$. In the latter case, $t_w([\rho_1]) = \pm t_w([\rho_2])$ for every $w \in F_2$.

It is not known whether there is a positive number δ_0 such that

$$\mathrm{Rep}_2(\mathbb{F}_p)_{\delta_0} = \{\rho \in \mathrm{Rep}_2(\mathbb{F}_p) \mid \rho(F_2) = \mathrm{SL}_2(\mathbb{F}_p)\}.$$

So we do not know if Corollary 3 holds for a fixed universal constant β and every surjective $\rho_1, \rho_2 \in \mathrm{Rep}_2(\mathbb{F}_p)$.

1.3. Proof strategy.

Note that since the natural quotient map

$$\bar{\iota} : \mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p) \to \mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$$

has a kernel of cardinality at most 4, it is enough to prove a similar result as in Theorem 1 for $\mathrm{PSL}_2(\mathbb{F}_p)$. Let us recall that for a random variable X with finite support and probability law μ, the *Renyi Entropy* of X is

$$H_2(X) := -\log \|\mu\|_2^2.$$

Let $X = (X_L, X_R)$ be a random variable into $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$ with distribution μ. We assume that μ is symmetric. Suppose that

$$\mathcal{L}(\pi_R[\mu]), \mathcal{L}(\pi_L[\mu]) > c_0 > 0,$$

μ takes a minimum of α_0 on its support, and that the support of μ generates $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$. Let L_1 be a large natural number. Using the Bourgain-Gamburd method (see [3]), we deduce that the only objection to μ having a spectral gap depending on c_0 and L_1 is if there exist an automorphism ϕ of $\mathrm{PSL}_2(\mathbb{F}_p)$ and an integer $\ell \geq L_1 \log |\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)|$, such that

$$\mathbb{P}(X_\ell \in \Gamma_\phi) \geq |\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)|^{-1/26}, \tag{1}$$

where $\Gamma_\phi = \{(x, \phi(x)) | x \in \mathrm{PSL}_2(\mathbb{F}_p)\}$ is the graph of ϕ. Considering the group of outer automorphisms of $\mathrm{PSL}_2(\mathbb{F}_p)$ has only two elements, we reduce the general case to the case where ϕ is inner, say that $\phi(x) = zxz^{-1}$. Notice $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$ acts on $\mathrm{PSL}_2(\mathbb{F}_p)$ by the left and right multiplications; that means $(u, v) \cdot x = uxv^{-1}$. The graph Γ_ϕ is the stabilizer subgroup of $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$ with respect to z^{-1}; that means

$$\Gamma_\phi = \{(u, v) \in \mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p) \mid (u, v) \cdot z^{-1} = z^{-1}\}.$$

Thus equation (1) turns in to

$$\mathbb{P}(X_\ell \cdot z^{-1} = z^{-1}) \geq |\,\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)|^{-1/26} \tag{2}$$

We would like to show that (2) is not possible. Equation (2) suggests that the Rényi entropy of $X_\ell \cdot z^{-1}$ should be *small*. This brings us to studying $H_2(X_\ell \cdot Y)$ where Y is a random variable on $\mathrm{PSL}_2(\mathbb{F}_p)$ that has *small* Renyi entropy. We note that $X_\ell \cdot Y = (X_L)_\ell Y (X_R)_\ell^{-1}$, and we know that $(X_L)_\ell$ and $(X_R)_\ell$ are almost equidistributed for $\ell \gg_{c_0} \log |\,\mathrm{PSL}_2(\mathbb{F}_p)|$. But we do not know how $(X_L)_\ell$ and $(X_R)_\ell$ are correlated. All we know is that the range of X generates $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$. The following lemma is instrumental in resolving this issue.

Lemma 4. *Suppose G and H are two finite groups, and G acts on H. Let $X^{(1)}, X^{(2)}$ be i.i.d. random variables into a finite group G and $Y^{(1)}, Y^{(2)}$ be i.i.d. random variables into a finite group H. Then we have that*

$$H_2((X^{(1)} \cdot Y^{(1)})^{-1}(X^{(2)} \cdot Y^{(2)})) \geq H_2((X^{(1)} \cdot Y^{(1)})^{-1}(X^{(1)} \cdot Y^{(2)}))$$

Lemma 4 is aligned with the general principle which says that *reducing the degree of freedom should decrease the Rényi entropy.* Applying Lemma 4 to our random variable X, we get

$$\begin{aligned} H_2((X^{(1)} \cdot Y^{(1)})^{-1}(X^{(2)} . Y^{(2)})) \geq & H_2((X^{(1)} \cdot Y^{(1)})^{-1}(X^{(1)} \cdot Y^{(2)})) \\ = & H_2((X_L^{(1)} Y^{(1)} (X_R^{(1)})^{-1})^{-1} (X_L^{(1)} Y^{(2)} (X_R^{(1)})^{-1}) \\ = & H_2(X_R^{(1)} (Y^{(1)})^{-1} Y^{(2)} (X_R^{(1)})^{-1}) \end{aligned} \tag{3}$$

Based on (3), we get a lower bound for $H_2((X^{(1)} \cdot Y^{(1)})^{-1}(X^{(2)} \cdot Y^{(2)}))$ using conjugation by X_R. Since we have a control on the spectral gap of X_R, after $\ell_0 := O_{c_0}(1)$ steps random walk $(X_R)_{\ell_0}$ gets close to equidistribution. This implies that conjugation by $(X_R)_{\ell_0}$ spreads the weight almost equally in the conjugacy classes that intersect the range of $(Y^{(1)})^{-1} Y^{(2)}$. Considering every conjugacy class of $\mathrm{PSL}_2(\mathbb{F}_p)$ except $\{1\}$ has at least p elements and Y has *small* Renyi entropy, we obtain the following dichotomy:

Either Y is *almost* concentrated at one point or we *gain* Rényi entropy after conjugation by X_R.

By gaining Rényi entropy, we mean that for some $\varepsilon := \varepsilon(\alpha_0) > 0$ and $\ell_0 := O_{c_0,\alpha_0}(1)$ the following holds

$$H_2((X_R)_{\ell_0} (Y^{(1)})^{-1} Y^{(2)} (X_R)_{\ell_0}^{-1}) \geq H_2((Y^{(1)})^{-1} Y^{(2)}) + \varepsilon.$$

If Y is *almost* concentrated at one point, we use the assumption that the range of X generates $\mathrm{PSL}_2(\mathbb{F}_p)\times\mathrm{PSL}_2(\mathbb{F}_p)$ to show the existence of a positive number $\varepsilon_1 := \varepsilon_1(\alpha_0)$ such that

$$H_2(X\cdot Y) \geq H_2(Y) + \varepsilon_1.$$

Altogether we obtain the following lemma.

Lemma 5. *Let $G := \mathrm{PSL}_2(\mathbb{F}_p)\times\mathrm{PSL}_2(\mathbb{F}_p)$. Suppose $X := (X_L, X_R)$ is a random variable with values in G and probability law μ. Suppose $\alpha_0 := \min\{\mathbb{P}(X = x) \mid x \in \operatorname{supp}(\mu)\}$ and $\mathcal{L}(X_R) \geq c_0 > 0$. Suppose that the range of X generates G. Then there exist constants $L, C \gg_{c_0,\alpha_0} 1$ such that for every random variable Y on $\mathrm{PSL}_2(\mathbb{F}_p)$ and every $\ell \geq L \log|G|$*

$$H_2(X_\ell \cdot Y) \geq \frac{1}{12}\log|\mathrm{PSL}_2(\mathbb{F}_p)\times\mathrm{PSL}_2(\mathbb{F}_p)| - C, \tag{4}$$

where $X\cdot Y := X_L Y X_R^{-1}$ and X_ℓ is the ℓ-step random walk with respect to μ.

For large enough p, (4) implies

$$H_2(X_\ell \cdot Y) \geq \frac{1}{12.5}\log|\mathrm{PSL}_2(\mathbb{F}_p)\times\mathrm{PSL}_2(\mathbb{F}_p)|.$$

Therefore for every $\ell \geq L\log|\mathrm{PSL}_2(\mathbb{F}_p)\times\mathrm{PSL}_2(\mathbb{F}_p)|$, we have

$$\mathbb{P}(X_\ell\cdot z^{-1} = z^{-1}) \leq e^{-\frac{1}{2}H_2(X_\ell\cdot z^{-1})} \leq |\mathrm{PSL}_2(\mathbb{F}_p)\times\mathrm{PSL}_2(\mathbb{F}_p)|^{-1/25}.$$

Thus Lemma 5 gives us that Equation (2) cannot hold, and this completes the proof.

1.4. **Acknowledgment.** The authors would like to thank the anonymous referee(s) for their helpful comments.

2. Notation and preliminary results

2.1. **Conventions.** If $f : G \to V$ is a function from a finite group to a $\mathbb{C}$-vector space and μ is a measure on G, we define

$$\int_G f(x)d\mu(x) := \sum_{x\in G}\mu(x)f(x)$$

We endow $L^2(G)$ with the inner product

$$\langle f, g\rangle = \sum_{x\in G} f(x)\overline{g(x)}$$

where $f, g \in L^2(G)$. For $f \in L^2(G)$, $\check{f} \in L^2(G)$ is given by

$$\check{f}(x) = \overline{f(x^{-1})}.$$

Note that if X is a random variable with values in G and distribution μ, then the probability law of X^{-1} is $\check{\mu}$.

