Convergence of Time-Inhomogeneous Random Walks on Finite Groups with Applications to Universality for Random Groups

Thesis by

Elia Gorokhovsky

In Partial Fulfillment of the Requirements for the
Degree of
Bachelor of Science

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended June 14, 2023

ACKNOWLEDGEMENTS

This thesis extends work done during the summer of 2022, supported by the Samuel P. and Frances Krown SURF Fellowship and the Caltech Summer Undergraduate Research Fellowship program. It was completed as a senior thesis supervised by Omer Tamuz.

I thank Melanie Matchett Wood and Omer Tamuz for direct mentorship and Alexander Gorokhovsky, Seth Berman, and Sandra O'Neill for insightful conversations. I also thank Will Sawin, Roger Van Peski, and Hoi Nguyen for helpful comments on an earlier draft of this manuscript.

ABSTRACT

We study time-inhomogeneous random walks on finite groups in the case where each random walk step need not be supported on a generating set of the group. When the supports of the random walk steps satisfy a natural condition involving normal subgroups of quotients of the group, we show that the random walk converges to the uniform distribution on the group, and give bounds for the convergence rate using spectral properties of the random walk steps. As applications, we prove a general universality theorem for quotients of the free group on n generators as $n \to \infty$, and another universality theorem for cokernels of random integer matrices with dependent entries.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

The work in this thesis is motivated by a question in random group theory, but is of independent interest to the study of random walks on groups.

Random walks on finite groups are well-studied in the reversible, time-homogeneous, ergodic regime, where the random walk on a group G consists of a product $X_1X_2 \dots X_n$ for i.i.d. X_i drawn from a distribution supported on a generating set of G. Such random walks are known to converge to the uniform distribution π on G exponentially quickly. Namely, if we denote by ν_n the distribution of $X_1X_2 \dots X_n$, then

$$d_{TV}(\nu_n, \pi) \leq \sigma^n,$$

where σ is the second-largest singular value of the Markov operator of the random walk and d_{TV} denotes the total variation distance. See [Sal04] for an excellent review of these kinds of walks.

Some of the niceness assumptions can also be relaxed; for instance, Saloff-Coste and Zúñiga [SZ07] studied convergence of time-inhomogeneous Markov chains, including random walks on finite groups, in the case where each step of the random walk is irreducible. In that case, if we denote by σ_i the second-highest singular value of the ith step,

$$d_{TV}(\nu_n, \pi) \leq \prod_{i=1}^{n} \sigma^i.$$

The main result of this paper is the following theorem, which extends part of [SZ07, Theorem 3.5] to some time-inhomogeneous random walks where the measures driving each step need not be irreducible:

Theorem 1.1. *Let G be a finite group, and let $\mu_1, \mu_2, \dots, \mu_n$ be probability measures on G. For each subgroup H of G, let $I_H = \{i \mid H = \langle \operatorname{supp} \mu_i \rangle\}$. Let $\mathcal{S}$ be a finite set of normal subgroups such that I_H is nonempty for each $H \in \mathcal{S}$. Write $\nu_n = \mu_1 * \cdots * \mu_n$.*

Also, for each i, let σ_i be the second-largest singular value of $*\mu_i$ as an operator on $L^2(\langle \operatorname{supp} \mu_i \rangle)$. Let π be the uniform distribution on G. Then if $G = \langle \bigcup_{H \in \mathcal{S}} H \rangle$, *there are constants $c_H > 0$ depending only on G and $H \in \mathcal{S}$ such that*

$$d_{TV}(\nu_n, \pi) \leq \sum_{H \in \mathcal{S}} c_H \left(\prod_{i \in I_H} \sigma_i \right).$$

We prove a more general version of this result in Theorem 2.1.

In particular, if a time-inhomogeneous random walk on a finite group has steps supported on enough normal subgroups, then it converges to the uniform distribution on the group with an exponential rate controlled by subgroups that appear infrequently or mix very slowly. Adding more probability measures to the convolution ν_n may not improve the convergence rate, but it never makes the bound worse because convolution by a probability measure is non-expansive in the L^2 norm. A nice consequence of this is that $H_1, \dots, H_k$ need not be an exhaustive list of every normal subgroup for which I_H is nonempty. We take advantage of this fact in the proof of Theorem 1.3, which is an application of this result.

The conditions of Theorem 1.1 can be weakened so that not all the subgroups H_i need to be normal (see Theorem 2.5), but see Example 2.6 for why some hypothesis on the subgroups is necessary. Theorem 2.5 also gives quantitative bounds on the constants.

Our main interest in developing this theorem is an application to the theory of random groups. In the paper [LW20], Liu and Wood studied a class of Borel probability measures μ_u on the set of isomorphism classes of profinite groups satisfying certain finiteness conditions. Namely, μ_u is the limit as n goes to infinity of the quotient of the free profinite group $\hat{F}_n$ by $n + u$ random elements drawn from the Haar measure on $\hat{F}_n$. The pushforward of μ_u by the abelianization map is defined for isomorphism classes B of finite abelian groups by

$$(\mu_u)_{\mathrm{ab}}(B) = \frac{1}{|B|^u| \operatorname{Aut}(B)|} \prod_{k=u+1}^{\infty} \zeta(k)^{-1},$$

where ζ denotes the Riemann zeta function.

The pushforward of μ_0 under the map taking a group to the p-Sylow subgroup of its abelianization (also known as the pro-p abelianization of μ_0) was conjectured by Cohen and Lenstra [CL83] as a heuristic to describe the distribution of p-Sylow subgroups of class groups of random imaginary quadratic number fields. Since then, there has been some work done on extending this heuristic to non-abelian cases, such as [BE11; BBH21; LWZ19].

In the case of the abelianization of μ_u, Wood [Woo19] showed that the distribution in fact arises as a limiting distribution for sequences of cokernels of random matrices. Indeed, let $(M_n)_{n=1}^{\infty}$ be a sequence with each M_n a random $n \times (n+u)$ integer matrix with independent entries. Wood showed that, under very weak conditions on the distributions of the entries of the M_n, the distribution of the random group $\mathbb{Z}^n / M_n(\mathbb{Z}^{n+u})$ converges weakly as $n \to \infty$ to the abelianization of μ_u for $u \geq 0$. Nguyen and Wood [NW22] extended this to allow the conditions on the entries to weaken with n. They also showed that, under the assumption of independent and identically distributed entries, there is a stronger form of convergence to μ_u. The phenomenon that the limiting distribution of $\mathbb{Z}^n / M_n(\mathbb{Z}^{n+u})$ is rather insensitive to the distributions of the entries of M_n is an example of *universality*.

Thus, Liu and Wood [LW20] ask whether the distributions of non-abelian groups μ_u themselves or some pushforwards of them may be universal for large classes of sequences of random groups, analogously to the universality result in [Woo19]. In this paper we establish a universality class for μ_u:

Theorem 1.2. *Let $u \in \mathbb{Z}$. For $n = 1, 2, \ldots$, let F_n be the free group on n generators and let ν_n be a symmetric probability measure supported on the generators of F_n. Suppose $\nu_n(x) \geq \varepsilon / n$ for each generator x of F_n for all n, and let ℓ_n be a sequence of integers such that $\frac{\ell_n}{n \log n} \to \infty$ as $n \to \infty$. For $i = 1, 2, \ldots$, let $X_{n,i}$ be an independent copy of ℓ_n steps of the ν_n-random walk on F_n. Let H_n be the normal subgroup of F_n generated by $X_{n,1}, \ldots, X_{n,n+u}$. Then the distribution of F_n / H_n converges weakly to μ_u.*

Here, weak convergence refers to weak convergence in the topology on a set of isomorphism classes of nice enough profinite groups defined in [LW20, Section 3]. We in fact prove a more general version of this theorem below (Theorem 3.1).

Theorem 1.2 relies on the observation that as $\ell \to \infty$, the result of ℓ steps of a random walk on a profinite group becomes Haar equidistributed. In fact, Liu and Wood [LW20] showed that the random quotients of $\hat{F}_n$ in the definition of μ_u can be replaced by quotients of the free group F_n by $n + u$ ℓ-step simple random walks on F_n, if ℓ is taken to infinity first and then n. More generally, the fact that random walks on finite groups become uniformly distributed is the key fact that allows us to prove universality results about μ_u and its pushforwards. In Section 3.2, we use the random walk results from Chapter 2 to extend the result of [Woo19] to cokernels of random matrices with dependent entries.

In her 2022 ICM talk, Wood [Woo23, Open Problem 3.10] asks if the universality class of μ_u can be extended to cokernels of matrices with some dependent entries. There are a few specific results in this direction. Most recently, Nguyen and Wood [NW22, Theorem 1.1] show that the distribution μ_1 is universal for Laplacians of Erdős-Rényi random directed graphs. Friedman and Washington [FW89] showed that the cokernels of the random matrices $I - M$, where M is drawn at random from the multiplicative Haar measure on $\mathrm{GL}_{2g}(\mathbb{Z}_p)$, approach μ_0 as $g \to \infty$. We are not aware of any existing results showing universality of μ_u for cokernels of broader classes of random matrices with dependent columns. However, Wood showed in [Woo14] that cokernels of random *symmetric* matrices also exhibit a different universal limiting distribution under weak constraints. Thus, an open question is: exactly how much and what kind of dependence is allowed among the entries of random matrices before they are no longer in the universality class of μ_u?

The main application of Theorem 1.1 in this paper is a generalization of the following result, which extends the result of [Woo19] to matrices with some dependence in their rows and columns. We introduce a regularity condition on matrices, (w, h, ε)-balanced. Generally, it means that the matrix can be written as a block matrix where the blocks have height at most h, width at most w, are all independent, and each satisfy some regularity condition depending on ε. The key detail is that the blocks of the matrix may have dependent entries, as long as there is no dependence between blocks. With this condition, we have:

Theorem 1.3. *Let $u \geq 0$ be an integer. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let $(w_n)_n, (h_n)_n$ be sequences of real numbers such that $w_n = o(\log n)$, $h_n = O(n^{1-\alpha})$, and $\varepsilon_n \geq n^{-\beta}$ for some $0 < \alpha \leq 1$ and $0 < \beta < \alpha/2$.*

For each integer $n \geq 0$, let M_n be an $(w_n, h_n, \varepsilon_n)$-balanced $n \times (n + u)$ random matrix with entries in $\mathbb{Z}$. Then the distribution of $\mathrm{coker}(M_n)$ *converges weakly to μ_u as $n \to \infty$.*

1.1 Notation and Terminology

For a finite set S and $p > 0$, we use $L^p(S)$ to denote the space of signed measures (equivalently, functions) on S, equipped with the norm $||f||_{L^p}^p = \sum_{s \in S} |f(s)|^p$. For a point $f \in L^p(S)$ and a compact set $K \subset L^p(S)$, we write $d_{L^p}(f, K) = \inf_{g \in K} ||f - g||_{L^p}$. We denote by $d_{TV}(\mu, \nu)$ the total variation distance between probability measures μ, ν on S, given by $d_{TV}(\mu, \nu) = \max_{T \subseteq S} |\mu(T) - \nu(T)|$. Any measure μ defines a linear *convolution operator* $*\mu$ on $L^p(S)$ given by $\nu \mapsto \nu * \mu$.

For two finite or profinite groups G, G', we write $\mathrm{Hom}(G, G')$ for the set of (continuous) group homomorphisms from G to G' and $\mathrm{Sur}(G, G')$ for the set of (continuous) surjective group homomorphisms from G to G'. For a subset $S \subseteq G$, we denote by $\langle S \rangle$ the (closed) subgroup of G generated by S and by $\langle S \rangle^G$ the (closed) normal subgroup of G generated by S.

We use $\mathbb{P}[\cdot]$ for probability and $\mathbb{E}[\cdot]$ for expectation. We denote by $\operatorname{supp} \mu$ the support of a measure μ.

If a random variable X has law μ, we write $X \sim \mu$.

Chapter 2

RANDOM WALKS

This chapter is devoted to proving a general result about equidistribution of time-inhomogeneous random walks on finite groups. Section 2.1 states a simpler version of the result and gives a geometric proof for it using contractions in Euclidean space. Section 2.2 gives a stronger result (Theorem 2.5) with a less intuitive proof which relates L^2 distance of measures on a group to the L^2 distances of their restrictions to cosets in the group. We also introduce the machinery of quotient sequences, which simplifies the process of applying Theorem 2.5.

2.1 Geometric Proof

In this subsection we prove the following theorem:

Theorem 2.1. *Let G be a finite group, and let $\mu_1, \mu_2, \ldots, \mu_n$ be probability measures on G. For each subgroup H of G, let $I_H = \{i \mid H = \langle \operatorname{supp} \mu_i \rangle\}$. Let $H_1, \ldots, H_k$ be normal subgroups with $G = \left\langle \bigcup_{j=1}^k H_j \right\rangle$. Write $\nu_n = \mu_1 * \cdots * \mu_n$. Recall that π is the uniform distribution on G.*

Also, for each i, let σ_i be the second-largest singular value of $*\mu_i$ as an operator on $L^2(\langle \operatorname{supp} \mu_i \rangle)$. Then there are constants $c_j > 0$ depending only on G and $H_1, \ldots, H_k$ such that

$$||\nu_n - \pi||_{L^2} \leq \sum_{j=1}^{k} c_j \left(\prod_{i \in I_{H_j}} \sigma_i \right).$$

This result is also implied by Theorem 2.5, but the proof is instructive due to the geometric intuition and some intermediate results also used in the proof of Theorem 2.5. Throughout this subsection, fix the finite group G, subgroups $H_1, \ldots, H_k$, and measures $\mu_1, \ldots, \mu_n$.

