The purpose of this note is to prove that an ergodic, measure preserving action
of a sofic group with positive entropy is virtually free. In the amenable case, this
observation appears as a remark in the last section of [8]. Seward [6] proved a
corresponding result about the \(f\)-invariant for free groups, which is essentially a
special case of sofic entropy for actions of free groups [1]. The methods in [6] are
seemingly rather different than ours. I thank Benjy Weiss pointing out the remark
in [8], as well as sharing an old unpublished proof of the result for the amenable
case, Yair Glasner for pointing out Seward’s result, and Guy Salomon for interesting
discussions about sofic entropy.

1. Notation and definitions

To avoid some numerical constants, we use the notation

\[ F_1(x_1, \ldots, x_n) \ll F_2(x_1, \ldots, x_n) \]

to indicate that there exists an explicit constant \(K > 0\), independent of any of the
parameters \(x_1, \ldots, x_n\) so that \(F_1(x_1, \ldots, x_n) < K F_2(x_1, \ldots, x_n)\) for any choice of
\(x_1, \ldots, x_k\).

For a positive integer \(d \in \mathbb{N}\), write \([d] := \{1, \ldots, d\}\). We denote the group of
permutations of \([d]\) by \(S_d\). For a non-empty set \(C \subset [d]\), Denote by \(P_C\) the uniform
probability measure on \(C\) given by \(P_C(A) := \frac{|A|}{|C|}\) for \(A \subset C\). For \(f : [d] \to \mathbb{R}\) write

\[ E_C f := \frac{1}{|C|} \sum_{i \in C} f(i). \]

Throughout \(G\) is a countable group which acts on a standard probability space
\((X, \mathcal{B}, \mu)\) by probability preserving bijections. For a finite or countable partition
\(\alpha \subset \mathcal{B}\) of \(X\), and \(F \subset G\) we denote

\[ G \cdot \alpha := \bigvee_{g \in F} T_g^{-1} \alpha = \left\{ \bigcap_{g \in F} g^{-1}(A_g) : A_g \in \alpha \ \forall g \in F \right\} \]

For infinite \(A \subset G\) we denote \(A \cdot \alpha\) the \(\sigma\)-algebra generated by the sets

\[ \bigcup_{F \in A} F \cdot \alpha. \]

A slight abuse of notation: We will sometimes identify a partition \(\alpha\) of \(X\) with
the function \(x \mapsto \alpha(x)\) sending an element of \(X\) to it’s partition cell.

**Sofic groups:** Let \(F \subset G\) and \(\delta > 0\). A map \(\sigma : G \to S_d\) is called an \((F, \delta)\)-sofic
approximation if for any \(g, h \in F\)

\[ P_{[d]} (\{i \in [d] : \sigma_g(\sigma_h(i)) \neq \sigma_{gh}(i)\}) < \delta \]
and for any distinct \( g, h \in F \)
\[
P_d (\{i \in [d] : \sigma_g (i) = \sigma_h (i)\}) < \delta |F|^{-1}.
\]

Following Weiss \cite{Weiss} (also Gromov \cite{Gromov}, under a different name), a group \( G \) is sofic if there exists an \((F, \delta)\)-sofic approximation for any \( F \subseteq G \) and \( \delta > 0 \). A sequence \((\sigma_i : G \rightarrow S_d)_i \) is called a sofic-approximation sequence if for any \( \delta > 0 \) and \( F \subseteq G \) there exists \( i_0 \in \mathbb{N} \) such that \( \sigma_i \) is an \((F, \delta)\)-sofic approximation for all \( i > i_0 \).

From now we assume \( G \) is a sofic group and \((\sigma_i)_i \) is a fixed sofic approximation.

**Sofic entropy:**

Suppose \( \alpha \subset B \) is a finite partition of \( X \) which is dynamically generating, namely \( G \cdot \alpha = B \mod \mu \). The existence of a finite dynamically generating partition is not automatic, it’s a condition we assume to simplify the presentation.