For a subset A of a finite group G and a positive integer k, we let

$$\textstyle\prod_k A := \{a_1\cdots a_k \mid a_1,\ldots,a_k \in A\}.$$

2.2. **Basics of Fourier analysis for finite groups and quasi-random groups.** Suppose G is a finite group. Then $\widehat{G}$ denotes the set of irreducible unitary subrepresentations of the regular representation $L^2(G)$. For $f \in L^2(G)$, the Fourier inverse of f is defined as

$$\widehat{f}(\pi) := \frac{1}{|G|} \sum_{g \in G} f(g)\pi(g)^*,$$

where $\pi(g)^*$ is the adjoint of $\pi(g)$. For $f_1, f_2 \in L^2(G)$, the convolution of f_1 and f_2 is defined with respect to the counting measure (and not the *probability* counting measure)

$$f_1 * f_2(x) = \sum_{x_1 x_2 = x} f_1(x_1) f_2(x_2),$$

and we have

$$\widehat{f_1 * f_2}(\pi) = |G| \widehat{f_2}(\pi)\widehat{f_1}(\pi) \tag{5}$$

for every $\pi \in \widehat{G}$. The Parseval theorem states that

$$\|f\|_2^2 = \sum_{\pi \in \widehat{G}} \dim \pi \; \|\widehat{f}(\pi)\|_{\mathrm{HS}}^2, \tag{6}$$

where $\|f\|_2^2 := \frac{1}{|G|} \sum_{g \in G} |f(g)|^2$ and the Hilbert-Schmidt norm $\|T\|_{\mathrm{HS}}^2 = \mathrm{Tr}(TT^*)$.

Let's recall Gowers's notion of quasi-randomness and its consequences to the study of random-walks(see [11]).[1]

Definition. For a positive number c, we say a finite group G is *c-quasi-random* if $\dim \pi \geq |G|^c$ for every non-trivial $\pi \in \widehat{G}$.

The following mixing inequality is one of the main properties of a quasi-random group (see [2, Theorem 2.1], [11], and [9, Lemma 6.1]).

Lemma 6. *Suppose c is a positive number and G is a c-quasi-random group. Then for every $f_1 \in L^2(G)$ and $f_2 \in L^2(G)^\circ$ the following inequality holds,*

$$\|f_1 * f_2\|_2 \leq |G|^{(1-c)/2} \|f_1\|_2 \|f_2\|_2.$$

Proof. By the Parseval theorem (see (6)) and (5), we obtain that

$$\begin{aligned} \|f_1 * f_2\|_2^2 &= \sum_{\pi \in \widehat{G}} \dim \pi \; \|\widehat{f_1 * f_2}(\pi)\|_{\mathrm{HS}}^2 = |G|^2 \sum_{\pi \in \widehat{G}, \pi \neq 1} \dim \pi \; \|\widehat{f_2}(\pi)\widehat{f_1}(\pi)\|_{\mathrm{HS}}^2 \\ &\leq |G|^{2-c} \Big(\sum_{\pi \in \widehat{G}} \dim \pi \; \|\widehat{f_2}(\pi)\|_{\mathrm{HS}}^2 \Big) \Big(\sum_{\pi \in \widehat{G}} \dim \pi \; \|\widehat{f_1}(\pi)\|_{\mathrm{HS}}^2 \Big) \\ &\leq |G|^{2-c} \|f_1\|_2^2 \|f_2\|_2^2. \end{aligned}$$

Hence $\|f_1 * f_2\|_2^2 \leq |G|^{1-c} \|f_1\|_2^2 \|f_2\|_2^2$, and the claim follows. □

[1] It should be pointed out that in [19], Sarnak and Xue had implicitly used the concept of quasi-randomness in order to prove a spectral gap property.

As it was pointed out in the introduction, when X is a symmetric random-variable with law μ, $T_\mu : L^2(G) \to L^2(G), T_\mu(f) := f * \mu$ is a self-adjoint operator, and $\lambda(\mu)$ is the maximum of the absolute values of eigenvalues of $T_\mu|_{L^2(G)^\circ} : L^2(G)^\circ \to L^2(G)^\circ$. Hence one can see that

$$\mathcal{L}(\mu^{(\ell)}) = \ell \mathcal{L}(\mu)$$

for every positive integer ℓ.

Lemma 7. *Suppose c is a positive number and G is a c-quasi-random group. Suppose X is a symmetric random-variable with values in G. Then*

1. *$H_2(X) \geq (1 - \frac{c}{2}) \log |G|$ implies $\mathcal{L}(X) \geq \frac{c}{4} \log |G|$.*
2. *Suppose X_ℓ is the random variable after ℓ-step random walk with respect to X. Suppose $C > 0$, and $H_2(X_\ell) \geq (1 - \frac{c}{2}) \log |G|$ for a positive integer $\ell \leq C \log |G|$. Then $\mathcal{L}(X) \geq \frac{c}{4C}$.*

Proof. Suppose μ is the probability law of X; that means $\mu(x) := \mathbb{P}(X = x)$ for every $x \in G$. By Lemma 6, for every function $f \in L^2(G)^\circ$, the following holds,

$$\|\mu * f\|_2 \leq |G|^{(1-c)/2} \|\mu\|_2 \|f\|_2.$$

Hence

$$\mathcal{L}(\mu) \geq \frac{-1 + c}{2} \log |G| + \frac{1}{2} H_2(X) \geq \frac{c}{4} \log |G|.$$

This implies the first part. Next applying the first part for the random variable X_ℓ, we deduce

$$C \log |G| \mathcal{L}(\mu) \geq \ell \mathcal{L}(\mu) \geq \frac{c}{4} \log |G|,$$

and the second part follows. □

2.3. **Group action and convolution.** When a finite group G acts on a finite set H, we write $G \curvearrowright H$. An action of G on H induces an action of G on $L^2(H)$ by

$$(x \cdot f)(y) = f(x^{-1} \cdot y)$$

for $x \in G, y \in H, f \in L^2(H)$. This is a unitary action and this way we can view $L^2(H)$ as a $\mathbb{C}[G]$-module. Given $\mu \in L^2(G), f \in L^2(H)$, we define

$$\mu \boxtimes f = \sum_{x \in G} \mu(x)(x \cdot f) = \int_G (x \cdot f) d\mu(x).$$

Note that if X is a random variable with values in G and distribution μ, Y is a random variable with values in H and distribution η, and X and Y are independent, then the distribution of $X \cdot Y$ is $\mu \boxtimes \eta$. Notice that the group action properties give us the following relations. If $\mu, \nu \in L^2(G), f \in L^2(H)$ then

$$\mu \boxtimes (\nu \boxtimes f) = (\mu * \nu) \boxtimes f.$$

Moreover $\boxtimes : L^2(G) \times L^2(H) \to L^2(H)$, $(\mu, f) \mapsto \mu \boxtimes f$ is a bilinear map. We call $\boxtimes$ the *convolution associated to* $G \curvearrowright H$. Notice that for the counting probability measure μ_G on G,

$$\mu_G \boxtimes \cdot : L^2(H) \to L^2(H), \quad f \mapsto \mu_G \boxtimes f$$

is the orthogonal projection of $L^2(H)$ onto the space $L^2(H)^G$ of G-invariant functions in $L^2(H)$. Thus $f \mapsto f - \mu_G \boxtimes f$ is the orthogonal projection from $L^2(H)$ to the space $(L^2(H)^G)^\perp$ orthogonal to the space of G-invariant functions. We also observe that if $\mu_{\{1\}}$ is the point mass at the identity, then $f = \mu_{\{1\}} \boxtimes f$ for every $f \in L^2(H)$. Therefore the orthogonal projection from $L^2(H)$ to the space $(L^2(H)^G)^\perp$ is given by

$$f \mapsto (\mu_{\{1\}} - \mu_G) \boxtimes f. \tag{7}$$

Notice that every irreducible subrepresentation V of $(L^2(H)^G)^\perp$ is non-trivial, and so by Maschke's theorem [6, Theorem 4.1.1], there is a G-module isometric embedding $i : V \to L^2(G)^\circ$. Hence for every probability measure μ on G and $f \in V$, we have

$$\begin{aligned} \|\mu \boxtimes f\|_2 &= \|i_V(\mu \boxtimes f)\|_2 = \|\mu * i_V(f)\|_2 \\ &\le \lambda(\mu)\|i_V(f)\|_2 = \lambda(\mu)\|f\|_2. \end{aligned} \tag{8}$$

Considering $(L^2(H)^G)^\perp$ is a direct sum of pairwise orthogonal irreducible subrepresentations, by (8), we obtain that $\|\mu \boxtimes f\|_2 \le \lambda(\mu)\|f\|_2$ for every $f \in (L^2(H)^G)^\perp$. Combining this result with (7), we deduce that

$$\begin{aligned} \|(\mu - \mu_G) \boxtimes f\|_2 = \|(\mu * (\mu_{\{1\}} - \mu_G)) \boxtimes f\|_2 &= \|\mu \boxtimes ((\mu_{\{1\}} - \mu_G) \boxtimes f)\|_2 \\ \le \lambda(\mu)\|(\mu_{\{1\}} - \mu_G) \boxtimes f\|_2 &\le \lambda(\mu)\|f\|_2, \end{aligned} \tag{9}$$

for every $f \in L^2(H)$.

3. An inequality for the Renyi entropy of random variables

The aim of this section is to prove Lemma 4. Let's recall that in the setting of Lemma 4, we have two finite groups G and H, and G acts on H. There are i.i.d. random variables $X^{(1)}, X^{(2)}$ with values in G and distribution μ, and i.i.d. random variables $Y^{(1)}, Y^{(2)}$ with values in H and distribution η. In this section, we work with the (non-normalized) counting measure m_H on H.

We start with two sets of convolution identities. The first one is well-known and the second one is based on the fact that $*$ is bilinear.