Consider the space $\mathcal{M} = L^2(G)$ of $\mathbb{R}$-valued functions on G with the Euclidean norm (think about these as signed measures). Since G is finite, $\mathcal{M} \cong \mathbb{R}^G$. Let $\mathcal{M}_0 = \{\nu \in \mathcal{M} \mid \nu(G) = 0\}$. For each subgroup $H \leq G$, let $\mathcal{M}_H \subseteq \mathcal{M}$ be the space of functions on $\mathcal{M}$ which are constant on each left coset of H (i.e., for $\nu \in \mathcal{M}_H$ and $g_1, g_2 \in G$ with $g_1^{-1} g_2 \in H$, $\nu(g_1) = \nu(g_2)$).

Let $\mathcal{P} \subseteq \mathcal{M}$ be the set of signed measures ν on G with $\nu(G) = 1$ and $\mathcal{P}_H = \mathcal{P} \cap \mathcal{M}_H$ for $H \leq G$. Note that $\mathcal{M}_0 = \mathcal{P} - \mathcal{P} := \{\nu - \nu' \mid \nu, \nu' \in \mathcal{P}\}$. Also, note that if $H \trianglelefteq G$, then the normed space $\mathcal{M}_H$ is canonically isomorphic to $L^2(G/H)$ (by sending $\nu \in \mathcal{M}_H$ to the function sending gH to $\nu(gH)$), and $\mathcal{P}_H$ maps to the set of signed measures on G/H with total mass 1 under this isomorphism. Finally, note that $\mathcal{P}_G = \{\pi\}$.

Any measure μ_i on G acts by on $\mathcal{M}$ by convolution on the right. If μ_i is a probability measure, the convolution operator $M_i(\nu) = \nu * \mu_i$ also fixes $\mathcal{P}$.

Say $i \in I_{H_j}$. It is well-known that if supp μ_i is not contained in a coset of a normal subgroup of H_j, then the μ_i-random walk on H_j converges to the uniform distribution on H_j (see, for example, [Sal04]). In general, M_i induces a contraction on $L^2(H_j)$ when $L^2(H_j)$ is viewed as a subspace of $\mathcal{M}$, i.e., $||M_i\nu - \nu'||_{L^2} \leq ||\nu - \nu'||_{L^2}$ for all $\nu, \nu' \in L^2(H_j)$. In addition, for any measures ν, ν' on H_j with $\nu(H_j) = \nu'(H_j)$ we have

$$||\nu * (\mu_i|_{L^2(H_j)}) - \nu' * (\mu_i|_{L^2(H_j)})||_{L^2(H_j)} \leq \sigma_i ||\nu - \nu'||_{L^2(H_j)}. \tag{2.1}$$

The reason for this inequality is as follows. Convolution with any probability measure on H_j fixes constant measures on H_j. Hence, the largest eigenvalue of the convolution operator $M_i|_{L^2(H_j)}$ is 1, corresponding to the subspace of constant measures on H_i. Moreover, $(M_i|_{L^2(H_j)})^*$ also acts by convolution with a probability measure ($\check{\mu}_i$, given by $\check{\mu}_i(g) = \mu_i(g^{-1})$), so its largest eigenvalue is also 1, corresponding to the same eigenspace. Hence, the largest singular value of $M_i|_{L^2(H_j)}$ is 1. The orthogonal complement of the subspace of constant measures on H_j is $L^2(H_j) \cap \mathcal{M}_0$, the subspace of signed measures on H_j with zero total mass. Therefore, the operator norm (equivalently, the largest singular value) of the restriction $M_i|_{L^2(H_j) \cap \mathcal{M}_0}$ is σ_i.

In particular, (2.1) holds when replacing H_j by any left coset of H_j, since multiplication on the right by a random element of H_j fixes left cosets.

On G, the operator M_i does not contract the distance between a probability measure and π, but it does contract the distance between a measure and the subspace $\mathcal{M}_{H_j}$. The idea of the proof of Theorem 2.1 is that doing this for many different subspaces $\mathcal{M}_{H_j}$ can give convergence to π.

To do this, we will need three short lemmas:

Lemma 2.2. *Let V be a finite-dimensional normed space (with norm $|\cdot|$) and $W = \{v \in V \mid \ell_1(v) = \ell_2(v) = \cdots = \ell_k(v) = 0\}$ a vector subspace of V cut out by k linear functions $\ell_i \colon V \to \mathbb{R}$. Define $d(x, W) = \inf_{y \in W} |x - y|$. Then there exist constants $c_1, c_2 > 0$ such that for all $x \in V$,*

$$c_1 \sum_{i=1}^{k} |\ell_i(x)| \leq d(x, W) \leq c_2 \sum_{i=1}^{k} |\ell_i(x)|.$$

Proof. The map $L = (\ell_1, \ldots, \ell_k)$ is a linear map $V \to \mathbb{R}^k$ with $\ker L = W$. By the first isomorphism theorem for vector spaces, there is a vector space isomorphism $\hat{L} \colon V/W \cong L(V)$. The L^1 norm on $L(V)$ induces a norm $||\cdot||_{\hat{L}}$ on V/W given by $||v||_{\hat{L}} = ||\hat{L}(v)||_{L^1} = \sum_{i=1}^{k} |\ell_i(v)|$. Also, V/W inherits a normed space structure from V with norm $||v||_V = d(v, W)$. Since V/W is finite-dimensional, the norms $||\cdot||_{\hat{L}}$ and $||\cdot||_V$ are equivalent. The result follows. □

Lemma 2.3. *Let V be a finite-dimensional normed space as above and $W_1, \ldots, W_r$ be subspaces of V with $W_1 \cap \cdots \cap W_r \neq \emptyset$. For $i = 1, \ldots, r$, say $W_i = \{v \in V \mid \ell_{i,1} = \cdots = \ell_{i,k_i} = 0\}$ for linear functions $\ell_{i,j} \colon V \to \mathbb{R}$. Then there exist constants $c_1, \ldots, c_r > 0$ such that for all $x \in V$,*

$$d(x, W_1 \cap \cdots \cap W_r) \leq c_1 d(x, W_1) + \cdots + c_r d(x, W_r).$$

Proof. The subspace $W_1 \cap \cdots \cap W_r$ is cut out by all of the functions $\ell_{i,j} \colon V \to \mathbb{R}$. Hence, by 2.2 there is a constant $c_0' > 0$ such that for all $x \in V$,

$$d(x, W_1 \cap \cdots \cap W_r) \leq c_0' \sum_{i=1}^{r} \sum_{j=1}^{k_i} |\ell_{i,j}(x)|.$$

Similarly, for each $i = 1, \ldots, r$, there is a constant $c_i' > 0$ such that for all $x \in V$,

$$d(x, W_i) \geq c_i' \sum_{j=1}^{k_i} |\ell_{i,j}(x)|.$$

Combining these results gives us

$$d(x, W_1 \cap \cdots \cap W_r) \leq c_0' \sum_{i=1}^{r} \frac{d(x, W_i)}{c_i'},$$

which is the desired result with $c_i = \frac{c_0'}{c_i'}$. □

Lemma 2.4. *Let G be a finite group and $H \leq G$ be a subgroup. Let $\nu \in \mathcal{M}$. Let $\tilde{\nu} \in \mathcal{M}_H$ be the measure on G given by $\tilde{\nu}(gh) = \frac{\nu(gH)}{|H|}$ for $h \in H$. Then $d_{L^2}(\nu, \mathcal{M}_H) = ||\nu - \tilde{\nu}||_{L^2}$.*

Proof. We want to show that $\tilde{\nu}$ minimizes the distance $||\nu - \tilde{\nu}||_{L^2}^2$ among measures in $\mathcal{M}_H$. Identifying $\mathcal{M}_H$ with $L^2(G/H)$, we have

$$d_{L^2}(\nu, \mathcal{M}_H)^2 = \inf_{\mu \in L^2(G/H)} \sum_{g \in G/H} \sum_{h \in H} (\nu(gh) - \mu(g))^2.$$

Since the values $\mu(g)$ can be chosen independently, this is equivalent to minimizing each term $\sum_{h \in gH} (\nu(gh) - \mu(g))^2$ individually:

$$d_{L^2}(\nu, \mathcal{M}_H)^2 = \sum_{g \in G/H} \inf_{\mu \in L^2(G/H)} \sum_{h \in H} (\nu(gh) - \mu(g))^2.$$

Since the mean of a finite collection of real numbers (here, $\{\nu(gh) \mid h \in H\}$) minimizes the sum of squared deviations, the infimum above is attained when $\mu(g) = \frac{\sum_{h \in H} \nu(gh)}{|H|} = \frac{\nu(gH)}{|H|} = \tilde{\nu}(gh)$.

In other words,

$$d_{L^2}(\nu, \mathcal{M}_H)^2 = \sum_{g \in G/H} \sum_{h \in H} (\nu(gh) - \tilde{\nu}(gh))^2 = ||\nu - \tilde{\nu}||_{L^2}^2.$$

□

Now we can combine these lemmas with (2.1) to prove Theorem 2.1.

Proof of Theorem 2.1. We bound the L^2 distance between the convolution ν_n and π by bounding the L^2 distance between ν_n and $\text{span}\{\pi\} = \mathcal{M}_G$ in $\mathcal{M}$. Note that $\mathcal{M}_0 \perp \mathcal{M}_G$; in particular, the affine space $\mathcal{P}$ is orthogonal to $\mathcal{M}_G$. Hence, for $\nu_n \in \mathcal{P}$, $||\nu_n - \pi||_{L^2} = d_{L^2}(\nu, \mathcal{M}_G)$. We will use Lemma 2.3 to bound $d_{L^2}(\nu, \mathcal{M}_G)$.

First, we claim that $\mathcal{M}_G = \bigcap_{j=1}^k \mathcal{M}_{H_j}$. Indeed, clearly $\mathcal{M}_G \subseteq \mathcal{M}_H$ for any subgroup $H \leq G$. On the other hand, suppose $\nu \in \bigcap_{j=1}^k \mathcal{M}_{H_j}$. Fix $g \in G$. Since $G = \langle \bigcup_{j=1}^k H_j \rangle$, g can be written in the form $g = h_1 \dots h_r$, so that for each i there is a $j_i \in \{1, \dots, k\}$ with $h_i \in H_{j_i}$. For $i = 1, \dots, r$, since $e \in H_{j_i}$, both e and h_i are in H_{j_i}, so $h_1 \dots h_{i-1}$ and $h_1 \dots h_i$ are in the same left coset of H_{j_i}, $h_1 \dots h_{i-1} H_{j_i}$. Since ν is constant on left cosets of H_{j_i}, $\nu(h_1 \dots h_{i-1}) = \nu(h_1 \dots h_i)$. Applying this argument inductively yields $\nu(e) = \nu(h_1 \dots h_k) = \nu(g)$. Since this holds for all $g \in G$, $\nu \in \mathcal{M}_G$. Hence, $\mathcal{M}_G \supseteq \bigcap_{j=1}^k \mathcal{M}_{H_k}$ as desired.

Now Lemma 2.3 tells us that there are constants $c'_1, \dots, c'_k$ such that

$$d_{L^2}(\nu_n, \mathcal{M}_G) \leq \sum_{j=1}^{k} c'_j d_{L^2}(\nu_n, \mathcal{M}_{H_j}).$$

So, we just need to bound the distance $d_{L^2}(\nu_n, \mathcal{M}_{H_j})$ for each j. Fix some $j \in \{1, \dots, k\}$. We proceed by induction on n.

First, if δ_0 is the Dirac measure on the identity of G, we have $d_{L^2}(\delta_0, \mathcal{M}_{H_j}) \leq 1 - \frac{1}{|G|} = ||\delta_0 - \pi||_{L^2}$.

Now suppose the claim holds for some fixed n. Let $\tilde{\nu}_n$ be the projection of ν_n onto $\mathcal{M}_{H_j}$, so that $||\nu_n - \tilde{\nu}_n||_{L^2} = d_{L^2}(\nu_n, \mathcal{M}_{H_j})$.

There are two cases:

Suppose $\langle \operatorname{supp} \mu_{n+1} \rangle \neq H_j$, so $n + 1 \notin I_{H_j}$. Since the transition matrix of a Markov chain is an L^2 contraction, $||\nu_{n+1} - \tilde{\nu}_n * \mu_{n+1}||_{L^2} \leq ||\nu_n - \tilde{\nu}_n||_{L^2}$.

Now note that since H_j is normal, its collections of left cosets and of right cosets coincide, so $\tilde{\nu}_n$ is constant on each right coset of H_j as well. Since multiplication by a random element of G permutes the right cosets of H_j, $\tilde{\nu}_n * \mu_{n+1}$ remains constant on right cosets of H_j, hence on left cosets. So $\tilde{\nu}_n * \mu_{n+1} \in \mathcal{M}_{H_j}$, and

$$d_{L^2}(\nu_{n+1}, \mathcal{M}_{H_j}) \leq ||\nu_{n+1} - \tilde{\nu}_n * \mu_{n+1}||_{L^2} \leq ||\nu_n - \tilde{\nu}_n||_{L^2} = d_{L^2}(\nu_n, \mathcal{M}_{H_j}).$$

Now suppose $\langle \operatorname{supp} \mu_{n+1} \rangle = H_j$, so $n + 1 \in I_{H_j}$.