As in \cite{Ghys}* Page 727* For \( \sigma : G \rightarrow S_d \), \( F \subseteq G \), we denote by \( \text{Map}_\mu (\alpha, F, \delta, \sigma) \) the set of maps \( \phi : \{1, \ldots, d\} \rightarrow X \) which satisfy the following:

1. \( |E_d f \circ \phi - E_\mu f| < \delta \) for any \( F \cdot \alpha \)-measurable \( f : X \rightarrow [0, 1] \).
2. \( P_d (\{i \in [d] : (F \cdot \alpha) \circ \phi \circ \sigma_g (i) \neq (F \cdot \alpha) \circ \sigma_h (i)\}) < \delta \) for all \( g \in F \).

We say that a pair \((\phi, \sigma)\) is a “sufficiently good sofic model for the action \( G \rhd (X, B, \mu) \)” if \( \sigma : G \rightarrow S_d \) is a sufficiently good sofic approximation and \( \phi \in \text{Map}_\mu (\alpha, F, \delta, \sigma) \) for sufficiently small \( \delta > 0 \) and sufficiently large \( F \subseteq G \).

Bowen \cite{Bowen} introduced a notion of “sofic entropy” for sofic groups. The sofic entropy of \((X, B, \mu, \sigma)\) with respect to a sofic approximation sequence \( \Sigma = (\sigma_i)_i \) is given by:

\[
h_\mu (X, B, \Sigma) = \inf_{F \subseteq G, \delta > 0} \limsup_{i \rightarrow \infty} \frac{1}{d_i} \log |\alpha^* \text{Map}_\mu (\alpha, F, \delta, \sigma_i)|,
\]

where \( \alpha^* \text{Map}_\mu (\alpha, F, \delta, \sigma_i) := \{\alpha \circ \phi \ : \ \phi \in \text{Map}_\mu (\alpha, F, \delta, \sigma_i)\} \)

If there exist \( F \subseteq G \) and \( \delta > 0 \) so that \( \text{Map}_\mu (\alpha, F, \delta, \sigma_i) = 0 \) for all large \( i \) we define
\[
h_{\mu} (X, B, \Sigma) = -\infty.
\]

Bowen showed in \cite{Bowen} that quantity \( h_\mu (X, B, \Sigma) \) does not depend on the choice of finite generating partition \( \alpha \). The are other ways to define sofic entropy, for instance as in \cite{Ghys}. These lead to an equivalent notion of sofic entropy in case there is a finite dynamically generating partition. Note that the finite set \( F \subseteq G \) above plays the role of two different parameters in appearing in the definition of sofic entropy in \cite{Ghys}:
It is both the set of elements \( g \) for which the action of \( \Sigma \) mimics the action of \( G \), and the collection of “observables” with respect to which \( P_{\mu} \circ \phi^{-1} \) approximates \( \mu \).

2. Positive Entropy Implies Finite Stabilizer

Let \( G = (V, E) \) be a finite directed graph. A set \( W \subseteq V \) is called \( \epsilon \)-dominating if there are at most \( \epsilon |V| \) vertices in \( V \) with no edge directed at some \( w \in W \).

**Lemma 2.1** (Small \( \epsilon \)-dominating random subsets in high degree graphs). For any \( \epsilon > 0 \) there exists \( k_0 \) sufficiently large such that for all \( k > k_0 \) and any \( M > 0 \) there exists \( N = N(k, M) \) such that any directed graph with at least \( N \) vertices and maximal in degree at most \( M \) such that all but \( \epsilon N \) vertices have out degree at least \( k \), a Bernoulli-randomly chosen set of intensity \( \frac{1}{\sqrt{k}} \) is \( 2 \epsilon \)-dominating with probability at least \( 1 - \epsilon \).
Proof. (Sketch) Let $G = (V, E)$ be a graph as above. Choose a random set $C$ by selecting each vertex independently with probability $\frac{1}{\sqrt{N}}$. For $v \in V$ let $n(v)$ be number of edges $(v, w) \in E$ with $w \in C$. This is Binomial $B(\frac{1}{\sqrt{N}}, \deg(v))$. Thus the expected number of vertices with no edge pointing at some $w \in C$ at most $(\epsilon + (1 - \frac{1}{\sqrt{N}})^k) |V| \approx \epsilon|V|$. For $v, w \in V$, the random variables $n(v)$ and $n(w)$ are independent, unless there is a common vertex $u$ which has an incoming edge from both $u$ and $w$. Because the maximal in-degree is at most $M$, each $u \in V$ can account for at most $M^2$ such pairs, so there are at most $M^2|V|$ pairs which are not independent, so using second moment estimate, Chebyshev’s inequality give that the probability that more than $\frac{M^2}{\sqrt{|V|}}$ are not covered is at most $\frac{M^2}{\sqrt{|V|}}$. □