Lemma 8. *In the above setting, for $f, g, h \in L^2(H)$, the following identities hold:*

(1) $f * (g * h) = (f * g) * h$, $\langle f * g, h\rangle = \langle f, h * \check{g}\rangle$ *and* $\langle f * g, h\rangle = \langle g, \check{f} * h\rangle$,
(2) $(\mu \boxtimes f) * (\mu \boxtimes g) = \int_{G^2} (u \cdot f) * (v \cdot g)\, d(\mu^{\otimes 2})(u, v)$.

Proof. Both parts easily follow from switching the order of summations. Here we only discuss the second part. The second part follows from the fact that $\mu \boxtimes f = \int_G (x \cdot f) d\mu(x)$ and $*$ is bilinear:

$$\begin{aligned}(\mu \boxtimes f) * (\mu \boxtimes g) =& \Big(\int_G (u \cdot f) d\mu(u) \Big) * \Big(\int_G (v \cdot g) d\mu(v) \Big) \\ =& \int_{G^2} (u \cdot f) * (v \cdot g) \; d\mu^{\otimes 2}(u, v).\end{aligned}$$

This completes the proof. □

We will be working with a new norm on $L^2(H)$ that we denote by $\vert\!\vert\!\vert \cdot \vert\!\vert\!\vert$ and it can be viewed as a non-commutative version of Gowers's U^2-norm. We will show that this norm is preserved by *shifted-automorphism* group actions $G \curvearrowright H$.

Definition. Suppose G and H are two groups. An action $G \curvearrowright H$ is called a *shifted-automorphism group action* if there are a group homomorphism $\phi : G \to \mathrm{Aut}(H)$ and a function $c : G \to H$ such that $x \cdot y = c(x)(\phi(x))(y)$ for every $x \in G$ and $y \in H$.

Notice that for every group H, the group action $H \times H \curvearrowright H$ given by $(x_L, x_R) \cdot y := x_L y x_R^{-1}$ is a shifted-automorphism group action as

$$x_L y x_R^{-1} = c(x_L, x_R) \phi(x_L, x_R)(y)$$

where $c(x_L, x_R) = x_L x_R^{-1}$ and $\phi : H \times H \to \mathrm{Aut}(H)$ is a group homomorphism given by

$$\phi(x_L, x_R)(y) := x_R y x_R^{-1}.$$

Lemma 9. *Suppose H is a finite group. Let $\vert\!\vert\!\vert f \vert\!\vert\!\vert := \|\check{f} * f\|_2^{1/2}$ for $f \in L^2(H)$. Then the following statements hold.*

1. *$\vert\!\vert\!\vert \cdot \vert\!\vert\!\vert$ is a norm and $\|f\|_2 \leq \vert\!\vert\!\vert f \vert\!\vert\!\vert$ for every non-negative $f \in L^2(H)$.*
2. *Suppose $G \curvearrowright H$ is a shifted-automorphism group action. Then for every $x \in G$ and $f \in L^2(H)$, we have $\vert\!\vert\!\vert x \cdot f \vert\!\vert\!\vert = \vert\!\vert\!\vert f \vert\!\vert\!\vert$.*
3. *Suppose $G \curvearrowright H$ is a shifted-automorphism group action and μ is a probability measure on G. Then we have that $\vert\!\vert\!\vert \mu \boxtimes f \vert\!\vert\!\vert \leq \vert\!\vert\!\vert f \vert\!\vert\!\vert$ for every $f \in L^2(H)$.*

Proof. (1) For every $g \in L^2(H)$, the convolution operator

$$T_g : L^2(H) \to L^2(H), \quad T_g(h) := g * h$$

is an integral operator with the kernel $K_g : H \times H \to \mathbb{C}$, $K_g(x, y) := g(xy^{-1})$. Therefore the Hilbert-Schmidt norm $\|T_g\|_{\mathrm{HS}}$ is equal to $\|K_g\|_2$ (see [4, Chapter II, Proposition 4.7]). Notice that

$$\|K_g\|_2^2 = \sum_{x,y \in H} |g(x^{-1}y)|^2 = |H| \|g\|_2^2,$$

and so $\|T_g\|_{\mathrm{HS}} = |H|^{1/2}\|g\|_2$. Hence

$$\|\!| f \|\!| = \|\check{f} * f\|_2^{1/2} = |H|^{-1/4}\|T_{\check{f}*f}\|_{\mathrm{HS}}^{1/2}. \tag{10}$$

By Lemma 8, we have $T_f^* = T_{\check{f}}$. Hence by (10), we obtain that

$$\|\!| f \|\!| = |H|^{-1/4}\|T_f^* \circ T_f\|_{\mathrm{HS}}^{1/2}. \tag{11}$$

For an operator $T : L^2(H) \to L^2(H)$, let $\|\!| T \|\!| := \|T^* \circ T\|_{\mathrm{HS}}^{1/2}$. Notice that if $\sigma_1, \ldots, \sigma_n$ are the singular values of T, then

$$\|\!| T \|\!| = (\sigma_1^4 + \cdots + \sigma_n^4)^{1/4}. \tag{12}$$

By (12) and [13, Theorem 7.4.24], we have that $\|\!| \cdot \|\!|$ is a unitarily invariant norm on $\mathrm{End}_{\mathbb{C}}(L^2(H))$. Therefore by (10) for every $f, g \in L^2(H)$, we have that

$$\|\!| f+g \|\!| = |H|^{-1/4}\|\!| T_f + T_g \|\!| \le |H|^{-1/4}\|\!| T_f \|\!| + |H|^{-1/4}\|\!| T_g \|\!| = \|\!| f \|\!| + \|\!| g \|\!|.$$

For every $c \in \mathbb{C}$ and $f \in L^2(H)$, clearly we have $\|\!| cf \|\!| = |c|\|\!| f \|\!|$, and $\|\!| f \|\!| = 0$ implies that $f = 0$. Hence $\|\!| \cdot \|\!|$ is a norm on $L^2(H)$. For two non-negative functions f and g, we have

$$\begin{aligned}\|f * g\|_2^2 = &\sum_x (f*g)(x)^2 = \sum_x \Big(\sum_{x_1x_2=x} f(x_1)g(x_2)\Big)^2 \\ \ge &\sum_x \sum_{x_1x_2=x} f(x_1)^2 g(x_2)^2 = \|f\|_2^2\|g\|_2^2,\end{aligned}$$

and so $\|f * g\|_2 \ge \|f\|_2\|g\|_2$. Therefore for a non-negative function f, we have

$$\|\!| f \|\!| = \|\check{f} * f\|_2^{1/2} \ge (\|\check{f}\|_2\|f\|_2)^{1/2} = \|f\|_2.$$

(2) Notice that $\{\mu_{\{y\}}\}_{y\in H}$ is an orthonormal basis of $L^2(H)$, and for y, y' in H and $f \in L^2(H)$, the (y, y')-matrix entry of T_f is $f(yy'^{-1})$. Hence the (y, y')-matrix entry of $T_{x\cdot f}$ is

$$f(x^{-1} \cdot (yy'^{-1})) = f\big(c(x^{-1})\ (\phi(x^{-1}))(y)\ (\phi(x^{-1}))(y')^{-1}\big), \tag{13}$$

where $\phi : G \to \mathrm{Aut}(H)$ and $c : G \to H$ give us the shifted-automorphism group action $G \curvearrowright H$. By (13), we obtain that the (y, y')-entry of $T_{x\cdot f}$ is equal to the $((\phi(x^{-1}))(y), (\phi(x^{-1}))(y'))$-matrix entry of $T_{c_x\cdot f}$ where $c_x := c(x^{-1})^{-1} \in H$ and $(c_x \cdot f)(y) = f(c_x^{-1}y)$ for every $y \in H$. For every $x \in G$, let $\sigma(x) : L^2(H) \to L^2(H)$ be the unitary operation given by $(\sigma(x)(f))(y) := f(\phi(x)^{-1}(y))$. Then by the above discussion, we deduce that

$$T_{x\cdot f} = \sigma(x) \circ T_{c_x\cdot f} \circ \sigma(x)^{-1}. \tag{14}$$

For every $y \in H$, let $l(y) : L^2(H) \to L^2(H), (l(y))(f) := y \cdot f$. Then $l(y)$ is a unitary map and $T_{y\cdot f} = l(y) \circ T_f$. Hence by (14), we obtain that

$$T_{x\cdot f} = \sigma(x) \circ l(c_x) \circ T_f \circ \sigma(x)^{-1}. \tag{15}$$

Therefore by [13, Theorem 7.4.24] and (15), we conclude that $\|T_{x\cdot f}\| = \|T_f\|$, and so

$$\|x \cdot f\| = |H|^{-1/4}\|T_{x\cdot f}\| = |H|^{-1/4}\|T_f\| = \|f\|.$$

(3) For every probability measure μ on G, by the first two parts we have:

$$\|\mu \boxtimes f\| \le \|\sum_{x\in G} \mu(x) x \cdot f\| \le \sum_{x\in G} \mu(x)\|x \cdot f\| = \|f\|.$$

This completes the proof. □

The following inequality plays an important role in proving Lemma 4.

Lemma 10. *We have that* $\|\mu \boxtimes f\|^2 \le \| \int_G \widecheck{(u \cdot f)} * (u \cdot f)\, d\mu(u)\|_2$.