By Lemma 2.4, $\tilde{\nu}_n|_{gH_j}$ is uniform on gH_j with total mass $\nu_n(gH_j)$. In particular, $\tilde{\nu}_n|_{gH_j} * \mu_{n+1}|_{gH_j} = \tilde{\nu}_n|_{gH_j}$. By ($P(r)$), this means

$$||\nu_{n+1}|_{L^2(gH_j)} - \tilde{\nu}_n|_{L^2(gH_j)}||_{L^2(gH_j)} \leq \sigma_{n+1} ||\nu_n|_{L^2(gH_j)} - \tilde{\nu}_n|_{L^2(gH_j)}||_{L^2(gH_j)}$$

Note that for any $\nu \in \mathcal{M}$, $||\nu||_{L^2(G)} = \sum_{g \in G/H_j} ||\nu|_{L^2(gH_j)}||_{L^2(gH_j)}$. So,

$$||\nu_{n+1} - \tilde{\nu}_n||_{L^2(G)} \leq \sigma_{n+1} ||\nu_n - \tilde{\nu}_n||_{L^2(G)}.$$

Hence,

$$d_{L^2}(\nu_{n+1}, \mathcal{M}_{H_j}) \leq ||\nu_{n+1} - \tilde{\nu}_n||_{L^2(G)} \leq \sigma_{n+1} ||\nu_n - \tilde{\nu}_n||_{L^2(G)} = \sigma_{n+1} d_{L^2}(\nu_n, \mathcal{M}_{H_j}).$$

By induction, we get

$$d_{L^2}(\nu_n, \mathcal{M}_{H_j}) \leq \left(1 - \frac{1}{|G|}\right) \prod_{i \in I_{H_j}} \sigma_i.$$

Combining these results for all j yields

$$||\nu_n - \pi||_{L^2} = d_{L^2}(\nu_n, \mathcal{M}_G) \leq \sum_{j=1}^{k} c'_j \left(\prod_{i \in I_{H_j}} \sigma_i \right).$$

□

2.2 Algebraic Proof

This subsection is devoted to proving a stronger version of Theorem 2.1:

Theorem 2.5. *Let G be a finite group and suppose we have a sequence of groups $H_1, \dots, H_k, G_1, \dots, G_k$ and morphisms $Q_1, \dots, Q_k, \tilde{Q}_1, \dots, \tilde{Q}_k$ as follows:*

- *We have $H_1 \trianglelefteq G$. Define $G_1 = G/H_1$. The map Q_1 is the canonical projection $G \twoheadrightarrow G_1$.*
- *For $j > 1$, we have $H_j \trianglelefteq G_{j-1}$. Define $G_j = G_{j-1}/H_j$. The map Q_j is the canonical projection $G_{j-1} \twoheadrightarrow G_j$.*
- *For $j \geq 1$, the map $\tilde{Q}_j : G \to G_j$ is the composition $Q_j \circ Q_{j-1} \circ \cdots \circ Q_1$.*

Let $\mu_1, \dots, \mu_n$ be probability measures on G. Let $\nu_n = \mu_1 * \cdots * \mu_n$. For each $j = 1, \dots, k$, let $I_j = \{i \mid \langle \operatorname{supp}(\tilde{Q}_{j-1})_* \mu_i \rangle = H_j\}$. Let π be the uniform distribution on G.

For $i \in I_j$, let σ_i be the second largest singular value of the $(\tilde{Q}_{j-1})_* \mu_i$-random walk on H_j. Then if $G_k = \{e\}$ and each I_j is nonempty, we have

$$||\nu_n - \pi||_{L^2}^2 \leq \sum_{j=1}^{k} \frac{|G_j| - 1}{|G|} \left(\prod_{i \in I_j} \sigma_i^2 \right) = \sum_{j=1}^{k} \frac{\prod_{i=j}^{k} |H_i| - 1}{|G|} \left(\prod_{i \in I_j} \sigma_i^2 \right).$$

In the case where $k = 1$ and $H_1 = G$, we recover the first part of [SZ07, Theorem 3.5]. It is not possible to fully remove the normality assumption, as the following example shows:

Example 2.6. *Consider the alternating group A_5. Recall that A_5 is generated by the 3-cycles $(1\ 2\ 3), (1\ 2\ 4), (1\ 2\ 5)$. Consider the following three-step time-inhomogeneous "random walk" on A_5: X_1 is uniformly distributed on $\langle (1\ 2\ 3) \rangle$, X_2 is uniformly distributed on $\langle (1\ 2\ 4) \rangle$, and X_3 is uniformly distributed on $\langle (1\ 2\ 5) \rangle$.*

The step distributions μ_1, μ_2, μ_3 on the respective cyclic groups all have second-largest singular value zero. However, the product $X_1X_2X_3$ is not uniformly distributed on A_5. Indeed, when $X_1X_2X_3$ acts on the tuple $(1,2,3,4,5)$, 3 can never end up in the fourth or fifth position, whereas if $X_1X_2X_3$ were uniform on A_5, 3 would end up in the fourth and fifth position with probability $1/5$ each.

The proof of Theorem 2.5 relies on the following observation that L^2 distance between a measure and the uniform measure on G can be decomposed "along" a quotient.

Lemma 2.7. *Let G be a finite group and $H \leq G$. Let π be the uniform distribution on G and let μ be any measure on G. Let $P\colon G \twoheadrightarrow G/H$ be the set map sending each element of G to the corresponding left coset of H. For each subset $S \subseteq G$, let π^μ_S be the uniform measure on S with total mass $\mu(S)$. Then*

$$||\mu - \pi||^2_{L^2(G)} = \frac{1}{|H|}||P_*\mu - P_*\pi||^2_{L^2(G/H)} + \sum_{gH\in G/H} ||\mu|_{gH} - \pi^\mu_{gH}||^2_{L^2(gH)}.$$

Proof. We have

$$\begin{aligned}
||\mu - \pi||^2_{L^2(G)} &= \sum_{g\in G} (\mu(g) - |G|^{-1})^2 \\
&= \sum_{gH\in G/H} \sum_{h\in gH} (\mu(h) - |G|^{-1})^2 \\
&= \sum_{gH\in G/H} \sum_{h\in gH} \left(\mu(h) - \frac{\mu(gH)}{|H|} + \frac{\mu(gH)}{|H|} - \frac{1}{|G|}\right)^2 \\
&= \sum_{gH\in G/H} \sum_{h\in gH} \left(\mu(h) - \frac{\mu(gH)}{|H|}\right)^2 \\
&\qquad + \sum_{gH\in G/H} \sum_{h\in gH} \left(\frac{\mu(gH)}{|H|} - \frac{1}{|G|}\right)^2 \\
&\qquad + 2 \sum_{gH\in G/H} \sum_{h\in gH} \left(\mu(h) - \frac{\mu(gH)}{|H|}\right)\left(\frac{\mu(gH)}{|H|} - \frac{1}{|G|}\right).
\end{aligned}$$

We consider each of these three sums independently. First, notice that

$$\sum_{gH\in G/H} \sum_{h\in gH} \left(\mu(h) - \frac{\mu(gH)}{|H|}\right)^2 = \sum_{gH\in G/H} ||\mu|_{gH} - \pi^\mu_{gH}||^2_{L^2(gH)}.$$

Next, we have

$$\sum_{gH\in G/H}\sum_{h\in gH}\left(\frac{\mu(gH)}{|H|}-\frac{1}{|G|}\right)^2 = \sum_{gH\in G/H}|H|\left(\frac{\mu(gH)}{|H|}-\frac{1}{|G|}\right)^2$$
$$= \frac{1}{|H|}\sum_{gH\in G/H}\left(\mu(gH)-\frac{1}{[G:H]}\right)^2$$
$$= \frac{1}{|H|}||P_*\mu - P_*\pi||^2_{L^2(G/H)}.$$

Finally,

$$2\sum_{gH\in G/H}\sum_{h\in gH}\left(\mu(h)-\frac{\mu(gH)}{|H|}\right)\left(\frac{\mu(gH)}{|H|}-\frac{1}{|G|}\right)$$
$$= 2\sum_{gH\in G/H}\left(\frac{\mu(gH)}{|H|}-\frac{1}{|G|}\right)\left(\sum_{h\in gH}\mu(h)-\frac{\mu(gH)}{|H|}\right)$$
$$= 2\sum_{gH\in G/H}\left(\frac{\mu(gH)}{|H|}-\frac{1}{|G|}\right)(\mu(gH)-\mu(gH))$$
$$= 0,$$

completing the proof. □

Combined with induction, this lemma allows us to prove the theorem:

Proof of Theorem 2.5. Let $H_0 = \{e\}$ and let $\tilde{Q}_0 = \mathrm{id}_G$.

We will prove the following statement by reverse induction on r:

$$||(\tilde{Q}_r)_*\nu_n - (\tilde{Q}_r)_*\pi||^2_{L^2(G_r)} \le \sum_{j=r+1}^{k}\frac{|G_j|-1}{|G_r|}\left(\prod_{i\in I_j}\sigma_i^2\right). \qquad (P(r))$$

When $r = k$, the right hand side of $P(r)$ is 0, and since both $(\tilde{Q}_r)_*\nu_n$ and $(\tilde{Q}_r)_*\pi$ are the unique probability measure on $G_r = \{e\}$, the left hand side is also 0, so $P(r)$ holds.

Now suppose $P(r+1)$ holds. We will show $P(r)$ holds.

Since $(\tilde{Q}_r)_*\pi$ is the uniform distribution on G_r, Lemma 2.7 applied to G_r and H_{r+1} says

$$||(\tilde{Q}_r)_*\nu_n - (\tilde{Q}_r)_*\pi||^2_{L^2(G_r)} = \frac{1}{|H_{r+1}|}||(\tilde{Q}_{r+1})_*\nu_n - (\tilde{Q}_{r+1})_*\pi||^2_{L^2(G_{r+1})}$$
$$+ \sum_{gH_{r+1}\in G_{r+1}} ||(\tilde{Q}_r)_*\nu_n|_{gH_{r+1}} - \pi^{(\tilde{Q}_r)_*\nu_n}_{gH_{r+1}}||^2_{L^2(gH_{r+1})}.$$

By the inductive hypothesis,

$$\begin{aligned}\frac{1}{|H_{r+1}|}||(\tilde{Q}_{r+1})_*\nu_n - (\tilde{Q}_{r+1})_*\pi||^2_{L^2(G_{r+1})} &\leq \sum_{j=r+2}^{k} \frac{|G_j|-1}{|G_{r+1}||H_{r+1}|}\left(\prod_{i\in I_j}\sigma_i^2\right) \\ &= \sum_{j=r+2}^{k} \frac{|G_j|-1}{|G_r|}\left(\prod_{i\in I_j}\sigma_i^2\right).\end{aligned}$$

Now consider the second term, $\sum_{gH_{r+1}\in G_{r+1}} ||(\tilde{Q}_r)_*\nu_n|_{gH_{r+1}} - \pi^{(\tilde{Q}_r)_*\nu_n}_{gH_{r+1}}||^2_{L^2(gH_{r+1})}$. By Lemma 2.4, this is precisely $d_{L^2}((\tilde{Q}_r)_*\nu_n, \mathcal{M}_{gH_{r+1}})^2$. Following the argument in the proof of Theorem 2.1, we get

$$\sum_{gH_{r+1}\in G_{r+1}} ||(\tilde{Q}_r)_*\nu_n|_{gH_{r+1}} - \pi^{(\tilde{Q}_r)_*\nu_n}_{gH_{r+1}}||^2_{L^2(gH_{r+1})} \leq \left(1-\frac{1}{|G_r|}\right)\prod_{i\in I_j}\sigma_i^2$$

Hence,

$$\begin{aligned}||(\tilde{Q}_r)_*\nu_n - (\tilde{Q}_r)_*\pi||^2_{L^2(G_r)} &\leq \sum_{j=r+2}^{k} \frac{|G_j|-1}{|G_r|}\left(\prod_{i\in I_j}\sigma_i^2\right) + \left(1-\frac{1}{|G_r|}\right)\prod_{i\in I_{r+1}}\sigma_i^2 \\ &=\leq \sum_{j=r+2}^{k} \frac{|G_j|-1}{|G_r|}\left(\prod_{i\in I_j}\sigma_i^2\right) + \left(\frac{|G_{r+1}|-1}{|G_r|}\right)\prod_{i\in I_{r+1}}\sigma_i^2 \\ &= \sum_{j=r+1}^{k} \frac{|G_j|-1}{|G_r|}\left(\prod_{i\in I_j}\sigma_i^2\right),\end{aligned}$$

completing the induction. When $r = 0$, we get

$$||\nu_n - \pi||^2_{L^2} \leq \sum_{j=1}^{k} \frac{|G_j|-1}{|G|}\left(\prod_{i\in I_j}\sigma_i^2\right).$$

□

Remark 2.8. *We can use some inequalities to get a similar nice bound on the L^1 or total variation distance.*

Jensen's inequality says that $\frac{1}{|G|^2}||\nu_n - \pi||^2_{L^1} \leq \frac{1}{|G|}||\nu_n - \pi||^2_{L^2}$, so

$$||\nu_n - \pi||^2_{L^1} \leq \sum_{j=1}^{k}(|G_j|-1)\left(\prod_{i\in I_j}\sigma_i^2\right).$$

By subadditivity of square root,

$$||\nu_n - \pi||_{L^1} \leq \sum_{j=1}^{k}\sqrt{|G_j|-1}\left(\prod_{i\in I_j}\sigma_i\right).$$

Finally, using $d_{TV}(\nu_n, \pi) = \frac{1}{2}||\nu_n - \pi||_{L^1}$, *we get*

$$d_{TV}(\nu_n, \pi) \leq \frac{1}{2} \sum_{j=1}^{k} \sqrt{|G_j| - 1} \left(\prod_{i \in I_j} \sigma_i \right).$$

Combining this remark with Theorem 2.1 yields Theorem 1.1.

Chapter 3

RANDOM GROUPS

This chapter is devoted to showing universality results for certain kinds of random groups, using the moment method of Wood. As a warm-up, Section 3.1 contains a proof that the distributions μ_u from [LW20] are universal for quotients of free groups by random walks. The main result of this chapter is the proof in Section 3.2 applying the results from Chapter 2 to extend the argument of [Woo19] to cokernels of random matrices with some dependence in the entries.

Liu and Wood [LW20, Section 3] describe a topology on the set of isomorphism classes of profinite groups satisfying a sufficiently nice property. In this section, weak convergence of a sequence of random groups always means weak convergence in this topology. We use only one fact about this topology, which is that weak convergence in the topology is equivalent to convergence of moments.