The following simple lemma is a twist on an observation known as the “Mass Transport Principle”. In this case, as in many other applications it is almost trivial, yet surprisingly useful. See for instance [5] for a more general context of the mass transport principle:

Lemma 2.2 (Mass Transport Principle). If $\sigma : G \rightarrow S_d$ is an $(F, \delta)$-sofic approximation, $F_0 \subset F$ is a finite subset of the group $G$ and $f : |d| \rightarrow \mathbb{R}$ is an arbitrary function, we have

$$\left| \sum_{i=1}^{d} \sum_{g \in F_0} f(\sigma_g(i)) - |F_0| \sum_{j=1}^{d} f(d) \right| < \delta |F_0| \|f\|_\infty.$$

Proof. Let

$$d_j := \# \{i \in [d] : \exists g \in F_0 : \sigma_g(i) = j \}.$$

Then

$$\sum_{i=1}^{d} \sum_{g \in F_0} f(\sigma_g(i)) = \sum_{j=1}^{d} d_j f(j).$$

Also,

$$d_j := \# \{\sigma_g^{-1}(j) : g \in F_0\} \leq |F_0|.$$

Because $\sigma$ is an $(F, \delta)$-sofic approximation, it follows that

$$P_d(\{j \in [d] : d_j \neq |F_0|\}) < \delta.$$

It follows that

$$\left| \sum_{j=1}^{d} d_j f(j) - |F_0| \sum_{j=1}^{d} f(j) \right| \leq \sum_{j=1}^{d} |(|F_0| - d_j)|f(j)| \leq \delta |F_0| \|f\|_\infty.$$

□

Theorem 2.3. Suppose $G \acts \times \mathbb{B}, \mu$ is an ergodic $G$-action with positive sofic entropy. Then the stabilizer is finite $\mu$-almost-surely.

Proof. Note that by ergodicity $|\text{stab}(x)|$ is equal to a constant off a $\mu$-null set, and in particular it is either finite with probability 1 or infinite with probability 1. Suppose $\text{stab}(x)$ is infinite $\mu$-almost-surely. Let $\alpha$ be a dynamically generating partition.
Choose \( M \gg \epsilon^{-2} \). From the assumption that \( \text{stab}(x) \) is infinite \( \mu \)-almost-surely, it follows that there exists \( F_0 \subset G \) sufficiently big so that

\[
\mu \left( \left\{ x : \left| \text{stab}(x) \cap F_0 \right| < M \right\} \right) \ll \epsilon
\]

Choose \( \delta_1 \ll |F_0|^{-1} \epsilon \). From the assumption that \( \alpha \) is a dynamically generating partition it follows that we can find \( F_1 \subset G \) sufficiently big so that for any \( g \in F_0 \)

\[
\mu \left( \left\{ x : g \notin \text{stab}(x) \right. \text{ and } (F_1 \cdot \alpha)(x) = (F_1 \cdot \alpha)(g(x)) \right\} \right) \ll \delta_1.
\]

By making \( F_0 \) and then \( F_1 \) bigger, we assume without loss of generality that \( F_0 \) and \( F_1 \) are symmetric \( F_0 = F_0^{-1} \) and \( F_1 = F_1^{-1} \). We also assume \( 1 \in F_0 \subset F_1 \).