Proof. Observe that $\mu \boxtimes f = \int_G (x \cdot f) d\mu(x)$ and $\widecheck{\mu \boxtimes f} = \int_G \widecheck{x \cdot f} d\mu(x)$. Therefore similar to the second part of Lemma 8, we have

$$(\mu \boxtimes f) * (\widecheck{\mu \boxtimes f}) = \int_{G^2} (u \cdot f) * (\widecheck{v \cdot f})\, d\mu^{\otimes 2}(u, v). \tag{16}$$

By (16) and the semi-linearity of the dot product, we obtain that

$$\begin{aligned} &\|(\mu \boxtimes f) * (\widecheck{\mu \boxtimes f})\|_2^2 = \\ &\int_{G^4} \big\langle (u_1 \cdot f) * (\widecheck{v_1 \cdot f}), (u_2 \cdot f) * (\widecheck{v_2 \cdot f})\big\rangle\, d\mu^{\otimes 4}(u_1, v_1, u_2, v_2). \end{aligned} \tag{17}$$

Let $c : G^4 \to \mathbb{C}, c(u_1, v_1, u_2, v_2) := \big\langle (u_1 \cdot f) * (\widecheck{v_1 \cdot f}), (u_2 \cdot f) * (\widecheck{v_2 \cdot f})\big\rangle$. Then by the Cauchy-Schwarz inequality, we have $\|c\|_1 \le \|c\|_2$ with respect to the probability measure $\mu^{\otimes 4}$, and so by (17), we deduce that

$$\|(\mu \boxtimes f) * (\widecheck{\mu \boxtimes f})\|_2^2 \le \left(\int_{G^4} |c|^2\, d\mu^{\otimes 4}\right)^{1/2}. \tag{18}$$

Another application of the Cauchy-Schwarz inequality implies that

$$\begin{aligned} |c(u_1, v_1, u_2, v_2)|^2 \le& \|(u_1 \cdot f) * (\widecheck{v_1 \cdot f})\|_2^2 \|(u_2 \cdot f) * (\widecheck{v_2 \cdot f})\|_2^2 \\ &= = c(u_1, v_1, u_1, v_1) c(u_2, v_2, u_2, v_2). \end{aligned} \tag{19}$$

By (18) and (19), we obtain that

$$\begin{aligned} \|(\mu \boxtimes f) * (\widecheck{\mu \boxtimes f})\|_2^2 \le& \left(\int_{G^4} c(u_1, v_1, u_1, v_1) c(u_2, v_2, u_2, v_2)\, d\mu^{\otimes 4}\right)^{1/2} \\ =& \int_{G^2} c(u, v, u, v) d\mu^{\otimes 2}. \end{aligned} \tag{20}$$

Notice that by the first part of Lemma 8 for every $u, v \in G$, we have

$$c(u, v, u, v) = \big\langle (u \cdot f) * (\widecheck{v \cdot f}), (u \cdot f) * (\widecheck{v \cdot f})\big\rangle = \big\langle (\widecheck{u \cdot f}) * (u \cdot f), (\widecheck{v \cdot f}) * (v \cdot f)\big\rangle. \tag{21}$$

By (20), (21), and the semi-linearity of the dot product, we deduce that

$$\begin{aligned}\|(\mu\boxtimes f)*(\widetilde{\mu\boxtimes f})\|_2^2 \leq& \int_{G^2}\big\langle(\widetilde{u\cdot f})*(u\cdot f),(\widetilde{v\cdot f})*(v\cdot f)\big\rangle\, d\mu^{\otimes 2}(u,v)\\ =&\big\langle\int_G(\widetilde{u\cdot f})*(u\cdot f)\,d\mu(u),\int_G(\widetilde{v\cdot f})*(v\cdot f)\,d\mu(v)\big\rangle. \qquad (22)\end{aligned}$$

By (22), we conclude that $\|(\mu\boxtimes f)*(\widetilde{\mu\boxtimes f})\|_2 \leq \|\int_G(\widetilde{u\cdot f})*(u\cdot f)\,d\mu(u)\|_2$ which completes the proof. □

We finish this section by proving Lemma 4.

Proof of Lemma 4. Note that the law of $(X^{(1)}\cdot Y^{(1)})^{-1}(X^{(1)}\cdot Y^{(2)})$ is given by

$$\int_G \widetilde{(u\cdot\eta)}*(u\cdot\eta)d\mu(u),$$

and the law of $(X^{(1)}\cdot Y^{(1)})(X^{(2)}\cdot Y^{(2)})^{-1}$ is given by

$$(\mu\boxtimes\eta)*(\widetilde{\mu\boxtimes\eta}).$$

Therefore by Lemma 10 we conclude that

$$\begin{aligned}H_2((X^{(1)}\cdot Y^{(1)})^{-1}(X^{(2)}\cdot Y^{(2)})) =& H_2((X^{(1)}\cdot Y^{(1)})(X^{(2)}\cdot Y^{(2)})^{-1})\\ \geq& H_2((X^{(1)}\cdot Y^{(1)})^{-1}(X^{(1)}\cdot Y^{(2)})),\end{aligned}$$

which finishes the proof. □

Corollary 11. *Suppose H is a finite group, μ is a probability measure on $H\times H$, and η is a probability measure on H. Consider the action*

$$H\times H\curvearrowright H \text{ given by } (x_L,x_R)\cdot y := x_L y x_R^{-1}$$

and the conjugation action $H\curvearrowright H$. Accordingly define $\boxtimes$ from $L^2(H\times H)\times L^2(H)$ to $L^2(H)$ and $\boxtimes$ from $L^2(H)\times L^2(H)$ to $L^2(H)$. Then

$$\|\mu\boxtimes\eta\|^2\leq\|\pi_R[\mu]\boxtimes(\check{\eta}*\eta)\|_2$$

where $\pi_R: H\times H\to H$ is the projection to the right component.

Proof. Suppose $X^{(1)} := (X_L^{(1)}, X_R^{(1)})$ and $X^{(2)} := (X_L^{(2)}, X_R^{(2)})$ are two independent random variables with probability law μ, and $Y^{(1)}$ and $Y^{(2)}$ are two independent random variables with probability law η. Applying Lemma 4 for the given group action $H\times H\curvearrowright H$, we obtain that

$$H_2((X^{(1)}\cdot Y^{(1)})(X^{(2)}\cdot Y^{(2)})^{-1})\geq H_2((X^{(1)}\cdot Y^{(1)})^{-1}(X^{(1)}\cdot Y^{(2)})). \qquad (23)$$

Notice that $(X^{(1)}\cdot Y^{(1)})^{-1}(X^{(1)}\cdot Y^{(2)}) = X_R^{(1)}{Y^{(1)}}^{-1}Y^{(2)}{X_R^{(1)}}^{-1}$, and so the distribution of this random variable is given by

$$\pi_R[\mu]\boxtimes(\check{\eta}*\eta). \qquad (24)$$

We also notice that the distribution of $(X^{(1)} \cdot Y^{(1)})(X^{(2)} \cdot Y^{(2)})^{-1}$ is given by $(\mu \boxtimes \eta) * \widecheck{(\mu \boxtimes \eta)}$. Hence by (23), we conclude that

$$\|(\mu \boxtimes \eta) * \widecheck{(\mu \boxtimes \eta)}\|_2 \leq \|\pi_R[\mu] \boxtimes (\check{\eta} * \eta)\|_2$$

which completes the proof. □

4. An escaping lemma

The main goal of this section is to prove Lemma 5. Our approach is inspired by a method of Lindenstrauss and Varjú developed in [16]. In this section, we will be working with the action $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p) \curvearrowright \mathrm{PSL}_2(\mathbb{F}_p)$ given by $(u, v)\cdot x = uxv^{-1}$, and the conjugation action $\mathrm{PSL}_2(\mathbb{F}_p) \curvearrowright \mathrm{PSL}_2(\mathbb{F}_p)$ given by $x \cdot y = xyx^{-1}$. We reformulate the statement of Lemma 5 in terms of distributions.

Lemma 12. *Let μ be a measure on $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$. Suppose that the minimum of μ on its support is at least $\alpha_0 > 0$ and $\mathcal{L}(\pi_R[\mu]) \geq c_0 > 0$, where $\pi_R : \mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p) \to \mathrm{PSL}_2(\mathbb{F}_p)$ is the projection to the right component. Suppose that the support of μ generates $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$. Then there exist constants $L, C \gg_{c_0,\alpha_0} 1$ such that for every probability measure η on $\mathrm{PSL}_2(\mathbb{F}_p)$ and every integer $\ell \geq L \log p$*

$$\|\mu^{*(\ell)} \boxtimes \eta\|_2 \leq C|\,\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)|^{-1/24}.$$

We start by proving that if $\check{\eta} * \eta$ is almost a point mass at the identity, then η is almost a point mass. Here, a measure is viewed as almost a point mass if a single point is responsible for a κ portion of its L^2-norm. This lemma is essentially proved in [16, Lemma 5].

Lemma 13. *Suppose that η is a probability measure on a finite group H and $\|\eta\|_\infty \leq \kappa\|\eta\|_2$ where κ is a positive number. Then we have that*

$$\|\eta\|^2 \geq \sqrt{2 - \kappa^2}\ \ \check{\eta} * \eta(1).$$

Proof. Notice that

$$\check{\eta} * \eta(1) = \sum_{y \in H} \check{\eta}(y^{-1})\eta(y) = \sum_{y \in H} \eta(y)^2 = \|\eta\|_2^2. \tag{25}$$

Next by a simple computation we obtain

$$\sum_{y \in H\setminus\{1\}} \check{\eta}*\eta(y)^2 = \sum_{y\neq 1} \Big(\sum_{y_1 y_2 = y} \check{\eta}(y_1)\eta(y_2)\Big)^2 \geq \sum_{y\neq 1} \sum_{y_1 y_2 = y} \check{\eta}(y_1)^2\eta(y_2)^2$$

$$= \Big(\sum_{y_1} \check{\eta}(y_1)^2\Big)\Big(\sum_{y_2} \eta(y_2)^2\Big) - \sum_{y_2} \check{\eta}(y_2^{-1})^2\eta(y_2)^2 = \|\eta\|_2^4 - \|\eta\|_4^4. \tag{26}$$

On the other hand, because $\kappa\|\eta\|_2 \geq \|\eta\|_\infty$, we deduce that

$$\|\eta\|_4^4 = \sum_{y} \eta(y)^4 \leq \|\eta\|_\infty^2\|\eta\|_2^2 \leq \kappa^2\|\eta\|_2^4. \tag{27}$$

By (25), (26), and (27), we obtain that

$$\begin{aligned}\|\check{\eta} * \eta\|_2^2 \geq & \check{\eta} * \eta(1)^2 + (\|\eta\|_2^4 - \|\eta\|_4^4) \geq \|\eta\|_2^4 + (\|\eta\|_2^4 - \kappa^2 \|\eta\|_2^4) \\ = & (2-\kappa^2)\, \check{\eta} * \eta(1)^2. \end{aligned} \tag{28}$$

This completes the proof. □

Next we show that if Z is a random variable with values in $\mathrm{PSL}_2(\mathbb{F}_p)$ and uniform distribution and Y is an independent random variable with values in $\mathrm{PSL}_2(\mathbb{F}_p)$, then

$$H_2(ZYZ^{-1}) \geq \min\{-2\log(\mathbb{P}(Y=1)), \log p\} - 2\log 2.$$

Notice that since $\mathbb{P}(ZYZ^{-1}=1) = \mathbb{P}(Y=1)$, $H_2(ZYZ^{-1}) \leq -2\log(\mathbb{P}(Y=1))$. Similar to the previous lemmas in this section, we formulate the lemma in terms of distributions.