Distributions on these kinds of profinite groups can be studied by analyzing their *moments*. For a random group A and a group G, the G-*moment* of A is the expected number of surjective homomorphisms from A to G, denoted $\mathbb{E}[\#\operatorname{Sur}(A, G)]$. Sawin [Saw20] showed that, provided the G-moments of A grow at most polynomially in $|G|$, for any sequence of random profinite groups A_n, if $\lim_{n\to\infty} \mathbb{E}[\#\operatorname{Sur}(A_n, G)] \to \mathbb{E}[\#\operatorname{Sur}(A, G)]$ for an appropriate collection of finite groups G, then A_n converge weakly to A in the topology of [LW20, Section 3]. This was proved originally for abelian groups in [Woo14], and an analogous statement is true more generally for random objects in "diamond categories", i.e., categories satisfying an analogue of the diamond isomorphism theorem [SW22].

The measure μ_u defined by [LW20] has G-moment $\frac{1}{|G|^u}$ for each finite group G. In particular, the abelianization of μ_u has moments $\frac{1}{|G|^u}$ for finite abelian G. To prove that a sequence of random groups converges weakly to μ_u, it suffices to show that their moments converge to $\frac{1}{|G|^u}$.

Here is a brief outline of the proof strategy used in [Woo19; NW22], and other work to compute the moments of cokernels of random matrices, which we adapt to prove our universality results. Let V_n be a free group (or free abelian group, free nilpotent group, free profinite etc.) on n generators. Let $X_{1,n}, \ldots, X_{m,n}$ be some random elements of V_n (with m depending on n), and let $H_n = \langle X_{1,n}, \ldots, X_{m,n} \rangle^{V_n}$ be the (compact) normal subgroup generated by the $X_{i,n}$. We are interested in finding the moment at a finite group G of the random group V_n/H_n. For example, in the case of random matrix cokernels, $V_n = \mathbb{Z}^n$ and $X_{1,n}, \dots, X_{n+u,n}$ are columns of a random $n \times (n+u)$ matrix.

There is a one-to-one correspondence between surjections $V_n/H_n \to G$ and surjections $V_n \to G$ vanishing on H_n. So,

$$\begin{aligned}\mathbb{E}[\#\operatorname{Sur}(V_n/H_n, G)] &= \mathbb{E}\left[\sum_{f\in\operatorname{Sur}(V_n,G)} 1_{H_n\subseteq\ker f}\right]\\ &= \sum_{f\in\operatorname{Sur}(V_n,G)} \mathbb{E}[1_{H_n\subseteq\ker f}]\\ &= \sum_{f\in\operatorname{Sur}(V_n,G)} \mathbb{P}[H_n \subseteq \ker f].\end{aligned}$$

If the $X_{i,n}$ are independent, then

$$\mathbb{P}[H_n \subseteq \ker f] = \prod_{i=1}^{m} \mathbb{P}[X_{i,n} \in \ker f].$$

This independence is a crucial assumption of [NW22]. More generally, one can consider the induced map $f^m \colon (V_n)^m \to G^m$ and see that

$$\mathbb{P}[H_n \subseteq \ker f] = \mathbb{P}[f^m(X_{1,n}, \dots, X_{m,n}) = 0 \in G^m].$$

Thus, the problem translates to understanding random elements in the finite group G^m. There are roughly $|G|^n$ surjections $V_n \to G$, so if $f^m(X_{1,n}, \dots, X_{m,n})$ is close to uniform in G^m, then $\mathbb{P}[H_n \subseteq \ker f] \approx \frac{1}{|G|^m}$ and $\mathbb{E}[\#\operatorname{Sur}(V_n/H_n, G)] \approx \frac{1}{|G|^{m-n}}$. If $m = n+u$, then in the limit we recover exactly the moments we want for convergence to μ_u.

In the cases we consider, $|G^m|$ grows exponentially with n, so it is not obvious that $f^m(X_{1,n}, \dots, X_{m,n})$ should be anywhere close to uniform in G^m. However, when the $X_{i,n}$ are close to independent, the problem reduces to showing closeness to the uniform distribution in smaller groups, which is much more tractable.

There are two more key regularity issues that need to be addressed. The first is that the $X_{i,n}$ need to have nice enough distributions that their projections can be close to independent. For example, if $V_n = \mathbb{Z}^n$ and every component of $X_{n,i}$ is divisible by a with probability 1, then for any group G with $|G| = a$, every surjection $V_n \to G$ descends to a surjection $V_n/H_n \to G$, giving us more surjections than we want. In

[NW22], and in this work, this worst-case example is avoided by demanding that the $X_{n,i}$ satisfy some anti-concentration conditions.

There is also the problem of the map f. Most of the time (this will be formalized later), f is sufficiently regular that the regularity of the $X_{i,n}$ implies regularity of the images $f(X_{i,n})$. However, this is not always the case. In [Woo14; Woo19], Wood gets around this issue by splitting up surjections into nice "codes" and pathological non-codes, and categorizing non-codes by how far they are from being codes. In Section 3.1, the $X_{n,i}$ are sufficiently regular and this issue does not appear, but in Section 3.2, we have to extend the approach of [Woo19] to work with dependent relators.

3.1 Nonabelian Groups

This subsection is devoted to constructing a universality class for μ_u, which is given by the following theorem:

Theorem 3.1. *Let $u \in \mathbb{Z}$. For $n = 1, 2, \ldots$, Let F_n be the free group on n generators. For $i, j = 1, 2, \ldots$, let $\mu_{n,i,j}$ be probability measures on F_n with $\mu_{n,i,j}(e) > 0$ and $\langle \operatorname{supp} \mu_{n,i,j} \rangle = F_n$. Let $\varepsilon_n = \inf_{i,j>0} \min_{x \in \operatorname{supp} \mu_{n,i,j}} \mu_{n,i,j}(x)$. Assume ℓ_n are positive integers such that $\lim_{n\to\infty} \frac{\ell_n \varepsilon_n^2}{\log n} = \infty$ for all i.*

For each n, for $i = 1, \ldots, n+u$, let $X_{n,i}$ be a random element of F_n drawn from $\mu_{n,i,1} * \cdots * \mu_{n,i,\ell_{n,i}}$.

Let $A_n = F_n / \langle X_{n,1}, \ldots, X_{n,n+u} \rangle^{F_n}$. Then the distributions of the random groups A_n converge weakly to μ_u as $n \to \infty$.

Moreover, if each $\mu_{n,i,j}$ is symmetric (i.e., $\mu_{n,i,j}(x) = \mu_{n,i,j}(x^{-1})$ for all $x \in F_n$), then we only need that $\lim_{n\to\infty} \frac{\ell_n \varepsilon_n}{\log n} = \infty$.

In particular, suppose $\mu_{n,i,j}$ is independent of i, j, supported on the generators of F_n, and symmetric. Then if there is some $\varepsilon > 0$ such that $\varepsilon_n \geq \varepsilon / n$ for all n, we recover Theorem 1.2.

To prove this theorem, we will work with moments. The argument is analogous to some parts of the proof in [Woo19] that the abelianization of μ_u is universal for abelian groups. First, we make an observation about the problem of finding moments for a random quotient.

Let V be a group and G be a finite group. Let $X_1, X_2, \ldots, X_{n+u}$ be independent random elements of V, and let $H = \langle X_1, X_2, \ldots, X_{n+u}\rangle^V$. Say we want to count $\mathbb{E}(\#\operatorname{Sur}(V/H, G))$.

There is a bijection between surjections $V/H \to G$ and surjections $V \to G$ vanishing on H. So,

$$\begin{aligned}\mathbb{E}[\#\operatorname{Sur}(V/H, G)] &= \mathbb{E}\left[\sum_{f\in \operatorname{Sur}(V,G)} 1_{H\subseteq \ker F}\right]\\ &= \sum_{f\in \operatorname{Sur}(V,G)} \mathbb{E}[1_{H\subseteq \ker f}]\\ &= \sum_{f\in \operatorname{Sur}(V,G)} \mathbb{P}[H \subseteq \ker f].\end{aligned}$$

By independence,

$$\mathbb{E}[\#\operatorname{Sur}(V/H, G)] = \sum_{f\in \operatorname{Sur}(V,G)} \mathbb{P}[H \subseteq \ker f] = \sum_{f\in \operatorname{Sur}(V,G)} \prod_{i=1}^{n+u} \mathbb{P}[f(X_i) = e]. \tag{3.1}$$

This reduces the problem of computing moments to understanding the distribution of the image of one random relator at a time. Since the random relators in Theorem 3.1 come from a random walk, their images in a finite group G also come from a random walk, so should converge to a uniform distribution. The following lemma formalizes this notion.

Lemma 3.2. *Fix n. Let F_n be the free group on n generators. For $j = 1, 2, \ldots,$ let μ_j be probability measures on F_n with $\mu_j(e) > 0$. Let $\varepsilon = \inf_{j>0} \min_{x\in \operatorname{supp} \mu_j} \mu_j(x)$. Let ℓ be a positive integer.*

Let X be a random element of F_n drawn from $\mu_1 * \cdots * \mu_\ell$. Let G be a finite group and let $f\colon F_n \twoheadrightarrow G$ be a surjective homomorphism. Then

$$\left|\mathbb{P}[f(X) = e] - \frac{1}{|G|}\right| \le \exp\left(-\frac{\varepsilon^2 \ell}{2|G|^2}\right)$$

If each μ_j is symmetric, then the bound can be improved to $\exp\left(-\frac{\varepsilon\ell}{|G|^2}\right)$.

Proof. The expression $\left|\mathbb{P}[f(X) = e] - \frac{1}{|G|}\right|$ is bounded above by the L^2 distance between the distribution of $f(X)$ and the uniform distribution on G.

Applying Theorem 2.5 (equivalently, [SZ07, Theorem 3.5]) and Remark 2.8 to the pushforward $f_*(\mu_1 * \cdots * \mu_\ell)$ yields

$$\left|\mathbb{P}[f(X) = e] - \frac{1}{|G|}\right| \leq \prod_{j=1}^{\ell} \sigma_j,$$

where σ_j is the second-largest singular value of the $f_*\mu_j$-random walk on G.

To bound σ_j, let M_j be the transition matrix for the $f_*\mu_j$-random walk. Then M_j^* is the transition matrix for the time-reversed random walk driven by $(f_*\mu_j)^\vee$. So, $M_j M_j^*$ corresponds to taking one step from $f_*\mu_j$, then one step from the reversed measure.

Since $\mu_j(e) > 0$, $\operatorname{supp}(f_*\mu_j * (f_*\mu_j)^\vee) \supseteq \operatorname{supp} f_*\mu_j$. Moreover,

$$\min_{x \in \operatorname{supp}(f_*\mu_j * (f_*\mu_j)^\vee)} (f_*\mu_j * (f_*\mu_j)^\vee)(x) \geq \varepsilon^2.$$

By [Sal04, Theorem 6.2],

$$\sigma_j^2 \leq 1 - \frac{\varepsilon^2}{D^2},$$

where D is the diameter of the Cayley graph of G with respect to the generating set $\operatorname{supp}(f_*\mu_j * (f_*\mu_j)^\vee)$. In particular, $D < |G|$, so

$$\sigma_j \leq \sqrt{1 - \frac{\varepsilon^2}{|G|^2}} \leq 1 - \frac{\varepsilon^2}{2|G|^2} \leq \exp\left(-\frac{\varepsilon^2}{2|G|^2}\right).$$

The result follows.

If μ_j is symmetric, then M_j is symmetric and its eigenvalues coincide with its singular values. Hence by [Sal04, Theorem 6.2] we get

$$\sigma_j \leq 1 - \frac{\varepsilon}{D^2} \leq \exp\left(-\frac{\varepsilon}{|G|^2}\right).$$

□

To combine the error terms for each individual relator, we use a trick from [Woo19]. The following lemma is a more general restatement of [Woo19, Lemma 2.4], but the proof is essentially the same.

Lemma 3.3. *Let G be a finite group, $y_1, \ldots, y_{n+u}$ fixed elements of G, $Y_1, \ldots, Y_{n+u}$ independent random elements of G, and $d > 0$ such that for each i,*

$$\left|\mathbb{P}[Y_i = y_i] - \frac{1}{|G|}\right| \leq d \leq \frac{\log 2}{|G|(n + u - 1)}.$$

Then

$$\left|\mathbb{P}[Y_i = y_i \textit{ for all } i] - \frac{1}{|G|^{n+u}}\right| \leq \frac{2(n+u)d}{|G|^{n+u-1}}.$$

Proof. This follows directly from [Woo19, Lemma 2.3]. □

From here, the proof of Theorem 3.1 is straightforward using moments.

Proof of Theorem 3.1. We want to apply Lemma 3.12.

Since $\frac{\varepsilon_n^2 \ell_n}{\log n} \to \infty$, $\varepsilon_n^2 \ell_n - c \log n \to \infty$ for all constants $c \in \mathbb{R}$. Hence,

$$\begin{aligned}\frac{\sqrt{|G|}}{2} \exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2}\right) \cdot \frac{|G|(n+u-1)}{\log 2} &= \frac{\sqrt{|G|}}{2} \exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2} + \log(n+u-1) + \log\frac{|G|}{\log 2}\right) \\ &\to 0\end{aligned}$$

as $n \to \infty$. In particular, it is less than 1 for large enough n.