Now choose an even bigger \( F_2 \subset G \), namely it should satisfy

\[
|F_2| \gg |F_1|^4 \epsilon^{-2} \text{ and } F_1^4 \subset F_2.
\]

Finally choose \( F = F_2^3 \) and \( \delta \ll \epsilon|F_2|^{-1} \). Let \( \sigma \) be an \((F, \delta)\)-sofic approximation.

Choose a set \( \text{Map} \subset \text{Map}_\mu(\alpha, F, \delta, \sigma) \), so that

\[
\alpha^* \text{Map} = \alpha^* \text{Map}_\mu(\alpha, F, \delta, \sigma) \text{ and } |\text{Map}| = |\alpha^* \text{Map}|.
\]

Consider the graph \( G_{F_2} \) whose vertex set is \([d]\) such the edge \((i, j)\) exists if and only if \( \sigma_g(i) = j \) for some \( g \in F_2 \).

We will describe a procedure for choosing a random function \( \tau : [d] \to G \). First, choose a random subset \( C \subset [d] \) so that \( i \in C \) with probability \( \frac{1}{\sqrt{|F_2|}} \) independently for \( i = 1, \ldots, d \). Given the subset \( C \) as above, for each \( i \in [d] \), \( \tau(i) \) will be chosen uniformly from the set \( N_i := \{ g \in F_2 : \sigma_g(i) \in C \} \).

If the set \( N_i \) above is empty, \( \tau(i) \) is chosen uniformly from \( F_2 \). We denote by \( \nu \) the probability measure on the measure space on which \( \tau \) and \( C \) are defined. Also define \( \sigma_\tau : [d] \to [d] \) by

\[
\sigma_\tau(i) := \sigma_{\tau(i)}(i).
\]

Because \( \sigma : G \to S_d \) is an \((F, \delta)\)-sofic approximation, all but \( \delta d \) vertices in the graph \( G_{F_2} \) have degree at least \((1 - \delta)|F_2|\), and all vertices have in-degree at most \( |F_2| \). It follows using Lemma 2.1 that a randomly chosen set \( C \subset [d] \) as above will be \( \epsilon \)-dominating with high probability. It follows that with high probability with respect to \( \nu \) we get \( C \) and \( \tau \) so that \( P_\delta(\sigma_\tau(i) \notin C) \ll \epsilon \).

For \( \phi : [d] \to X \) and \( i \in [d] \) let

\[
\overline{\text{stab}_\phi}(i) := \{ g \in F_0 : (\alpha \circ \phi \circ \sigma_g)|_{F_1} = (\alpha \circ \phi)|_{F_1} \}
\]

The set \( \overline{\text{stab}_\phi}(i) \subset G \) should be viewed as a “guess” for elements which are in the stabilizer of \( \phi(i) \), using “local observations”.

By (2) it follows that for any \( g \in F_0 \),

\[
\mu \left( \left\{ x : (F \cdot \alpha)(x) \neq (F \cdot \alpha)(g(x)) \text{ and } (F_1 \cdot \alpha)(x) = (F_1 \cdot \alpha)(g(x)) \right\} \right) \ll \delta_1.
\]
For any $\phi \in \text{Map}$, because $(\phi, \sigma)$ is a sufficiently good sofic model for $G \subset (X, E, \mu)$, it follows that for any $g \in F_0$

(4) $P[d_i ((\phi \circ \sigma_g)|_{F_2} \neq (\phi|_{F_2} \text{ and } (\phi \circ \sigma_g)|_{F_1} = (\phi|_{F_1}) \ll \delta_1$.

Using (1), for any $\phi \in \text{Map}$ we have:

(5) $P[d_i] \left( |\text{stab}_\phi(i)| < M \right) \ll \epsilon$.