Lemma 14. *Let $H := \mathrm{PSL}_2(\mathbb{F}_p)$ and μ_H be the probability counting measure on H. Suppose η is a probability measure on H. Consider the conjugation action $H \curvearrowright H$ given by $z \cdot y := zyz^{-1}$, and let $\mu_H \boxtimes \eta$ be the convolution associated to the the conjugation action. Then we have that*

$$\|\mu_H \boxtimes \eta\|_2 \leq \eta(1) + p^{-1/2}.$$

Proof. For every $x \in H$, let $\mathrm{Cl}(x)$ be the conjugacy class of x. Since H is generated by its elements of order p, for every proper subgroup $K \subsetneq H$ we have that $[H:K] \geq p$. Hence for every $x \in H$, either $|\mathrm{Cl}(x)| \geq p$ or $\mathrm{Cl}(x) = \{x\}$. The latter holds exactly when x is in the center $Z(H)$ of H. Notice that $Z(H) = \{1\}$, and so for every $x \in H \setminus \{1\}$

$$|\mathrm{Cl}(x)| \geq p. \tag{29}$$

Let's recall that for every $A \subseteq H$, μ_A denotes the probability counting measure on A. Notice that for every $x \in H$, we have

$$\mu_H \boxtimes \mu_{\{x\}} = \mu_{\mathrm{Cl}(x)}.$$

Hence by the bilinearity of $\boxtimes$, we have

$$\mu_H \boxtimes \eta = \sum_{x \in H} \eta(x)\mu_{\mathrm{Cl}(x)} = \sum_{c \in \mathrm{Conj}(H)} \eta(c)\mu_c, \tag{30}$$

where $\mathrm{Conj}(H)$ is the set of all the conjugacy classes of H. By (30), the triangle inequality, and (29), we conclude that

$$\begin{aligned}\|\mu_H \boxtimes \eta\|_2 \leq \sum_{c \in \mathrm{Conj}(H)} \eta(c)\|\mu_c\|_2 = & \sum_{c \in \mathrm{Conj}(H)} \eta(c)|c|^{-1/2} \\ \leq & \eta(1) + \eta(H \setminus \{1\})\, p^{-1/2}. \end{aligned}$$

This completes the proof. □

Before proving Lemma 5, we show a lemma on symmetric generating sets of a finite group.

Lemma 15. *Suppose S is a symmetric generating set of a group G. Suppose G acts on a set Ω and there are distinct points $x_1, x_2 \in \Omega$ such that for every $s \in S$, $s \cdot x_1 = x_2$. Then G has a subgroup of index 2.*

Proof. By assumption for every $s \in S$, we have $s^{-1} \cdot x_2 = x_1$. Because S is symmetric, we deduce that for every $s \in S$, $s \cdot x_2 = x_1$. Hence for every $s_1, \ldots, s_k \in S$, we have that $(s_1 \ldots s_k) \cdot x_1 \in \{x_1, x_2\}$. On other hand, because S is a symmetric generating set of G, $G = \{1\} \cup \bigcup_{k=1}^{\infty} \prod_k S$; and so the G-orbit of x_1 has exactly two points. Therefore the stabilizer subgroup of G with respect to x_1 is of index 2. This completes the proof. $\square$

Notice that $\mathrm{PSL}_2(\mathbb{F}_p)$ is generated by its p-elements, and so if $p > 2$, $\mathrm{PSL}_2(\mathbb{F}_p)$ does not have a subgroup of order 2.

In the rest of this section, we prove Lemma 12 which is a reformulation of Lemma 5.

Proof of Lemma 12. Let's recall that α_0 is the minimum of μ in its support. Choose $0 < \kappa_0 < 1$ such that $\sqrt{\alpha_0^2 + (1-\alpha_0)^2} + \sqrt{1-\kappa_0^2} < 1$. Let η be the probability law of Y. We are going to consider two cases mostly depending on whether or not η is almost a point mass or not.

Case 1. *Suppose that $\|\eta\|_\infty / \|\eta\|_2 > \kappa_0$.*

In this case, there exists $x_0 \in \mathrm{PSL}_2(\mathbb{F}_p)$ such that $\eta(x_0)^2 > \kappa_0^2 \|\eta\|_2^2$. Let $\eta_{x_0}^{\perp} := \eta \mathbb{1}_{\mathrm{PSL}_2(\mathbb{F}_p) \setminus \{x_0\}}$ where $\mathbb{1}_{\mathrm{PSL}_2(\mathbb{F}_p) \setminus \{x_0\}}$ is the characteristic function of $\mathrm{PSL}_2(\mathbb{F}_p) \setminus \{x_0\}$. Notice that

$$\eta = \eta(x_0)\mu_{\{x_0\}} + \eta_{x_0}^{\perp} \quad \text{and} \quad \mu_{\{x_0\}} \perp \eta_{x_0}^{\perp}. \tag{31}$$

By (31), we have that $\|\eta\|_2^2 = \eta(x_0)^2 + \|\eta_{x_0}^{\perp}\|_2^2$, and so

$$\|\eta_{x_0}^{\perp}\|_2^2 < (1-\kappa_0^2)\|\eta\|_2^2. \tag{32}$$

Moreover, by (31), we obtain that $\mu \boxtimes \eta = \eta(x_0)\, \mu \boxtimes \mu_{\{x_0\}} + \mu \boxtimes \eta_{x_0}^{\perp}$. Notice that

$$\begin{aligned} \mu \boxtimes \mu_{\{x_0\}} &= \sum_{y \in \mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)} \mu(y) \mu_{\{y \cdot x_0\}} \\ &= \sum_{y G_{x_0} \in G/G_{x_0}} \mu(y G_{x_0}) \mu_{y \cdot x_0}, \end{aligned} \tag{33}$$

where $G := \mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$ and G_{x_0} is the stabilizer subgroup of G with respect to x_0. Since the support of μ is a symmetric generating set of $\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)$ and the minimum of μ on its support is α_0, by Lemma 15 and (33) we conclude that

$$\|\mu \boxtimes \mu_{\{x_0\}}\|_2 \leq (\alpha_0^2 + (1-\alpha_0)^2)^{1/2}. \tag{34}$$

Therefore by the triangle inequality, (34), and (32), we conclude that

$$\begin{aligned}\|\mu \boxtimes \eta\|_2 &\leq \eta(x_0)\|\mu \boxtimes \mu_{\{x_0\}}\|_2 + \|\mu \boxtimes \eta_{x_0}^{\perp}\|_2 \\ &\leq \eta(x_0)(\alpha_0^2 + (1-\alpha_0)^2)^{1/2} + \|\eta_{x_0}^{\perp}\|_2 \\ &\leq \left((\alpha_0^2 + (1-\alpha_0)^2)^{1/2} + (1-\kappa_0^2)^{1/2}\right)\|\eta\|_2 \qquad (35)\end{aligned}$$

Case 2. *Suppose that $\|\eta\|_\infty/\|\eta\|_2 \leq \kappa_0$.*

Choose $0 < \kappa_1 < 1$ such that $(2-\kappa_0^2)^{-1/2} + 2\kappa_1 < 1$. Since $\mathcal{L}(\pi_R[\mu]) > c_0$, there is a positive integer ℓ_0 which is bounded by a function of c_0 (and κ_1) such that $\lambda(\pi_R[\mu]^{*(\ell_0)}) < \kappa_1$. Set $H := \mathrm{PSL}_2(\mathbb{F}_p)$ and $\nu := \mu^{*(\ell_0)}$, and so $\lambda(\pi_R[\nu]) < \kappa_1$. Then by Corollary 11, we obtain that

$$|||\nu\boxtimes\eta|||^2 \leq \|\pi_R[\nu]\boxtimes(\check{\eta}*\eta)\|_2 \leq \|(\pi_R[\nu]-\mu_H)\boxtimes(\check{\eta}*\eta)\|_2 + \|\mu_H\boxtimes(\check{\eta}*\eta)\|_2. \qquad (36)$$

By (36), (9), and Lemma 14, we deduce that

$$|||\nu \boxtimes \eta|||^2 \leq \kappa_1 |||\eta|||^2 + \check{\eta} * \eta(1) + p^{-1/2}. \qquad (37)$$

By (37) and Lemma 13, we have that

$$|||\nu \boxtimes \eta|||^2 \leq \kappa_1 |||\eta|||^2 + (2-\kappa_0^2)^{-1/2}|||\eta|||^2 + p^{-1/2}. \qquad (38)$$