Lemma 3.2 says that for any surjection $f\colon F_n \to G$,

$$\left|\mathbb{P}[f(X_{n,i}) = e] - \frac{1}{|G|}\right| \leq \frac{\sqrt{|G|}}{2} \exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2}\right),$$

so by Lemma 3.3, for large enough n,

$$\left|\mathbb{P}[f(X_{n,i}) = e \text{ for all } 1 \leq i \leq n+u] - \frac{1}{|G|^{n+u}}\right| \leq \frac{C}{|G|^n}(n+u)\exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2}\right)$$

for some constant C depending on $|G|$ and u. In particular, since $\#\operatorname{Hom}(F_n, G) = |G|^n$,

$$\begin{aligned}&\left|\mathbb{E}[\#\operatorname{Sur}(F_n/\langle X_{n,1}, \ldots, X_{n,n+u}\rangle^{F_n}, G)] - \frac{1}{|G|^u}\right| \\ &= \left|\left(\sum_{f \in \operatorname{Sur}(F_n,G)} \mathbb{P}[f(X_{n,i}) = e \text{ for all } 1 \leq i \leq n+u]\right) - \left(\sum_{f \in \operatorname{Hom}(F_n,G)} \frac{1}{|G|^{n+u}}\right)\right| \\ &\leq \sum_{f \in \operatorname{Sur}(F_n,G)} \frac{C}{|G|^n}(n+u)\exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2}\right) + \sum_{f \in \operatorname{Hom}(F_n,G)\setminus \operatorname{Sur}(F_n,G)} \frac{1}{|G|^{n+u}} \\ &\leq C(n+u)\exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2}\right) + \sum_{f \in \operatorname{Hom}(F_n,G)\setminus \operatorname{Sur}(F_n,G)} \frac{1}{|G|^{n+u}}.\end{aligned}$$

[Woo19] showed in Theorem 2.9 that $\sum_{f \in \operatorname{Hom}(F_n,G)\setminus \operatorname{Sur}(F_n,G)} \frac{1}{|G|^{n+u}} \to 0$ as $n \to \infty$, and since $\frac{\varepsilon_n^2 \ell_n}{\log n} \to \infty$ as $n \to \infty$, $C(n+u)\exp\left(-\frac{\varepsilon_n^2 \ell_n}{2|G|^2}\right) \to 0$ as $n \to \infty$. Hence,

$$\lim_{n \to \infty} \mathbb{E}[\#\operatorname{Sur}(F_n/\langle X_{n,1}, \ldots, X_{n,n+u}\rangle^{F_n}, G) = \frac{1}{|G|^u},$$

which is the G-moment for μ_u. The theorem follows from [Saw20, Theorem 1.2].

If each $\mu_{n,i,j}$ is symmetric and $\frac{\varepsilon_n \ell_n}{\log n} \to \infty$ as $n \to \infty$, then the same result holds using the stronger bound in Lemma 3.2. □

3.2 Abelian Groups with Dependent Relations

In this section, μ_u will refer to the abelianized version for brevity. The goal of this section is to prove the following theorem, which is also Theorem 1.3:

Theorem 3.4. *Let $u \geq 0$ be an integer. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let $(w_n)_n$, $(h_n)_n$ be sequences of real numbers such that $w_n = o(\log n)$, $h_n = O(n^{1-\alpha})$, and $\varepsilon_n \geq n^{-\beta}$ for some $0 < \alpha \leq 1$ and $0 < \beta < \alpha/2$.*

For each integer $n \geq 0$, let M_n be an $(w_n, h_n, \varepsilon_n)$-balanced $n \times (n+u)$ random matrix with entries in $\mathbb{Z}$. Then the distribution of $\operatorname{coker}(M_n)$ *converges weakly to μ_u as $n \to \infty$.*

To prove convergence results, we will use the moment method of Wood (see [Woo14; Woo19]) as follows. Let $X_1, X_2, \ldots$ be a sequence of random finitely generated abelian groups and Y be a random finitely generated abelian group. Let $a > 0$ be an integer and A the set of isomorphism classes of abelian groups with exponent dividing A. If for every $G \in A$ we have

$$\lim_{n\to\infty} \mathbb{E}[\#\operatorname{Sur}(X_n, G)] = \mathbb{E}[\#\operatorname{Sur}(Y, G)] \leq |\wedge^2 G|$$

then for every $H \in A$ we have

$$\lim_{n\to\infty} \mathbb{P}[X_n \otimes \mathbb{Z}/a\mathbb{Z} \cong H] = \mathbb{P}[Y \otimes \mathbb{Z}/a\mathbb{Z} \cong H]$$

[Woo19, Theorem 3.1]. If this holds for all choices of a, it should be understood as weak convergence in the set of isomorphism classes of finitely generated abelian groups with the topology generated by the open sets $U_{a,H} = \{X \text{ finitely generated abelian } | X \otimes \mathbb{Z}/a\mathbb{Z} \cong H\}$. Moreover, note that if $Y \sim \mu_u$, then [Woo19, Lemma 3.2] gives

$$\mathbb{E}[\#\operatorname{Sur}(Y, G)] = |G|^{-u}.$$

Following this strategy, we will obtain Theorem 1.3 as a corollary of Theorem 3.18, which states that if X_n are the cokernels of $n \times (n+u)$ random matrices satisfying appropriate conditions, then $\lim_n \mathbb{E}[\#\operatorname{Sur}(X_n, G)] = |G|^{-u}$.

When X_n is the cokernel of a random matrix, the problem of counting surjections from X_n into G can be attacked with combinatorics. Say $X_n = \mathbb{Z}^n/N$, where N is a random subgroup of $\mathbb{Z}^n$. Then surjections $X_n \to G$ are in bijection with surjections $\mathbb{Z}^n \to G$ which vanish on N. It follows from linearity of expectation that

$$\mathbb{E}[\#\operatorname{Sur}(\mathbb{Z}^n/H, G)] = \sum_{f \in \operatorname{Sur}(\mathbb{Z}^n, G)} \mathbb{P}[f(N) = 0].$$

In the case of cokernels of random matrices, N is the subgroup generated by the columns of the random matrix, viewed as elements of $\mathbb{Z}^n$. Given a map $f\colon \mathbb{Z}^n \to G$, we get a map $f\colon (\mathbb{Z}^n)^m \to G^m$ applying f to each component. Then, viewing M as an element of $(\mathbb{Z}^n)^m$, we have that $f(N) = 0$ if and only if $f(M) = 0$. Thus, we want to bound the probabilities $f(M) = 0$. Past work that assumes random matrices with independent entries (e.g., [Woo19]) has observed that if Z is a random tuple in $\mathbb{Z}^n$ with independent, sufficiently regular components, then for most $f \in \operatorname{Sur}(\mathbb{Z}^n, G)$, the element $f(Z) \in G$ is close to uniformly distributed. Applying this independently to each column allows us to compute $\mathbb{P}[f(M) = 0]$. In this work, we apply the same principle to consider several columns of a random matrix at a time.

We will start by defining an appropriate notion of regularity for random matrices.

Balanced elements

The following definition captures the idea that a random element in a group is not too concentrated in a particular coset.

Definition 3.5. Let G be a group. A G-valued random variable X is *ε-balanced* if for any proper subgroup $H < G$ and element $g \in G$, we have $\mathbb{P}[X \in gH] \leq 1 - \varepsilon$.

This definition agrees with the definition for cyclic groups in [Woo19].

In this paper, we consider $n \times m$ integer matrices as elements of the abelian group $(\mathbb{Z}^n)^m$. For each subset E of $\{1, \dots, n\} \times \{1, \dots, m\}$, we have a quotient map π_S from $(\mathbb{Z}^n)^m$ onto $\mathbb{Z}^S$ given by taking the entries of a matrix indexed by pairs in S. We say that a subset of the entries of a random matrix M with indices S is (jointly) ε-balanced if $\pi_S(M)$ is ε-balanced in $\mathbb{Z}^S$.

The new definition of ε-balanced has some desirable properties that help construct new examples of ε-balanced random variables.

Lemma 3.6. *(1) If $\pi\colon G \to Q$ is a surjective homomorphism of groups and X is ε-balanced in G, then $\pi(X)$ is ε-balanced in Q.*

(2) *If G, G' are groups, X is ε-balanced in G, Y is ε-balanced in G', and X and Y are independent, then (X, Y) is ε-balanced in $G \times G'$.*

Proof. (1) Let $qK \subsetneq Q$ be a coset of a proper subgroup of Q. We have $\pi^{-1}(qK) = \bigcup_{g \in \pi^{-1}(q)} g\pi^{-1}(K)$. Since π is surjective, $\pi^{-1}(K)$ is a proper subgroup of G (or else $K = \pi(\pi^{-1}(K)) = \pi(G) = Q$). Now for any $g, g' \in \pi^{-1}(q)$, we have $\pi(g) = \pi(g')$, so $gg^{-1} \in \ker\pi \subset \pi^{-1}(K)$ and $g\pi^{-1}(K) = g'\pi^{-1}(K)$. Hence, if $g \in \pi^{-1}(q)$, we have $\pi^{-1}(qK) = g\pi^{-1}(K)$, and this is a coset of a proper subgroup of G. Since X is ε-balanced,

$$\mathbb{P}[\pi(X) \in qK] \leq \mathbb{P}[X \in \pi^{-1}(qK)] \leq 1 - \varepsilon,$$

as desired.

(2) Let kH be a coset of a proper subgroup of $G \times G'$. Note that

$$\mathbb{P}[(X,Y) \in kH] = \mathbb{P}[(X,e) \in (e,Y^{-1})kH] = \mathbb{P}[(X,e) \in (e,Y^{-1})kH \cap (G \times \{e\})].$$

Recall that the intersection of two cosets in a group is either empty or a coset of their intersection. In particular, $(e, Y^{-1})kH \cap (G \times \{e\})$ is either empty or a coset of a subgroup of $G \times \{e\}$.

Thus there are two cases:

i. If $(e, y^{-1})kH \cap (G \times \{e\}) \subsetneq G \times \{e\}$ for all $y \in G'$:
Condition on $Y = y$ for some fixed $y \in G'$. Since X and Y are independent,

$$\mathbb{P}[(X,e) \in (e,Y^{-1})kH \cap (G \times \{e\}) \mid Y = y] = \mathbb{P}[(X,e) \in (e,y^{-1})kH \cap (G \times \{e\})].$$

Since $(e, y^{-1})kH \cap (G \times \{e\}) \subsetneq G \times \{e\}$, either $(e, y^{-1})kH \cap (G \times \{e\}) = \varnothing$ or $(e, y^{-1})kH \cap (G \times \{e\})$ is a coset of a proper subgroup of $G \times \{e\}$. In the former case, $\mathbb{P}[(X, e) \in (e, y^{-1})kH \cap (G \times \{e\})] = 0$. In the latter case, notice that (X, e) is ε-balanced in $G \times \{e\}$ by (1). Hence $\mathbb{P}[(X, e) \in (e, y^{-1})kH \cap (G \times \{e\})] \leq 1 - \varepsilon$.
In both cases, $\mathbb{P}[(X, y) \in kH] = \mathbb{P}[(X, e) \in (e, y^{-1})kH \cap (G \times \{e\})] \leq 1 - \varepsilon$. Hence, we have

$$\begin{aligned}\mathbb{P}[(X,Y) \in kH] &= \sum_{y \in G'} \mathbb{P}[(X,e) \in (e,Y^{-1})kH \cap (G \times \{e\}) \mid Y = y]\mathbb{P}[Y = y] \\ &\leq (1 - \varepsilon) \sum_{y \in G'} \mathbb{P}[Y = y] \\ &= 1 - \varepsilon.\end{aligned}$$

ii. If $G \times \{e\} \subseteq (e, y^{-1})kH$ for some $y \in Y$, then $(e, e) \in (e, y^{-1})kH$, so in particular $(e, y^{-1})kH$ is a subgroup of $G \times G'$ and we must have $(e, y^{-1})kH = H$. We claim that $H = G \times H'$ for some proper subgroup H' of G'.

Indeed, let $\pi\colon G \times G' \to G'$ be the projection and let $H' = \pi(H)$. On one hand, clearly $H \subseteq G \times \pi(H)$. On the other, if $(g, h') \in G \times H'$, then $h' = \pi(g', h)$ for some $(g', h) \in H$. Then $(g, h') = (g(g')^{-1}, e)(g', h)$. Since $(g(g')^{-1}, e) \in G \times \{e\} \subseteq H$, we have $(g, h') \in H$. Hence $H = G \times H'$. Note that $H' \lneq G'$, or else $H = G \times G'$ is not a proper subgroup.

Then

$$\mathbb{P}[(X, Y) \in kH] = \mathbb{P}[Y \in H'] \leq 1 - \varepsilon.$$

Hence, in both cases we have $\mathbb{P}[(X, Y) \in kH] \leq 1 - \varepsilon$ and since this holds for every proper coset kH, we have that (X, Y) is balanced.

□

Note that Lemma 3.6 gives us a nice way to build up ε-balanced matrices. If the entries of a random matrix can be partitioned into independent subsets and each of these subsets of the entries is jointly ε-balanced, then the whole matrix is ε-balanced. For example, any matrix with independent, ε-balanced entries (as in [Woo19]) is ε-balanced as a matrix.

When a random variable is ε-balanced, we can get an upper bound on the associated singular value.

Lemma 3.7. *Suppose G is a finite group and X is ε-balanced in G with distribution μ. Let σ be the second largest singular value of the operator $*\mu$ on $L^2(G)$. Then*

$$\sigma \leq \exp\left(-\frac{\varepsilon}{2|G|^3}\right).$$

Proof. Note that σ is the square root of the second largest eigenvalue of the operator $*\nu := *\mu * \check{\mu}\colon L^2(G) \to L^2(G)$, where $*\check{\mu}$ is the adjoint to the operator $*\mu$, given by $\hat{\mu}(x) = \mu(-x)$. The operator $*\nu$ is the transition operator for a random walk on G, where each step is a difference of two independent copies of X.