Call $i \in [d]$ good for $\phi$ if the following conditions are satisfied:

1. $(\phi \circ \sigma_g)|_{F_2} = (\phi|_{F_2}$ for all $g \in \text{stab}(i)$
2. $|\text{stab}(i)| \geq M$.
3. $\sigma_g \sigma_h(i) = \sigma_h \sigma_g(i)$ for all $g, h \in F$
4. $\alpha(g(\phi(i))) = \alpha(\phi(\sigma_g(i)))$ for all $g \in F$.

Otherwise, say that $i$ is bad for $\phi$.

Because $\sigma$ is an $(F, \delta)$-sofic approximation the third property fails only on a set of size $\ll \epsilon d$. If $\phi \in \text{Map}$ the last property also fails only in for a set of size $\ll \epsilon d$.

We denote by $\Psi_\phi : [d] \rightarrow \mathbb{R}$ the indicator function of bad points:

(6) $\Psi_\phi(i) := \begin{cases} 0 & \text{if } i \text{ is good for } \phi \\ 1 & \text{if } i \text{ is bad for } \phi \end{cases}$

From (4) and (5) it follows that for any $\phi \in \text{Map}

(7) $E[d_i] \Psi_\phi = P[d_i] [i \text{ is good for } \phi] \ll \delta_1 |F_0| + \epsilon \ll \epsilon$

Now:

$$\int \tau \sum_{\phi \in \text{Map}} E[d] \Psi_\phi \circ \sigma_\tau \nu(\tau) = \sum_{\phi \in \text{Map}} E[d] \left[ \int \tau \Psi_\phi \circ \sigma_\tau \nu(\tau) \right] =$$

$$\sum_{\phi \in \text{Map}} E[d] \left[ \frac{1}{|F_2|} \sum_{g \in F_2} \Psi_\phi \circ \sigma_g \right]$$

The last equaibility above is due to the fact that

$$f(\tau(i)) = \sum_{g \in F_2} \tau(i)g f(\sigma_g(i)),$$

and $\nu(\tau(i)) = \int_{|F_2|}$ for all $g \in F_2$, $i \in [d]$. Thus,

$$E[d] f \circ \sigma_\tau = E[d] \frac{1}{|F_2|} \sum_{g \in F_2} f \circ \sigma_g.$$  

Using the Mass Transport Principle (Lemma 2.2), we thus have:

$$\int \tau \left[ \sum_{\phi \in \text{Map}} E[d] \Psi_\phi \circ \sigma_\tau \right] \nu(\tau) \leq \sum_{\phi \in \text{Map}} (E[d] \left[ |\Psi_\phi| + \delta |F_2| \right] \ll |\text{Map}| (\epsilon + \delta |F_2|) \ll |\text{Map}| \epsilon$$

Using Markov inequality, it follows that for some explicit constant $K > 0$

$$\nu \left( \left\{ \tau : [d] \rightarrow G : \# \{ \phi \in \text{Map} : E[d] \Psi_\phi \circ \sigma_\tau < K \epsilon \} > \frac{1}{2} |\text{Map}| \right\} \right) \gg 1.$$
In particular there exists $\tau : [d] \to F_2$ and $C \subset [d]$ such that $|C| \ll \frac{d}{\sqrt{|F_2|}}$ and a set $Map_0 \subset Map$ so that $|Map_0| > \frac{1}{2}$ such that for any $\phi \in Map_0$

\begin{equation}
P_{|d|}(\sigma_{\tau}(i) \text{ is bad for } \phi) \ll \epsilon
\end{equation}

and

\begin{equation}
P_{d}(\sigma_{\tau}(i) \notin C) \ll \epsilon.
\end{equation}

We now fix $C$ and $\tau$ satisfying the above.

For $j \in [d]$ we define $\widehat{stab}_\phi(j) \in G$ as follows:

\begin{equation}
\widehat{stab}_\phi(j) := \begin{cases}
\tau^{-1}(j)g\tau(j) : g \in \widehat{stab}_\phi(\sigma_{\tau}(j)) & \text{if } \sigma_{\tau}(j) \in C \\
0 & \text{otherwise}
\end{cases}
\end{equation}

The set $\widehat{stab}_\phi(j) \subset F$ should again be viewed as a “guess” for elements which are in the stabilizer of $\phi(j)$, based on local information obtained by “sampling only around points in $C$”. Note that for any $j \in [d]$ we have

$\widehat{stab}_\phi(j) \subset F_2F_0F_2 \subset F$.