Next we combine the inequalities in (35) and (38). Suppose β is a positive number less than 1 which is more than

$$\max\left\{(2-\kappa_0^2)^{-1/2} + 2\kappa_1, (\alpha_0^2 + (1-\alpha_0)^2)^{1/2} + (1-\kappa_0^2)^{1/2}\right\}.$$

Then by (35) and (38), at least one of the following three inequalities hold. Either

$$\|\nu\boxtimes\eta\|_2 \leq \beta\|\eta\|_2, \quad \text{or} \quad |||\nu\boxtimes\eta|||^2 \leq \beta|||\eta|||^2, \quad \text{or} \quad |||\eta|||^2 \leq \kappa_1^{-1}p^{-1/2}. \qquad (39)$$

Applying (39) repeatedly, by Lemma 9, we conclude that for every integer $\ell \geq 2\log p/(-\log\beta)$ at least one of the following three inequalities hold. Either

$$\begin{aligned}\|\nu^{*(\ell)} \boxtimes \eta\|_2 <& \beta^{\ell/2} \leq 1/p \quad \text{or} \\ |||\nu^{*(\ell)} \boxtimes \eta|||^2 <& \beta^{\ell/2} \leq 1/p \quad \text{or} \qquad (40) \\ |||\nu^{*(\ell)} \boxtimes \eta|||^2 <& \kappa_1^{-1}p^{-1/2}.\end{aligned}$$

By Lemma 9 part (1) and (40), for every $\ell \geq 2\log p/(-\log\beta)$ the following holds

$$\|\nu^{*(\ell)} \boxtimes \eta\|_2 \leq \kappa_1^{-1/2}p^{-1/4}.$$

Since $|\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)| \leq p^6$, we conclude that

$$\|\nu^{*(\ell)} \boxtimes \eta\|_2^2 \leq \kappa_1^{-1}|\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)|^{-1/12},$$

which completes the proof. □

5. Proving the main result.

In this section, we will be using the Bourgain-Gamburd method to prove Theorem 1 based on Lemma 5. Bourgain and Gamburd in their seminal work (see [3]) laid out a blueprint for finding a bound for the spectral gap of a random-walk in a (single scale) finite group. One of the important results that they proved is the following proposition (see [3], [20], [8], [9]).

Proposition 16. *Let G be a finite group. Suppose X and Y are two independent random variables with values in G, and $K \geq 2$. If*

$$H_2(XY) \leq \frac{H_2(X)+H_2(Y)}{2} + \log K,$$

then there exist $A \subseteq G$ and a universal fixed positive number R with the following properties

(1) (Approximate structure) *A is a K^R-approximate subgroup; by this we mean A is symmetric, $1 \in A$, and there is a subset B of $A \cdot A$ such that $|B| \leq K^R$ and $A \cdot A \subseteq B \cdot A \cap A \cdot B$.*

(2) (Controlling the size) *$|\log |A| - H_2(X)| \leq R \log K$.*

(3) (Almost equidistribution) *For every $a \in A$,*

$$\mathbb{P}(X'X = a) \geq \frac{1}{K^R|A|}$$

where X' is an independent random variable that has an identical distribution with X^{-1}.

where R is a fixed number.

In this section, first using Proposition 16, we will prove the following lemma.

Lemma 17. *Let $\varepsilon > 0$ and $H := \mathrm{SL}_2(\mathbb{F}_p)$. Suppose $Y := (Y_L, Y_R)$ is a random variable with values in $H \times H$, distribution $\mathcal{P}$, and the following properties:*

(1) (Close to a coupling) *For every $y \in H$, $\mathbb{P}(Y_L = y) \leq 2|H|^{-1}$ and $\mathbb{P}(Y_R = y) \leq 2|H|^{-1}$.*

(2) (Room for improvement) *$H_2(Y) \leq (1-\varepsilon)\log|H \times H|$.*

Then there is a positive number γ_0 depending on ε and a universal positive constant R such that for every positive $\gamma \leq \gamma_0$ at least one of the following statements hold.

(0) (Exceptional cases) *$|H|^{R\gamma} \leq 2$.*

(1) (Gaining entropy) *$H_2(Y_2) \geq H_2(Y) + \gamma \log|H \times H|$ where Y_2 is the 2-step random walk with respect to $\mathcal{P}$.*

(2) (Graph of an automorphism) *$\mathbb{P}(\bar{\iota}(Y_2) \in \Gamma_\phi) \geq |H \times H|^{-R\gamma}$ for some automorphism ϕ of $\mathrm{PSL}_2(\mathbb{F}_p)$, where*

$$\bar{\iota} : H \times H \to \mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p), \quad \bar{\iota}(x_L, x_R) := (\{\pm x_L\}, \{\pm x_R\})$$

and $\Gamma_\phi := \{(x, \phi(x)) \mid x \in \mathrm{PSL}_2(\mathbb{F}_p)\}$.

Using Lemma 17, Lemma 5, and Lemma 7, Theorem 1 will be proved.

Proof of Lemma 17. Suppose $\gamma \leq \gamma_0$, where γ_0 is a sufficiently small positive number to be specified later. Let's assume that we are not in the *exceptional cases* and we do not *gain enough entropy*; that means $|H|^{R\gamma} > 2$ and $H_2(Y_2) < H_2(Y) + \gamma \log |H \times H|$. Then by Proposition 16, we obtain that there is an $|H \times H|^{R\gamma}$-approximate subgroup A such that

$$|\log |A| - H_2(Y)| \leq R\gamma \log |H \times H| \quad \text{and} \quad \mathbb{P}(Y_2 \in A) \geq |H \times H|^{-R\gamma}. \tag{41}$$

Notice that by the *close to a coupling* condition, we have

$$\mathbb{P}((Y_L)_2 \in \pi_L(A)) \leq 2|\pi_L(A)||H|^{-1}.$$

Therefore, we obtain

$$|H \times H|^{-R\gamma} \leq \mathbb{P}(Y_2 \in A) \leq \mathbb{P}((Y_L)_2 \in \pi_L(A)) \leq 2|\pi_L(A)||H|^{-1}.$$

A similar result holds for the projection to the right copy of H. Hence

$$|\pi_L(A)| \geq |H|^{1-3R\gamma} \quad \text{and} \quad |\pi_R(A)| \geq |H|^{1-3R\gamma}. \tag{42}$$

Notice that by a result of Frobenius [7], degree of every non-trivial irreducible representation of $\mathrm{SL}_2(\mathbb{F}_p)$ is at least $(p-1)/2$, and so there exists a positive number c_1 such that for all primes $p \geq 5$, $\mathrm{SL}_2(\mathbb{F}_p)$ is c_1-quasi-random. Therefore by a result of Gowers (see [11] and [2]) and (42), for $\gamma < c_1/(9R)$, we have that

$$\textstyle\prod_3 \pi_L(A) = \mathrm{SL}_2(\mathbb{F}_p) \quad \text{and} \quad \prod_3 \pi_R(A) = \mathrm{SL}_2(\mathbb{F}_p). \tag{43}$$

Claim 1. *In the above setting, for a small enough γ_0 depending only on ε, we have*

$$\textstyle\prod_9 A \cap (H \times \{\pm 1\}) \subseteq \{(\pm 1, \pm 1)\} \quad \textit{and} \quad \prod_9 A \cap (\{\pm 1\} \times H) \subseteq \{(\pm 1, \pm 1)\}.$$

Proof of Claim 1. It is clear that by symmetry it is enough to prove only one of the inclusions. Suppose to the contrary that, we have (x, e) is in $\prod_9 A$, for some $x \in \mathrm{SL}_2(\mathbb{F}_p) \setminus \{\pm 1\}$ and $e \in \{\pm 1\}$. Since $\pi_L(\prod_3 A) = \mathrm{SL}_2(\mathbb{F}_p)$, we obtain that

$$\mathrm{Cl}(x) \times \{e\} \subseteq \textstyle\prod_{15} A. \tag{44}$$

By [21, Theorem 2.2 and 2.3], we have that $\prod_3 \mathrm{Cl}(x) = \mathrm{SL}_2(\mathbb{F}_p)$, for every prime $p \geq 5$. Therefore, by (44), we deduce that $\mathrm{SL}_2(\mathbb{F}_p) \times \{1\} \subseteq \prod_{90} A$. Hence by (43), we obtain that

$$\textstyle\prod_{93} A = \mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p). \tag{45}$$

Because A is an $|H \times H|^{R\gamma}$-approximate subgroup,

$$|\textstyle\prod_{93} A| \leq |H \times H|^{92R\gamma}|A|. \tag{46}$$

Thus by (45) and (46), we conclude that

$$|A| \geq |H \times H|^{1-92R\gamma}. \tag{47}$$

By the condition on the *room for improvement* and (41), we deduce that

$$|A| \leq |H \times H|^{1-\varepsilon+R\gamma}. \tag{48}$$

By (47) and (48), we reach to a contradiction if $\gamma < \varepsilon/93R$. This completes proof of Claim 1.