In particular, note that this new random walk is time-reversible. By [Sal04, Theorem 6.2], for any symmetric generating set Σ of G, the eigenvalues of $*\mu * \check{\mu}$ are therefore bounded above by

$$\sigma^2 \leq 1 - \frac{m}{D^2},$$

where $m = \min_{x \in \Sigma} (\mu * \check{\mu})(x)$ and D is the diameter of the Cayley graph of (G, Σ). In particular, $D \leq |G|$. Since ν is symmetric, we can relax the assumption that Σ is symmetric by taking $\Sigma \cup \Sigma^{-1}$ in the theorem. In that case, m stays the same but D can only decrease.

The goal is to choose an appropriate Σ to bound m from below. Note that if X_1 and X_2 are ε-balanced, then so is $X_1 X_2^{-1}$ (via conditioning on X_2). In particular, ν is ε-balanced.

We proceed iteratively. Having chosen $x_1, \ldots, x_{n-1}$ (including the empty set $n = 1$), if $\langle x_1, \ldots, x_{n-1} \rangle = G$ then we are done. Otherwise, since ν is ε-balanced, $\nu(\langle x_1, \ldots, x_{n-1} \rangle) \leq 1 - \varepsilon$. Choose

$$x_n = \operatorname{argmax}_{x \in G \setminus \langle x_1, \ldots, x_{n-1} \rangle} \nu(x).$$

Since $\nu(\langle x_1, \ldots, x_{n-1} \rangle) \leq 1 - \varepsilon$, we have $\nu(G \setminus \langle x_1, \ldots, x_{n-1} \rangle) \geq \varepsilon$, so $\nu(x_n) \geq \frac{\varepsilon}{|G \setminus \langle x_1, \ldots, x_{n-1} \rangle|} \geq \frac{\varepsilon}{|G|}$.

Hence we have $m \geq \frac{\varepsilon}{|G|}$, so

$$\sigma \leq \sqrt{1 - \frac{\varepsilon}{|G|^3}} \leq 1 - \frac{\varepsilon}{2|G|^3} \leq \exp\left(-\frac{\varepsilon}{2|G|^3}\right),$$

as desired. □

Now we will use the ε-balanced condition to give a regularity condition for matrices.

Definition 3.8. Let S be a finite set. A *partition* of S is a collection $\mathcal{P} = \{P_1, \ldots, P_k\} \subseteq 2^S$, such that $S = P_1 \sqcup P_2 \sqcup \cdots \sqcup P_k$ and each P_i is nonempty. We say $|\mathcal{P}| = \max_i \# P_i$ and $\# \mathcal{P} = k$. If $\sigma \subseteq 2^{[n]}$, write $\cup \sigma$ for $\bigcup_{S \in \sigma} S$.

Note that $\# \mathcal{P} \cdot |\mathcal{P}| \geq \# S$.

The next definition specifies the kinds of restrictions we will give for the matrices in our universality class. The idea is that we can split up the columns of the matrix and then the rows, so that the resulting sections of the matrix are ε-balanced.

If M is an $n \times m$ matrix, $S = \{s_1 < \cdots < s_k\} \subset [n]$, and $T = \{t_1 < \cdots < t_\ell\} \subset [m]$, then $M_{S,T}$ is the $k \times \ell$ matrix $(M_{s_i, t_j})_{1 \leq i \leq k, 1 \leq j \leq \ell}$.

Definition 3.9. An $n \times m$ random matrix M with entries in a ring R is (w, h, ε)-balanced if there is a partition $\mathcal{Q} = \{Q_1, \dots, Q_r\}$ of $[m]$ and a partition $\mathcal{P} = \{P_1, \dots, P_\ell\}$ of $[n]$ with $|\mathcal{Q}| \le w$, $|\mathcal{P}| \le h$, and such that each random matrix M_{P_i,Q_j} is ε-balanced in the additive abelian group $(R^{\#P_i})^{\#Q_j}$ and the random matrices M_{P_i,Q_j} are independent.

If $|\mathcal{P}| = |\mathcal{Q}| = 1$ then we recover the definition of ε-balanced from [Woo19] and other related work.

Bounds for most maps f

It turns out that (w, h, ε)-balanced is a strong enough condition that we can get bounds on $\mathbb{P}[f(M) = 0]$ for the vast majority of maps f.

Definition 3.10. Let $\mathcal{P} = \{P_1, \dots, P_\ell\}$ be a partition of $[n]$ and G be a finite abelian group. A function $f \colon V \to G$ is a *$\mathcal{P}$-code* of distance w if for any $\sigma \subset [\#\mathcal{P}]$ with $| \cup \sigma| < w$, we have $f(V_{\setminus \cup \sigma}) = G$.

To approximate $\mathbb{P}[f(M) = 0]$ for codes f, we will split the matrices M into independent sets of columns. Each such set of r random columns gets mapped to something close to uniform in G^r. The following lemma is analogous to [Woo19, Lemma 2.1].

Lemma 3.11. *Let $n, r \ge 1$ be integers. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let N be the number of subgroups of G. Let $\varepsilon > 0$ and $\delta > 0$ be real numbers. Let $V = (\mathbb{Z}/a\mathbb{Z})^n$. Let $\mathcal{P} = \{P_i\}$ be a partition of $[n]$ with $|\mathcal{P}| = \ell$. Let $f \in \operatorname{Hom}(V, G)$ be a $\mathcal{P}$-code of distance w.*

Let M be an $n \times r$ random matrix in V^r such that the matrices $M_{P_i,[r]}$ are independent and ε-balanced as random elements of $((\mathbb{Z}/a\mathbb{Z})^{\#P_i})^r$.

Let $g_1, \dots, g_r \in G$. Then

$$|\mathbb{P}[f(M) = (g_1, \dots, g_r)] - |G|^{-r}| \le N \exp\left(-\frac{\varepsilon w}{2\ell N |G|^{3r}}\right)$$

Proof. Let $e_1, \dots, e_n$ be the standard generating set for V.

The idea is to treat $f(M)$ as a random walk in G^r. We have

$$f(M) = \sum_{i=1}^{\#\mathcal{P}} f(M_{P_i,[r]}),$$

where $M_{P_i,[r]}$ is interpreted as an ε-balanced random element of $\langle e_j \mid j \in P_i\rangle^r \cong ((\mathbb{Z}/a\mathbb{Z})^{\# P_i})^r$, a subgroup of $((\mathbb{Z}/a\mathbb{Z})^n)^r$.

Let $S = \{H \leq G \mid H = f(M_{P_i,[r]})$ for at least $w/\ell N$ values of $i\}$. Since there are at most N subgroups of G not in S, there are strictly fewer than w/ℓ values of i such that $f(M_{P_i,[r]}) \notin S$, and for these i we must have $|\bigcup_i P_i| \leq w$. Since f is a $\mathcal{P}$-code of distance w, it remains surjective if we discard all of these indices, which means the images of the $M_{P_i,[r]}$s with $f(M_{P_i,[r]}) \in S$ generate G. In other words, we have $\langle \bigcup_{H\in S} H\rangle = G$. The subgroups in S will be the ones we use in the random walk, applying Theorem 2.1.

By construction of S, for each H in S we have $\# I_H \geq w/\ell N$. By Lemma 3.6, the steps $f(M_{P_i,[r]})$ are ε-balanced, which means that by Lemma 3.7 we have $\sigma_i \leq \exp\left(-\frac{\varepsilon}{2|G|^{3r}}\right)$ (using the fact that each $f(M_{P_i,[r]})$ is supported on a subgroup of G^r).

Hence by Theorem 2.1 we have

$$|\mathbb{P}[f(M) = (g_1, \ldots, g_r)] - |G|^{-r}| \leq \sum_{H\in S} \exp\left(-\frac{\varepsilon w}{2\ell N|G|^{3r}}\right) \leq N \exp\left(-\frac{\varepsilon w}{2\ell N|G|^{3r}}\right),$$

as desired. □

To combine these estimates we will use a result which is a more general version of Lemma 3.3:

Lemma 3.12. *Let* $x_1, \ldots, x_m \geq -1$ *be real numbers such that* $\sum_{i=1}^m \max\{0, x_i\} \leq \log 2$. *Then*

$$\left|\prod_{i=1}^m (1 + x_i) - 1\right| \leq 2\sum_{i=1}^m |x_i|$$

and

$$\sum_{i=1}^m \min\{0, x_i\} \leq \prod_{i=1}^m (1 + x_i) - 1 \leq 2\sum_{i=1}^m \max\{0, x_i\}.$$

Proof. The first statement follows from the second statement because $\max\{0, x_i\} \leq |x_i|$ and $\min\{0, x_i\} \geq -|x_i|$. So, we will show the second statement.

First, assume $x_i \leq 0$ for all i. In that case,

$$\prod_{i=1}^m (1 + x_i) \geq 1 + \sum_{i=1}^m x_i.$$

Next, assume $x_i \geq 0$ for all i. Using the fact that $1 + x_i \leq e^{x_i}$, we get

$$\prod_{i=1}^{m}(1 + x_i) \leq e^{\sum_{i=1}^{m} x_i}.$$

We have $e^x - 1 = 2x$ at $x = 0$ and $\frac{d}{dx}(e^x - 1) \leq \frac{d}{dx}(2x)$ for $x \leq \log 2$, so $e^x - 1 \leq 2x$ for $0 \leq x \leq \log 2$. Hence, if $\sum_{i=1}^{m} x_i \leq \log 2$, then $\exp\left(\sum_{i=1}^{m} x_i\right) - 1 \leq 2\sum_{i=1}^{m} x_i$.

Now consider the general case. By replacing each negative x_i with zero, we can only increase the product $\prod_{i=1}^{m}(1 + x_i)$. On the other hand, by replacing each positive x_i with zero, we can only decrease it. Hence, for general x_i, we get

$$\begin{aligned}\sum_{i=1}^{m} \min\{0, x_i\} &\leq \prod_{i=1}^{m}(1 + \min\{0, x_i\}) - 1 \\ &\leq \prod_{i=1}^{m}(1 + x_i) - 1 \\ &\leq \prod_{i=1}^{m}(1 + \max\{0, x_i\}) - 1 \leq 2\sum_{i=1}^{m} \max\{0, x_i\}.\end{aligned}$$

□

Applying this lemma with x_i being the error in Lemma 3.11 multiplied by $|G|^r$ yields an estimate on the probability that the whole matrix maps to zero:

Lemma 3.13. *Let $u \in \mathbb{Z}$. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let N be the number of subgroups of G. Let $(w_n)_n$, $(h_n)_n$, $(\delta_n)_n$, $(\varepsilon_n)_n$ be sequences of real numbers such that $w_n = o(\log n)$, $h_n = O(n^{1-\alpha})$, and $\varepsilon_n \delta_n \geq n^{-\alpha+\beta}$ for some $0 < \beta \leq \alpha \leq 1$.*

For a natural number n, let $V = (\mathbb{Z}/a\mathbb{Z})^n$. Let M be an $(w_n, h_n, \varepsilon_n)$-balanced $n \times (n + u)$ random matrix with entries in $\mathbb{Z}/a\mathbb{Z}$. Let $\mathcal{P}$ be the row partition associated to M and let $f \in \mathrm{Hom}(V, G)$ be a $\mathcal{P}$-code of distance $n\delta_n$.

Then there are constants $K, c, \gamma > 0$ depending only on G, α, β, and the sequence h_n such that for all $g_1, \ldots, g_{n+u} \in G$,

$$|\mathbb{P}[f(M) = (g_1, \ldots, g_{n+u})] - |G|^{-n-u}| \leq \frac{K\exp(-cn^\gamma)}{|G|^{n+u}}$$

Proof. Let $\mathcal{P}$ and $\mathcal{Q}$ be the row and column partitions for M as in the definition of $(w_n, h_n, \varepsilon_n)$-balanced. Let $M_i = M_{[n],Q_i}$ for each i. Let $g_{Q_i} = (g_j \mid j \in Q_i)$. By independence,

$$\mathbb{P}[f(M) = (g_1, \ldots, g_{n+u})] = \prod_i \mathbb{P}[f(M_i) = g_{Q_i}].$$

For each i, let $x_i = |G|^{\#Q_i}\mathbb{P}[f(M_i) = g_{Q_i}] - 1$. By Lemma 3.11, we have

$$\begin{aligned}|x_i| &\le N|G|^{\#Q_i} \exp\left(-\frac{n\varepsilon_n\delta_n}{2Nh_n|G|^{3\#Q_i}}\right)\\ &\le N|G|^{w_n} \exp\left(-\frac{n\varepsilon_n\delta_n}{2Nh_n|G|^{3w_n}}\right).\end{aligned}$$

Hence we have

$$\log|x_i| \le \log N + w_n \log|G| - \frac{n\varepsilon_n\delta_n}{2Nh_n|G|^{3w_n}}.$$

Since $h_n = O(n^{1-\alpha})$ and $\varepsilon_n\delta_n \ge n^{-\alpha+\beta}$, there is a constant C depending only on the proportionality constant in h_n such that for large enough n we have $\frac{\varepsilon_n\delta_n}{h_n} \ge Cn^{\beta-1}$ so that $\frac{n\varepsilon_n\delta_n}{2Nh_n|G|^{3w_n}} \ge \frac{Cn^\beta}{2N|G|^{3w_n}}$

Since $w_n = o(\log n)$, for large enough n we have $w_n \le \frac{\beta\log n}{6\log|G|}$ so that $|G|^{3w_n} = e^{3w_n\log|G|} \le n^{\beta/2}$ and, for large enough n, $\frac{n\varepsilon_n\delta_n}{2Nh_n|G|^{3w_n}} \ge \frac{Cn^{\beta/2}}{2N}$.