We say that $i \in [d]$ is exceptional for $\phi$ and $\tau$ if one of the following conditions hold:

1. $|\widehat{stab}_\phi(i)| < M$
2. There exists $g \in \widehat{stab}_\phi(i)$ so that $\alpha(\phi(i)) \neq \alpha(\sigma_{\tau}(i))$.

Because $|\widehat{stab}_\phi(j)| = |\widehat{stab}_\phi(\sigma_{\tau}(j))|$, if $\sigma_{\tau}(i) \in C$, the condition $|\widehat{stab}_\phi(j)| < M$ holds unless $\sigma_{\tau}(i)$ is bad or $\sigma_{\tau}(i) \notin C$, which happens with probability $\ll \epsilon$.

Suppose for some $i \in [d]$ there exists $g \in \widehat{stab}_\phi(j)$ so that $(\alpha \circ \phi)(i) \neq (\alpha \circ \phi)(\sigma_{\tau}(i))$. Note that $g = \tau^{-1}(i)g_1\tau(i)$ for some $g_1 \in \widehat{stab}_\phi(\sigma_{\tau}(i))$. Because $\sigma$ is a good sofic approximation,

$P_{d}[\sigma_{\tau^{-1}(i)g_1}\tau(i) \neq (\sigma_{\tau^{-1}(i)g_1} \circ \sigma_{\tau})(i) \forall g_1 \in F_0] \ll \epsilon$

If $\sigma_{\tau}(i)$ is good and $g_1 \in \widehat{stab}_\phi(\sigma_{\tau}(i))$, then

$(\alpha \circ \phi \circ \sigma_{\tau^{-1}(i)g_1} \circ \sigma_{\tau}(i))(i) = (\alpha \circ \phi \circ \sigma_{\tau^{-1}(i)})(i)$ and $(\alpha \circ \phi \circ \sigma_{\tau^{-1}(i)} \circ \sigma_{\tau}(i))(i) = (\alpha \circ \phi)(i)$.

We conclude that for every $\phi \in Map_0$,

$P_{d}(i \text{ is exceptional for } \phi \text{ and } \tau) \leq

\leq P_d(\sigma_{\tau}(i) \notin C) + P_d(\sigma_{\tau}(i) \text{ is bad }) + P_d[\sigma_{\tau^{-1}(i)g_1}\tau(i) \neq (\sigma_{\tau^{-1}(i)g_1} \circ \sigma_{\tau})(i) \forall g_1 \in F_0] \ll \epsilon$

Let

$Map^0_0 := \left\{ \phi|_{CF^2_1} : \phi \in Map_0 \right\}$.

$Map^0_0$ is set of functions $\psi : CF^2_1 \to X$ which are the restriction to $CF^2_1$ of some function in $Map_0$.

For $\psi \in Map^0_0$ we denote

$Map^{\psi}_0 := \left\{ \phi \in Map_0 : \phi|_{CF^2_1} = \psi \right\}$. 
Given $\phi \in \text{Map}_0$, we define another graph $G = G_\phi$, whose vertices are $[d]$ and there is an edge from $i$ to $j$ if and only if $j = \sigma_g(i)$ for some $g \in \text{stab}_\phi(i)$. Notice that the graph $G_\phi$ only depends on $\phi|_{CF^2}$, so it makes sense to write $G_\psi$ for $\psi \in \text{Map}_{0}$.

Let $C_1 \subset [d]$ be a random subset so that $i \in C_1$ with probability $\frac{1}{\sqrt{M}}$ independently for $i \in [d]$. It follows using Lemma 2.1 that with positive probability a randomly chosen set $C_1 \subset [d]$ as above is $\epsilon$-dominating in $G_\psi$ and that

$$|C_1| \ll \frac{d}{\sqrt{M}} \ll \epsilon |d|.$$  

In particular, for any $\psi \in \text{Map}_0$ there exists $C_1$ as above, which we denote by $C_\psi$.