Claim 2. *Let* $\iota : \mathrm{SL}_2(\mathbb{F}_p) \to \mathrm{PSL}_2(\mathbb{F}_p), \iota(x) := \{\pm x\}$ *and* $\overline{A} := \iota(A)$*. Then there is an automorphism* $\phi : \mathrm{PSL}_2(\mathbb{F}_p) \to \mathrm{PSL}_2(\mathbb{F}_p)$ *such that* $\prod_3 \overline{A} = \Gamma_\phi$*.*

Proof of Claim 2. By (43), for every $x \in \mathrm{SL}_2(\mathbb{F}_p)$, there is $\widetilde{\phi}(x) \in \mathrm{SL}_2(\mathbb{F}_p)$ such that $(x, \widetilde{\phi}(x))$ is in $\prod_3 A$. Since A is a symmetric set, for every $x, y \in \mathrm{SL}_2(\mathbb{F}_p)$ we obtain that

$$(1, \widetilde{\phi}(x)\widetilde{\phi}(y)\widetilde{\phi}(xy)^{-1}) = (x, \widetilde{\phi}(x))(y, \widetilde{\phi}(y))(xy, \widetilde{\phi}(xy))^{-1} \in \textstyle\prod_9 A. \tag{49}$$

By (49) and Claim 1, we deduce that $\widetilde{\phi}(xy) = \pm\widetilde{\phi}(x)\widetilde{\phi}(y)$ for every x, y in $\mathrm{SL}_2(\mathbb{F}_p)$. Notice that for every $x \in \mathrm{SL}_2(\mathbb{F}_p)$, $(x, y) \in \prod_3 A$ implies that $(1, \widetilde{\phi}(x)y^{-1}) \in \prod_9 A$. Hence by Claim 1, we have $y = \pm\widetilde{\phi}(x)$. Let

$$\phi : \mathrm{PSL}_2(\mathbb{F}_p) \to \mathrm{PSL}_2(\mathbb{F}_p), \quad \phi(\{\pm x\}) := \{\pm\widetilde{\phi}(x)\}.$$

Notice that because $\widetilde{\phi}(xy) = \pm\widetilde{\phi}(x)\widetilde{\phi}(y)$, we have that ϕ is a well-defined group homomorphism. By (43), we deduce that ϕ is surjective, and Claim 1 implies that ϕ is injective. Altogether we conclude that ϕ is an automorphism of $\mathrm{PSL}_2(\mathbb{F}_p)$ and $\iota(\prod_3 A) = \Gamma_\phi$. This completes proof of Claim 2.

By Claim 2 and (41), we conclude that $\mathbb{P}(\bar{\iota}(Y_2) \in \Gamma_\phi) \geq |H \times H|^{-R\gamma}$, which completes the proof. □

Proof of Theorem 1. Because $\mathcal{L}(\pi_L[\mu])$ and $\mathcal{L}(\pi_R[\mu])$ are at least c_0, by the Cauchy-Schwarz inequality and (9), for $\ell_0 \geq (2\log|\mathrm{SL}_2(\mathbb{F}_p)|)/c_0$, we have

$$\begin{aligned} \|\pi_L[\mu]^{*(\ell_0)} - \mu_H\|_\infty =& \|((\pi_L[\mu]^{*(\ell_0)} - \mu_H) * \mu_{\{1\}}) * \mu_{\{1\}}\|_\infty \\ \leq& \|(\pi_L[\mu]^{*(\ell_0)} - \mu_H) * \mu_{\{1\}}\|_2 \leq |H|^{-2}. \end{aligned} \tag{50}$$

By (50) and its counterpart for $\pi_R[\mu]$, for every $x \in H$ we obtain that

$$\begin{aligned} &|\pi_L[\mu^{*(\ell_0)}](x) - |H|^{-1}| \leq 2|H|^{-1} \quad \text{and} \\ &|\pi_R[\mu^{*(\ell_0)}](x) - |H|^{-1}| \leq 2|H|^{-1}. \end{aligned} \tag{51}$$

Suppose ℓ is a positive integer which is at least ℓ_0 and will be specified later. Let Y be a random variable with values in $H \times H$ and distribution $\mu^{*(\ell)}$.

As it has been mentioned earlier, by a result of Frobenius every non-trivial representation of $\mathrm{SL}_2(\mathbb{F}_p)$ is of dimension at least $(p-1)/2$. Hence there is a positive number c_2 such that $\mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$ is an c_2-quasi-random group.

By Lemma 17, there is a positive number γ_0 depending on c_2 and a universal positive constant R such that for every positive number $\gamma \leq \gamma_0$, one of the following statements hold. Either $|H|^{R\gamma} \leq 2$, or

$$H_2(Y) > (1 - c_2/2)\log|H \times H|, \tag{52}$$

or there is an automorphism ϕ of $\mathrm{PSL}_2(\mathbb{F}_p)$ such that

$$\mathbb{P}(\iota(Y_2) \in \Gamma_\phi) \geq |H \times H|^{-R\gamma}, \tag{53}$$

where Γ_ϕ is the graph of ϕ, or

$$H_2(Y_2) \geq H_2(Y) + \gamma \log(|H \times H|). \tag{54}$$

Claim 1. *There are positive numbers p_0, γ_1, and L' all depending only on c_0 and α_0 such that Equation (53) does not hold for any automorphism ϕ, positive integer $\ell \geq L' \log|H \times H|$, positive number $\gamma = \gamma_1$, and prime $p \geq p_0$.*

Proof of Claim 1. Suppose to the contrary that (53) holds for an automorphism ϕ of $\mathrm{PSL}_2(\mathbb{F}_p)$. Let $X := (X_L, X_R)$ be a random variable with values in $\mathrm{SL}_2(\mathbb{F}_p)\times\mathrm{SL}_2(\mathbb{F}_p)$ and distribution μ. Let $\overline{X} := (\iota(X_L), \phi^{-1}(\iota(X_R)))$. Then by Lemma 5, there exist constants $L, C \gg_{c_0,\alpha_0} 1$ such that for every integer $\ell \geq L \log|H \times H|$ we have

$$H_2(\overline{X}_\ell \cdot Z) \geq \frac{1}{12} \log|\,\mathrm{PSL}_2(\mathbb{F}_p) \times \mathrm{PSL}_2(\mathbb{F}_p)| - C \tag{55}$$

where Z is a random variable with values in $\mathrm{PSL}_2(\mathbb{F}_p)$ and distribution $\mu_{\{1\}}$. Notice that

$$\begin{aligned}\mathbb{P}(\overline{X}_\ell \cdot Z = 1) =& \mathbb{P}(\iota((X_L)_\ell) = \phi^{-1}(\iota((X_R)_\ell))) = \mathbb{P}(\overline{\iota}(X_\ell) \in \Gamma_\phi) \\ \geq& |H \times H|^{-R\gamma_0}.\end{aligned} \tag{56}$$

By (56), we conclude that

$$H_2(\overline{X}_\ell \cdot Z) \leq R\gamma_0 \log|H \times H|. \tag{57}$$

By (55) and (57), we reach to a contradiction if $p \gg_{c_0,\alpha_0} 1$. This completes proof of Claim 1.

Claim 2. *Suppose ℓ_1 is an integer in the interval*

$$[L' \log|H \times H|, 2L' \log|H \times H|].$$

Let X be a random variable with distribution μ. Then

$$H_2(X_\ell) > (1 - c_2/2) \log|H \times H|$$

for every integer $\ell \geq 2^{\gamma_1^{-1}} \ell_1$.

Proof of Claim 2. Let $Y := X_{\ell_1}$, and consider the sequence of Renyi entropies $\{H_2(Y_{2^k})\}_k$. By Claim 1, for every positive integer k, either

$$\begin{aligned} H_2(Y_{2^k}) >& (1 - c_2/2) \log|H \times H| \quad \text{or} \\ H_2(Y_{2^{k+1}}) \geq& H_2(Y_{2^k}) + \gamma_1 \log(|H \times H|). \end{aligned}$$

Hence $H_2(Y_{2^k}) > (1 - c_2/2) \log(|H \times H|)$ if k is an integer more than γ_1^{-1}. This completes proof of Claim 2.

By Claim 2 and Lemma 7, we conclude that either $p \ll_{c_0,\alpha_0} 1$ or

$$\mathcal{L}(\mu) \geq \frac{c_2}{2^{\gamma_1^{-1}+1} L'} \gg_{\alpha_0,c_0} 1.$$

This completes proof of the main theorem. □

6. Proof of Corollary 3

Proof of Corollary 3. We can and will assume that $p \geq 5$. Suppose ρ_1 and ρ_2 are two distinct points of $\mathrm{Rep}_2(\mathbb{F}_p)_\delta$. Let

$$\Omega := \{(\rho_1(a), \rho_2(a))^{\pm 1}, (\rho_1(b), \rho_2(b))^{\pm 1}\}$$

Let H be the group generated by Ω. Notice that $\pi_L(H) = \mathrm{Im}(\rho_1) = \mathrm{SL}_2(\mathbb{F}_p)$ and $\pi_R(H) = \mathrm{Im}(\rho_2) = \mathrm{SL}_2(\mathbb{F}_p)$. Therefore for every $x \in \mathrm{SL}_2(\mathbb{F}_p)$, there is $\widetilde{\phi}(x) \in \mathrm{SL}_2(\mathbb{F}_p)$ such that $(x, \widetilde{\phi}(x)) \in H$. Notice that if $(y, 1) \in H$, then for every $x \in \mathrm{SL}_2(\mathbb{F}_p)$,

$$(xyx^{-1}, 1) = (x, \widetilde{\phi}(x))(y, 1)(x, \widetilde{\phi}(x))^{-1} \in H.$$

Therefore $\pi_L(H \cap \ker \pi_R)$ is a normal subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$, and so it is either $\mathrm{SL}_2(\mathbb{F}_p)$ or central. Similarly $\pi_R(H \cap \ker \pi_L)$ is either $\mathrm{SL}_2(\mathbb{F}_p)$ or central. Altogether, we have that either $H = \mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$ or $\bar{\iota}(H) = \Gamma_\phi$ where Γ_ϕ is the graph of an automorphism ϕ of $\mathrm{PSL}_2(\mathbb{F}_p)$. Notice that by [22, Theorem 3.2], there is an automorphism $\widehat{\phi}$ of $\mathrm{SL}_2(\mathbb{F}_p)$ such that $\phi(\{\pm x\}) = \{\pm \widehat{\phi}(x)\}$ for every $x \in \mathrm{SL}_2(\mathbb{F}_p)$.