Finally, since $\log N + w_n\log|G| = o(\log n)$, we also have that for n large enough, $\log|x_i| \le -\frac{C}{4N}n^{\beta/2}$ and $|x_i| \le \exp\left(-\frac{C}{4N}n^{\beta/2}\right)$. In particular, for n large enough,

$$\sum_{i=1}^m |x_i| \le m\exp\left(-\frac{C}{4N}n^{\beta/2}\right) \le n\exp\left(-\frac{C}{4N}n^{\beta/2}\right) \le \log 2.$$

By Lemma 3.12, we therefore have that for such n,

$$\begin{aligned}||G|^{n+u}\mathbb{P}[f(M) = (g_1,\dots,g_{n+u})] - 1| &= \left|\prod_{i=1}^m |G|^{\#Q_i}\mathbb{P}[f(M_i) = g_{Q_i}]| - 1\right|\\ &= \left|\prod_{i=1}^m (1+x_i) - 1\right|\\ &\le 2\sum_{i=1}^m |x_i|\\ &\le 2n\exp\left(-\frac{C}{4N}n^{\beta/2}\right)\\ &= 2n\exp\left(-\frac{C}{8N}n^{\beta/2}\right)\cdot\exp\left(-\frac{C}{8N}n^{\beta/2}\right).\end{aligned}$$

Since $\lim_{n\to\infty} 2n\exp\left(-\frac{C}{8N}n^{\beta/2}\right) = 0$, the expression $2n\exp\left(-\frac{C}{8N}n^{\beta/2}\right)$ is uniformly bounded above by some constant for all $n \ge 0$. Then the appropriate constant K can be chosen so that

$$||G|^{n+u}\mathbb{P}[f(M) = (g_1,\dots,g_{n+u})] - 1| \le K\exp\left(-\frac{C}{8N}n^{\beta/2}\right),$$

for all n, as desired. □

Bounds for the rest of the maps

This gives results for the case when f is a code, but we still need to account for non-codes. To do this, we will show that non-codes make up a negligible proportion of all maps $V \to G$ and thus contribute only a small error term to the sum $\mathbb{E}[\#\operatorname{Sur}(\operatorname{coker}(M), G)]$. However, it turns out that splitting maps into codes and non-codes is not enough to get this bound. Instead, we will categorize non-codes by how far they are from being codes.

Definition 3.14. The $(\mathcal{P}, \delta)$*-depth* of $f \in \operatorname{Hom}(V, G)$ is the maximal positive D such that there is a $\sigma \subset [\#\mathcal{P}]$ with $|\cup\sigma| < \ell(D)\delta n$ such that $D = [G : f(V_{\setminus\cup\sigma})]$, or 1 if there is no such D.

We can count the number of f that have given $(\mathcal{P}, \delta)$-depth:

Lemma 3.15. *If $D > 1$, then the number of $f \in \operatorname{Hom}(V, G)$ of $(\mathcal{P}, \delta)$-depth D is at most*

$$K\binom{n}{\lceil \ell(D)\delta n\rceil - 1}2^{\ell(D)\delta n}|G|^n D^{-n+\ell(D)\delta n},$$

where K is the number of subgroups of G of index D.

Proof. For each f of $(\mathcal{P}, \delta)$-depth D, there is a $\sigma \subset [\#\mathcal{P}]$ as described above. There must be some set $S \subset [n]$ with $\#S = \lceil \ell(D)\delta n\rceil - 1$ and $\cup\sigma \subseteq S$. There are $\binom{n}{\lceil \ell(D)\delta n\rceil - 1}$ choices of S, and for each choice of S, there are certainly at most $2^{\#S} = 2^{\lceil \ell(D)\delta n\rceil - 1} \leq 2^{\ell(D)\delta n}$ choices of $\cup\sigma$. Since $\mathcal{P}$ is a partition, $\cup\sigma$ uniquely determines σ, so there are at most $2^{\ell(D)\delta n}$ choices of σ for each choice of S.

Now we count how many f of $(\mathcal{P}, \delta)$-depth D have each choice of σ, so fix σ. There are a constant number of subgroups of G with index D, say K of them.

Fix a subgroup H of G with index D. We now count the number of f with $f(V_{\setminus\cup\sigma}) = H$. There are at most $|H|^{n-|\cup\sigma|}$ maps from $V_{\setminus\cup\sigma}$ to H, and for each such map, there are at most $|G|^{|\cup\sigma|}$ homomorphisms from V to G which restrict appropriately. Hence, there are at most

$$\begin{aligned}|H|^{n-|\cup\sigma|}|G|^{|\cup\sigma|} &= |G|^{n-|\cup\sigma|}D^{-n+|\cup\sigma|}|G|^{|\cup\sigma|}\\ &= |G|^n D^{-n+|\cup\sigma|} \leq |G|^n D^{-n+|\cup\sigma|} \leq |G|^n D^{-n+\ell(D)\delta n}\end{aligned}$$

maps f with $f(V_{\setminus\cup\sigma}) = H$. Combined with the counts of choices of σ and subgroups of G of index D, we get the lemma. □

For non-codes, we do not get precise estimates on $\mathbb{P}[f(M) = 0]$, but we can get upper bounds.

Lemma 3.16. *Let $r \geq 1$ be an integer. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let N be the number of subgroups of G. Let $\varepsilon > 0$ and $\delta > 0$ be real numbers. Let $V = (\mathbb{Z}/a\mathbb{Z})^n$. Let $\mathcal{P} = \{P_1, \ldots, P_m\}$ be a partition of $[n]$ with $|\mathcal{P}| = \ell$. Let $f \in \operatorname{Hom}(V, G)$ have $(\mathcal{P}, \delta)$-depth $D > 1$ with $[G : f(V)] < D$.*

Let M be an $n \times r$ random matrix in V^r such that the matrices $M_{P_i,[r]}$ are independent and ε-balanced as random elements of $((\mathbb{Z}/a\mathbb{Z})^{\#P_i})^r$.

Then

$$\mathbb{P}[f(M) = 0] \leq (1 - \varepsilon)\left(D^r|G|^{-r} + N \exp\left(-\frac{\varepsilon\delta n}{2N\ell(D^{-1}|G|)^{3r}}\right)\right)$$

Proof. Since f has $(\mathcal{P}, \delta)$-depth D, there is a $\sigma \subset [\#\mathcal{P}]$ with $|\cup\sigma| < \ell(D)\delta n$ such that $D = [G : f(V_{\backslash\cup\sigma})]$. Let $f(V_{\backslash\cup\sigma}) =: H$. Since $[G : f(V)] < D$, we cannot have that σ is empty.

Write $f(M) = \sum_{j\notin\sigma} f(M_{P_j,[r]}) + \sum_{j\in\sigma} f(M_{P_j,[r]})$. So,

$$\mathbb{P}[f(M) = 0] = \mathbb{P}[f(M) \in H]\mathbb{P}\left[\sum_{j\notin\sigma} f(M_{P_j,[r]}) = -\sum_{j\in\sigma} f(M_{P_j,[r]}) \mid f(M) \in H\right].$$

We bound the two probabilities on the right side separately. Note that since $\sum_{j\in\sigma} f(M_{P_j,[r]}) \in H$, we have $f(M) \in H$ exactly when $\sum_{j\notin\sigma} f(M_{P_j,[r]}) \in H$. Since $[G : f(V)] < [G : H]$, there must be some $i \in \sigma$ such that $f(M_{P_i,[r]})$ reduces to a nonzero element of G/H. Conditioning on all other $M_{P_k,[r]}$ for $k \neq i$, by the ε-balanced assumption we have that

$$\mathbb{P}\left[f(M) \in H\right] = \mathbb{P}\left[f(M_{P_i,[r]}) \equiv -\sum_{j\in\sigma\backslash\{k\}} f(M_{P_j,[r]}) \pmod{H}\right] \leq 1 - \varepsilon.$$

For the second probability, let $\mathcal{P}'$ be the partition of $[n] \setminus \cup\sigma$ induced by $\mathcal{P}$. Notice that $f|_{V_{\backslash\cup\sigma}}$ is a $\mathcal{P}'$-code of distance δn. Indeed, suppose there is some $\tau \subset [\#\mathcal{P}']$ with $|\tau| < \delta n$ inducing some $\tau' \subset [\#\mathcal{P}]$ with $f(V_{\backslash\cup(\sigma\cup\tau)}) \neq H$. Then the image of $f|_{V_{\backslash\cup(\sigma\cup\tau)}}$ would have degree strictly greater than D, contradicting maximality of D.

Now we can apply Lemma 3.11 to the submatrix $M_{[n]\backslash\cup\sigma,[r]}$ and the code f mapping it into H^r. If N' is the number of subgroups of H and $\ell' = |\mathcal{P}'|$, then conditioning on $M_{P_j,[r]}$ for $j \in \sigma$ gives

$$\mathbb{P}\left[\sum_{j\notin\sigma} f(M_{P_j,[r]}) = -\sum_{j\in\sigma} f(M_{P_j,[r]}) \mid f(M) \in H\right] \le |H|^{-r} + N' \exp\left(-\frac{\varepsilon\delta n}{2N'\ell'|H|^{3r}}\right)$$

$$\le D^r|G|^{-r} + N \exp\left(-\frac{\varepsilon\delta n}{2N\ell(D^{-1}|G|)^{3r}}\right),$$

and the lemma follows. □

Finally, we use Lemma 3.12 again to get a bound for the full $n \times (n+u)$ matrix:

Lemma 3.17. *Let $u \in \mathbb{Z}$. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let N be the number of subgroups of G. Let $(w_n)_n$, $(h_n)_n$, $(\delta_n)_n$, $(\varepsilon_n)_n$ be sequences of real numbers such that $w_n = o(\log n)$, $h_n = O(n^{1-\alpha})$, and $\varepsilon_n\delta_n \ge n^{-\alpha+\beta}$ for some $0 < \beta \le \alpha \le 1$.*

For a natural number n, let $V = (\mathbb{Z}/a\mathbb{Z})^n$. Let M be an $(w_n, h_n, \varepsilon_n)$-balanced $n \times n + u$ random matrix with entries in $\mathbb{Z}/a\mathbb{Z}$. Let $\mathcal{P}$ be the row partition associated to M and let $f \in \mathrm{Hom}(V, G)$ have $(\mathcal{P}, \delta_n)$-depth $D > 1$, with $[G : f(V)] < D$.

Then there is a constant $K > 0$ depending only on u, G, α, β, and the sequences h_n, w_n such that for all n,

$$\mathbb{P}[f(M) = 0] \le K \exp\left(-\varepsilon_n \frac{n}{\log n}\right) D^n |G|^{-n}.$$

Proof. Let Q be the column partition for M as in the definition of $(w_n, h_n, \varepsilon_n)$-balanced. Let $M_i = M_{[n],Q_i}$ for each i. By independence,

$$\mathbb{P}[f(M) = 0] = \prod_i \mathbb{P}[f(M_i) = 0].$$

For each i, let $x_i = \frac{|G|^{\#Q_i} D^{-\#Q_i}}{1-\varepsilon_n}\mathbb{P}[f(M_i) = 0] - 1$. By Lemma 3.16, we have

$$\max\{0, x_i\} \le N|G|^{\#Q_i} D^{-\#Q_i} \exp\left(-\frac{n\varepsilon_n\delta_n}{2Nh_n(D^{-1}|G|)^{3\#Q_i}}\right)$$

$$\le N|G|^{w_n} D^{-w_n} \exp\left(-\frac{n\varepsilon_n\delta_n}{2Nh_n(D^{-1}|G|)^{3w_n}}\right).$$

By the same argument as in the proof of Lemma 3.13, there is some constant C depending only on α and the sequence h_n such that for large enough n (where "large enough" depends on u, G, α, β, and the sequences h_n and w_n), we have

$$\sum_{i=1}^{m} \max\{0, x_i\} \le n \exp\left(\frac{C}{4N} n^{\beta/2}\right) \le \log 2.$$

By Lemma 3.12, we therefore have that for such n,

$$\begin{aligned}
\frac{(D^{-1}|G|)^{n+u}}{(1-\varepsilon_n)^{\#Q}}\mathbb{P}[f(M)=0]-1 &= \prod_{i=1}^{m}\frac{(D^{-1}|G|)^{\#Q_i}}{1-\varepsilon_n}\mathbb{P}[f(M_i)=0]-1\\
&= \prod_{i=1}^{m}(1+x_i)-1\\
&\leq 2\sum_{i=1}^{m}\max\{0,x_i\}\\
&\leq 2n\exp\left(-\frac{C}{4N}n^{\beta/2}\right)\\
&= 2n\exp\left(-\frac{C}{8N}n^{\beta/2}\right)\cdot\exp\left(-\frac{C}{8N}n^{\beta/2}\right).
\end{aligned}$$

Since $\lim_{n\to\infty} 2n\exp\left(-\frac{C}{8N}n^{\beta/2}\right)=0$, the expression $2n\exp\left(-\frac{C}{8N}n^{\beta/2}\right)$ is uniformly bounded above by some constant for all $n\geq 0$. Then the appropriate constant K' can be chosen so that

$$\frac{(D^{-1}|G|)^{n+u}}{(1-\varepsilon_n)^{\#Q}}\mathbb{P}[f(M)=0]-1\leq K'\exp\left(-\frac{C}{8N}n^{\beta/2}\right),$$

for all n. Hence we have

$$\begin{aligned}
\mathbb{P}[f(M)=0] &\leq D^{n+u}|G|^{-n-u}(1-\varepsilon_n)^{\#Q}\left(1+K'\exp\left(-\frac{C}{8N}n^{\beta/2}\right)\right)\\
&\leq D^{n+u}|G|^{-n-u}\exp(-\varepsilon_n\#Q)\left(1+K'\exp\left(-\frac{C}{8N}n^{\beta/2}\right)\right)\\
&\leq (K'+1)D^{n+u}|G|^{-n-u}\exp(-\varepsilon_n\#Q).
\end{aligned}$$

The lemma follows from the fact that for large enough n, we have $w_n\leq\log n$, so $\#Q\geq\frac{n}{w_n}\geq\frac{n}{\log n}$. □

Putting it all together

Finally, we can combine all these results to compute the limiting moments for cokernels of (w_n,h_n,ε_n)-balanced random matrices. The most relevant part of this proof is the part where we handle the non-codes. This will involve a careful choice of the sequence δ_n.