For $\psi \in \text{Map}_0$, let

$$\kappa_\psi(i) := \max\{j \in C_\psi : (i, j) \text{ is an edge in } G_\psi\},$$

where $\kappa_\psi$ is undefined if $\{j \in C_\psi : (i, j) \text{ is an edge in } G_\psi\} = \emptyset$. Note that if $i$ is not exceptional for $\phi$ and $\tau$ then $\kappa(i)$ is defined and $\alpha \circ \phi(i) = (\alpha \circ \phi)(\kappa_\psi(i))$.

It follows that for every $\phi \in \text{Map}_0$ we have:

$$P_d(\alpha \circ \phi)(i) \neq (\alpha \circ \phi)(\kappa_\psi(i)) \text{ or } \kappa_\psi(i) \text{ is undefined } \ll \epsilon.$$

Thus any $\phi_1, \phi_2 \in \text{Map}_0$ differ only on a set of size $\ll \epsilon d$. So for some constant $K > 0$ we have for any $\psi \in \text{Map}_0$,

$$|\text{Map}_0^\psi| \ll \binom{d}{K \epsilon d} |\alpha|^{K \epsilon d},$$

so using Stirling’s approximation, for $\epsilon \in (0, \frac{1}{2})$:

$$\log |\text{Map}_0^\psi| \ll (\log |\alpha| \epsilon - \epsilon \log \epsilon) d.$$

Also note that

$$|\text{Map}_0| = \bigcup_{\psi \in \text{Map}_0^\psi} |\text{Map}_0^\psi|$$

so

$$\log |\text{Map}_0| \leq \log |\text{Map}_0^\psi| + \max_{\psi \in \text{Map}_0^\psi} \log |\text{Map}_0^\psi| \ll$$

$$\ll |CF_1| \log |\alpha| + (\log |\alpha| \epsilon - \epsilon \log \epsilon) d.$$ 

Note that

$$|CF_1^2| \leq |C||F_1|^2 \ll \frac{d}{\sqrt{|F_2|}} |F_1|^2 \ll \epsilon d$$

We conclude: For any $\epsilon > 0$, there exists $\delta > 0$ and $F \subseteq G$ so that for any $(F, \delta)$-sofic approximation $\sigma : G \to S_d$

$$\log \text{Map}_\mu(\alpha, F, \delta, \sigma) \ll (\log |\alpha| + \epsilon \log \epsilon) d,$$

So the action $G \curvearrowright (X, B, \mu)$ does not have positive sofic entropy. \qed
References

[1] L. Bowen. The ergodic theory of free group actions: entropy and the f-invariant. Groups Geom. Dyn., 4(3):419–432, 2010.
[2] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. J. Amer. Math. Soc., 23(1):217–245, 2010.
[3] M. Gromov. Endomorphisms of symbolic algebraic varieties. J. Eur. Math. Soc. (JEMS), 1(2):109–197, 1999.
[4] D. Kerr and H. Li. Soficity, amenability, and dynamical entropy. Amer. J. Math., 135(3):721–761, 2013.
[5] G. Last and H. Thorisson. Invariant transports of stationary random measures and mass-stationarity. Ann. Probab., 37(2):790–813, 2009.
[6] B. Seward. Finite entropy actions of free groups, rigidity of stabilizers, and a howe-moore type phenomenon, 2012.
[7] B. Weiss. Sofic groups and dynamical systems. Sankhyā Ser. A, 62(3):350–359, 2000. Ergodic theory and harmonic analysis (Mumbai, 1999).
[8] B. Weiss. Actions of amenable groups. In Topics in dynamics and ergodic theory, volume 310 of London Math. Soc. Lecture Note Ser., pages 226–262. Cambridge Univ. Press, Cambridge, 2003.