Let μ be the probability counting measure on Ω. Since $\rho_1, \rho_2 \in \mathrm{Rep}_2(\mathbb{F}_p)_\delta$, $\mathcal{L}(\pi_L[\mu])$ and $\mathcal{L}(\pi_R[\mu])$ are at least δ. Hence if $H = \mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$, then by Theorem 1, $\lambda(\mu) \leq \lambda_0$ for some positive number λ_0 which is less than 1 and only depends on δ. Then for every positive integer ℓ, we have

$$\begin{aligned} &\mu^{*(\ell)}_{\{a^{\pm}, b^{\pm}\}}(\{w \in F_2 \,| \rho_1(w) = \rho_2(w)\}) \\ &\qquad \leq |\,\mathrm{SL}_2(\mathbb{F}_p)|^{-1} + |(\mu_\Omega^{*(\ell)} - \mu_H)(\{(x, x) \mid x \in \mathrm{SL}_2(\mathbb{F}_p)\})| \\ &\qquad \leq |\,\mathrm{SL}_2(\mathbb{F}_p)|^{-1} + \lambda_0^\ell |\,\mathrm{SL}_2(\mathbb{F}_p)|. \end{aligned}$$

If $\bar{\iota}(H) = \Gamma_\phi$, then for every $w \in F_2$ we have $\iota(\rho_2(w)) = \phi(\iota(\rho_1(w)))$. This means that for every $w \in F_2$ we have $\rho_2(w) = \pm \widehat{\phi}(\rho_1(w))$. Hence

$$\begin{aligned} &\mu^{*(\ell)}_{\{a^{\pm}, b^{\pm}\}}(\{w \in F_2 \,| \rho_1(w) = \rho_2(w)\}) \\ &\qquad \leq \mu_\Omega^{*(\ell)}(\{(x, \pm\widehat{\phi}(x)) \mid \pm\widehat{\phi}(x) = x\}) \\ &\qquad \leq \mu^{*(\ell)}_{\pi_L(\Omega)}(\{x \in \mathrm{SL}_2(\mathbb{F}_p) \mid \phi(x) = \pm x\}) \\ &\qquad \leq \frac{|\{x \in \mathrm{SL}_2(\mathbb{F}_p) \mid \phi(x) = \pm x\}|}{|\,\mathrm{SL}_2(\mathbb{F}_p)|} + 2^{-\delta\ell} |\,\mathrm{SL}_2(\mathbb{F}_p)|. \end{aligned}$$

Notice that since $\rho_1 \neq \rho_2$, $\widehat{\phi}$ is a non-trivial automorphism of $\mathrm{SL}_2(\mathbb{F}_p)$. Therefore

$$\{x \in \mathrm{SL}_2(\mathbb{F}_p) \mid \phi(x) = \pm x\}$$

is a proper subgroup of $\mathrm{SL}_2(\mathbb{F}_p)$. Hence we conclude that

$$\mu^{*(\ell)}_{\{a^{\pm}, b^{\pm}\}}(\{w \in F_2 \mid \rho_1(w) = \rho_2(w)\}) \leq p^{-1} + 2^{-\delta\ell} |\,\mathrm{SL}_2(\mathbb{F}_p)|.$$

Suppose $[\rho_1] \neq [\rho_2]$. In the above setting, if $H = \mathrm{SL}_2(\mathbb{F}_p) \times \mathrm{SL}_2(\mathbb{F}_p)$, then by a similar argument as above we have

$$\begin{aligned}
&\mu^{*(\ell)}_{\{a^{\pm 1}, b^{\pm 1}\}}(\{w \in F_2 \, | t_w([\rho_1]) = t_w([\rho_2])\}) \\
&\qquad\qquad \leq \mu^{*(\ell)}_{\Omega}(\{(x_1, x_2) \mid \mathrm{Tr}(x_1) = \mathrm{Tr}(x_2)\}) \\
&\qquad\qquad \leq \frac{|\{(x_1, x_2) \mid \mathrm{Tr}(x_1) = \mathrm{Tr}(x_2)\}|}{|\mathrm{SL}_2(\mathbb{F}_p)|^2} + \lambda_0^{\ell} |\mathrm{SL}_2(\mathbb{F}_p)|. \qquad (58)
\end{aligned}$$

Notice that for a given $a \in \mathbb{F}_p$, we have

$$\begin{aligned}
&|\{x \in \mathrm{SL}_2(\mathbb{F}_p) \mid \mathrm{Tr}(x) = a\}| = \\
&\qquad |\{(x_{11}, x_{12}, a - x_{11}, x_{21}) \in \mathbb{F}_p^4 \mid x_{11}(a - x_{11}) - x_{12}x_{21} = 1\}|,
\end{aligned}$$

and for a given x_{12}, x_{21} there are at most 2 choices for x_{11}. Hence

$$|\{x \in \mathrm{SL}_2(\mathbb{F}_p) \mid \mathrm{Tr}(x) = a\}| \leq 2p^2 \qquad (59)$$

for every $a \in \mathbb{F}_p$. This implies that

(60)
$$|\{(x_1, x_2) \mid \mathrm{Tr}(x_1) = \mathrm{Tr}(x_2)\}| = \sum_{a \in \mathbb{F}_p} |\{x \in \mathrm{SL}_2(\mathbb{F}_p) \mid \mathrm{Tr}(x) = a\}|^2 \leq 4p^5.$$

By (58) and (60), for $p \geq 5$, we obtain

$$\begin{aligned}
\mu^{*(\ell)}_{\{a^{\pm}, b^{\pm}\}}(\{w \in F_2 \mid \rho_1(w) = \rho_2(w)\}) \leq & \frac{4}{p}(1 - p^{-2})^{-2} + \lambda_0^{\ell} |\mathrm{SL}_2(\mathbb{F}_p)| \\
< & \frac{5}{p} + \lambda_0^{\ell} |\mathrm{SL}_2(\mathbb{F}_p)|.
\end{aligned}$$

In the above setting, if $\bar{\iota}(H) = \Gamma_\phi$, then as we discussed above for some automorphism $\widehat{\phi}$ of $\mathrm{SL}_2(\mathbb{F}_p)$, we have $\rho_2(w) = \pm\widehat{\phi}(\rho_1(w))$. Hence by [22, Theorem 3.2], $\mathrm{Tr}(\rho_2(w)) = \pm \mathrm{Tr}(\rho_1(w))$ for every $w \in F_2$. This completes the proof. □

References

[1] N. Alon. Eigenvalues and expanders. volume 6, pages 83–96. 1986. Theory of computing (Singer Island, Fla., 1984).

[2] László Babai, Nikolay Nikolov, and László Pyber. Product growth and mixing in finite groups. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 248–257. ACM, New York, 2008.

[3] Jean Bourgain and Alex Gamburd. Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. *Ann. of Math. (2)*, 167(2):625–642, 2008.

[4] John B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1985.

[5] Jozef Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.*, 284(2):787–794, 1984.

[6] Pavel Etingof, Oleg Golberg, Sebastian Hensel, Tiankai Liu, Alex Schwendner, Dmitry Vaintrob, and Elena Yudovina. *Introduction to representation theory*, volume 59 of *Student Mathematical Library*. American Mathematical Society, Providence, RI, 2011. With historical interludes by Slava Gerovitch.

[7] G. Frobenius. *Über Gruppencharaktere*. Preussische Akademie der Wissenschaften Berlin: Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin. Reichsdr., 1896.

[8] Alireza S Golsefidy and Brian Longo, Toward super-approximation in positive characteristic. *J. Lond. Math. Soc. (2)*, 105(2): 1200–1261, 2022.

[9] Alireza S Golsefidy, Keivan Mallahi-Karai, and Amir Mohammadi. Locally random groups, *Michigan Math. J.*, 72 :479–527, 2020.

[10] Alireza S Golsefidy, Keivan Mallahi-Karai, and Amir Mohammadi. Spectrally independent groups, 2021. (in preparation).

[11] W. T. Gowers. Quasirandom groups. *Combin. Probab. Comput.*, 17(3):363–387, 2008.

[12] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561, 2006.

[13] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[14] Emmanuel Kowalski. Crible en expansion. *Astérisque*, 348, 2012.

[15] Emmanuel Kowalski. *An introduction to expander graphs*, volume 26 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 2019.

[16] Elon Lindenstrauss and Péter P. Varjú. Spectral gap in the group of affine transformations over prime fields. *Ann. Fac. Sci. Toulouse Math. (6)*, 25(5):969–993, 2016.

[17] Alexander Lubotzky. *Discrete groups, expanding graphs and invariant measures*, volume 125 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1994. With an appendix by Jonathan D. Rogawski.

[18] Alexander Lubotzky. Expander graphs in pure and applied mathematics. *Bull. Amer. Math. Soc. (N.S.)*, 49(1):113–162, 2012.

[19] Peter Sarnak and Xiao Xi Xue. Bounds for multiplicities of automorphic representations. *Duke Math. J.*, 64(1):207–227, 1991.

[20] Péter P. Varjú. Expansion in $\mathrm{SL}_d(\mathcal{O}_K/I)$, I square-free. *J. Eur. Math. Soc. (JEMS)*, 14(1):273–305, 2012.

[21] L. N. Vaserstein and E. Wheland. Commutators and companion matrices over rings of stable rank 1. *Linear Algebra Appl.*, 142:263–277, 1990.

[22] Robert A. Wilson. *The finite simple groups*, volume 251 of *Graduate Texts in Mathematics*. Springer-Verlag London, Ltd., London, 2009.

Mathematics Department, University of California, San Diego, CA 92093-0112, USA
Email address: golsefidy@ucsd.edu

Mathematics Department, University of California, San Diego, CA 92093-0112, USA
Email address: scsriniv@ucsd.edu