Theorem 3.18. *Let $u\in\mathbb{Z}$. Let G be a finite abelian group and let a be a multiple of the exponent of G. Let $(w_n)_n,(h_n)_n$ be sequences of real numbers such that $w_n=o(\log n)$, $h_n=O(n^{1-\alpha})$, and $\varepsilon_n\geq n^{-\beta}$ for some $0<\alpha\leq 1$ and $0<\beta<\alpha/2$.*

Then there are $c, K, \gamma > 0$ such that the following holds for every natural number n large enough that it makes sense. Let M be an $(w_n, h_n, \varepsilon_n)$-balanced $n \times (n+u)$ random matrix with entries in $\mathbb{Z}/a\mathbb{Z}$. Then

$$|\mathbb{E}[\#\operatorname{Sur}(\operatorname{cok}(M), G)] - |G|^{-u}| \le Ke^{-cn^{\gamma}}.$$

Proof. Let $V = (\mathbb{Z}/a\mathbb{Z})^n$. As usual with this kind of approach, we want to estimate $\sum_{f \in \operatorname{Sur}(V,G)} \mathbb{P}[f(M) = 0]$. Let $\mathcal{P}, \mathcal{Q}$ be the row and column partitions witnessing the $(w_n, h_n, \varepsilon_n)$-balancedness of M.

Let $\delta_n = n^{-\alpha/2}$. Note that then $\varepsilon_n \delta_n \ge n^{-\beta-\alpha/2}$ with $-\beta - \alpha/2 > -\alpha$, so δ_n satisfies the conditions for Lemmas 3.13 and 3.17.

Just like in [Woo19, Theorem 2.9], we will allow K to change in each line as long as it remains a constant depending only on $a, u, \alpha, \beta, (h_n)_n, (w_n)_n, G$.

We have

$$\begin{aligned}
&\left|\mathbb{E}[\#\operatorname{Sur}(\operatorname{cok}(M), G)] - \frac{1}{|G|^u}\right| \\
&\quad = \left| \sum_{f \in \operatorname{Sur}(V,G)} \mathbb{P}[f(M) = 0] - \frac{1}{|G|^u} \right| \\
&\quad = \left| \sum_{f \in \operatorname{Sur}(V,G)} \mathbb{P}[f(M) = 0] - \sum_{f \in \operatorname{Hom}(V,G)} \frac{1}{|G|^{n+u}} \right| \\
&\quad \le \sum_{\substack{f \in \operatorname{Sur}(V,G) \\ f \text{ code of distance } n\delta_n}} \left| \mathbb{P}[f(M) = 0] - \frac{1}{|G|^{n+u}} \right| && (2) \\
&\qquad + \sum_{\substack{D > 1 \\ D \,|\, |G|}} \sum_{\substack{f \in \operatorname{Sur}(V,G) \\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}} \mathbb{P}[f(M) = 0] && (3) \\
&\qquad + \sum_{D \,|\, |G|} \sum_{\substack{f \in \operatorname{Sur}(V,G) \\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}} \frac{1}{|G|^{n+u}} && (4) \\
&\qquad + \sum_{f \in \operatorname{Hom}(V,G) \setminus \operatorname{Sur}(V,G)} \frac{1}{|G|^{n+u}} && (5)
\end{aligned}$$

Wood showed that (4) is bounded above by $Ke^{-n \log 2}$. By Lemma 3.13, we can bound (1):

$$\sum_{\substack{f\in \mathrm{Sur}(V,G)\\ f \text{ code of distance } n\delta_n}} \left|\mathbb{P}[f(M)=0]-\frac{1}{|G|^{n+u}}\right| \le \sum_{\substack{f\in \mathrm{Sur}(V,G)\\ f \text{ code of distance } n\delta_n}} \frac{K\exp(-cn^{\gamma})}{|G|^{n+u}}$$
$$\le |G|^n\frac{K\exp(-cn^{\gamma})}{|G|^{n+u}}$$
$$= K\exp(-cn^{\gamma}).$$

To bound (2) and (3) we use Lemma 3.15. For each $D>1$, there are at most

$$K\binom{n}{\lceil \ell(D)n\delta_n\rceil -1}2^{\ell(D)n\delta_n}|G|^nD^{-n+\ell(D)n\delta_n}$$

maps of $(\mathcal{P},\delta_n)$-depth D. A standard inequality says that $\binom{n}{k}\le\left(\frac{ne}{k}\right)^k$, so for $\lceil \ell(D)n\delta_n\rceil\ge 2$ (which is the case for n large enough, independent of D)

$$\binom{n}{\lceil \ell(D)n\delta_n\rceil -1} \le \left(\frac{ne}{\lceil \ell(D)n\delta_n\rceil -1}\right)^{\lceil \ell(D)n\delta_n\rceil -1}$$
$$\le \left(\frac{2ne}{\ell(D)n\delta_n}\right)^{\ell(D)n\delta_n}$$
$$= \left(\frac{2e}{\ell(D)\delta_n}\right)^{\ell(D)n\delta_n}$$
$$= \exp\left(\ell(D)n\delta_n\left(1+\log 2-\log \ell(D)-\log\delta_n\right)\right).$$

Hence, the number of maps of $(\mathcal{P},\delta_n)$-depth D is at most

$$K|G|^nD^{-n}\exp\left(\ell(D)n\delta_n\left(\log\frac{4eD}{\ell(D)}-\log\delta_n\right)\right) = K|G|^n\exp\left(\ell(D)n\delta_n\left(\log\frac{4eD}{\ell(D)}-\log\delta_n\right)-n\log D\right)$$
$$\le K|G|^n\exp\left(\ell(|G|)n\delta_n\left(\log\frac{4e|G|}{\ell(|G|)}-\log\delta_n\right)-n\log 2\right)$$

Since $\lim_{\delta\to 0}\delta\log\delta=0$ and $\delta_n\to 0$ as $n\to\infty$, for large enough n (depending only on β and $|G|$) we have $\ell(|G|)\delta_n\left(\log\frac{4e|G|}{\ell(|G|)}-\log\delta_n\right)\le\frac{1}{2}\log 2$, which means that for large enough n,

$$\sum_{D||G|}\sum_{\substack{f\in \mathrm{Sur}(V,G)\\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}}\frac{1}{|G|^{n+u}} \le \sum_{D||G|}K|G|^{-u}\exp\left(\ell(|G|)n\delta_n\left(\log\frac{4e|G|}{\ell(|G|)}-\log\delta_n\right)-n\log 2\right)$$
$$\le \sum_{D||G|}K\exp\left(-\frac{\log 2}{2}n\right)$$
$$\le K\exp\left(-\frac{\log 2}{2}n\right),$$

bounding (3) as desired.

Finally, we need to bound (2). From Lemma 3.17, we have that if f has $(\mathcal{P}, \delta_n)$-depth D,

$$\mathbb{P}[f(M) = 0] \leq K \exp\left(-\varepsilon_n \frac{n}{\log n}\right) D^n |G|^{-n},$$

which means

$$\begin{aligned}\sum_{\substack{f \in \mathrm{Sur}(V,G) \\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}} \mathbb{P}[f(M) = 0] &\leq K \exp\left(\ell(D) n \delta_n \left(\log \frac{4eD}{\ell(D)} - \log \delta_n\right) - \varepsilon_n \frac{n}{\log n}\right) \\ &\leq K \exp\left(\ell(|G|) n^{1-\alpha/2} \left(\log \frac{4e|G|}{\ell(|G|)} + \frac{\alpha}{2} \log n\right) - \frac{n^{1-\beta}}{\log n}\right).\end{aligned}$$

Since $\beta < \alpha/2$, we have that $n^{1-\alpha/2}(\log n)^2 = o(n^{1-\beta})$, so

$$\lim_{n\to\infty} \frac{\ell(|G|) n^{1-\alpha/2} \left(\log \frac{4e|G|}{\ell(|G|)} + \frac{\alpha}{2} \log n\right)}{n^{1-\beta}/\log n} = 0.$$

Hence for large enough n (depending only on G, α, and β), we have $\ell(|G|) n^{1-\alpha/2} \left(\log \frac{4e|G|}{\ell(|G|)} + \frac{\alpha}{2} \log n\right) \leq \frac{1}{2} \frac{n^{1-\beta}}{\log n}$ and

$$\sum_{\substack{f \in \mathrm{Sur}(V,G) \\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}} \mathbb{P}[f(M) = 0] \leq K \exp\left(-\frac{n^{1-\beta}}{2 \log n}\right).$$

For the same reason, for n large enough (depending only on β) we have

$$\sum_{\substack{f \in \mathrm{Sur}(V,G) \\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}} \mathbb{P}[f(M) = 0] \leq K \exp\left(-\frac{1}{2} n^{\frac{1-\beta}{2}}\right),$$

which means

$$\sum_{\substack{D>1 \\ D \| G|}} \sum_{\substack{f \in \mathrm{Sur}(V,G) \\ f \text{ of } (\mathcal{P},\delta_n)\text{-depth } D}} \mathbb{P}[f(M) = 0] \leq K \exp\left(-\frac{1}{2} n^{\frac{1-\beta}{2}}\right),$$

giving us a bound on (2).

Finally, take c and γ appropriately to obtain the desired result. □

Chapter 4

PARTIAL RESULTS AND FUTURE WORK

In this chapter, we give some partial results toward proving a universality theorem for random class-c nilpotent groups. Section 4.1 gives some formalism that simplifies application of Theorem 2.5. In Section 4.2, we use this formalism to prove an equidistribution result for nilpotent groups analogous to Lemmas 3.2 and 3.11. Future work will attack the problem of bounding the error caused by non-codes in the case of nilpotent groups.

4.1 Quotient Sequences

The result of Theorem 2.5 is quite powerful, but it is locked behind a technical condition. In order to apply Theorem 2.5, we can abstract the condition to make it easier to work with. Given the term $\frac{|G_j|-1}{|G|}$ in the result of Theorem 2.5, the following definition quite naturally captures the condition for the theorem.

Definition 4.1. Let G be a group. A *quotient sequence* of length k is a sequence $Q = (Q_i \colon G_{i-1} \to G_i)_{i=1}^k$ of surjections

$$G = G_0 \overset{Q_1}{\twoheadrightarrow} G_1 \overset{Q_2}{\twoheadrightarrow} \dots \overset{Q_{k-1}}{\twoheadrightarrow} G_{k-1} \overset{Q_k}{\twoheadrightarrow} G_k,$$

with $\tilde{Q}_{i,j} = Q_j \circ \cdots \circ Q_i$ and $\tilde{Q}_i = \tilde{Q}_{1,i}$. A subset S of G_i is called *sub-level j* with respect to Q if $\tilde{Q}_{i,j}(S) = \{e\} \in G_j$. S is called *level j* with respect to Q if $\tilde{Q}_{i-1,j}(S) = \ker Q_j$. Q is *complete* if $G_k = \{e\}$.

Remark 4.2. 1. *If S is sub-level i, then S is also sub-level j for all $j > i$.*

2. *Each group has a unique trivial complete quotient sequence $G \to \{e\}$.*

3. *Given a quotient sequence Q' of length ℓ starting at G_j and ending at G_{j+1}, Q can be refined by inserting Q' between G_j and G_{j+1}, yielding a quotient sequence of length $k + \ell - 1$.*

Lemma 4.3. *Let $Q \colon G = G_0 \to \cdots \to G_{k+1}$ be a quotient sequence. Let $H_1, \dots, H_k$ be normal subgroups of G_j such that $\ker Q_{j+1} = \langle \bigcup_{i=1}^k H_i \rangle$. Then there exists a quotient sequence*

$$Q' \colon G = G_0 \to \cdots \to G_j \to G_j/H_1 \to \cdots \to G_{j+1} \to \cdots \to G_{k+1}$$

such that H_i is level $j+i$ for each i.

Proof. By Remark 4.2, it is enough to show this when Q is a one-step quotient sequence $Q_1\colon G \to G_1$.

The idea is to inductively construct a quotient sequence for G out of the normal subgroups H_j. This is possible because surjections preserve normality.

Given $G'_0 = G, \dots, G'_j$, and maps $Q'_1, \dots, Q'_j$ let $G'_{j+1} = G'_j/\tilde{Q}'_j(H_{j+1})$. Let Q'_{j+1} be the natural projection and $\tilde{Q}'_{j+1} = Q'_{j+1} \circ \tilde{Q}'_j$. These determine a quotient sequence Q' of length k. By construction, H_j is level j in this quotient sequence.

Observe that for each j, $H_1, \dots, H_j \subseteq \ker \tilde{Q}'_j$, since each level-$i$ set is sub-level j for $j \geq i$. In particular, $H_1, \dots, H_k \subseteq \ker \tilde{Q}'_k$. Since these subgroups generate $\ker Q_1$, $\ker Q_1 \subseteq \ker \tilde{Q}_k$. On the other hand, since each $H_j \subseteq \ker Q_1$, $\ker Q_1 = \ker \tilde{Q}_k$. This gives an isomorphism $G'_k \cong G_1$, so Q' is a quotient sequence from G to G_1, as desired. □

The language of quotient sequences allows for a much more concise rephrasing of Theorem 2.5.

Theorem 4.4. *(Theorem 2.5 with quotient sequences) Let G be a finite group and $Q = (Q_i)_{i=1}^k$ be a complete quotient sequence of G. Let $\mu_1, \dots, \mu_n$ be probability measures on G. Let $\nu_n = \mu_1 * \cdots * \mu_n$. For each $j = 1, \dots, k$, let $I_j = \{i \mid \operatorname{supp} \mu_i \text{ is level } j\}$. Let π be the uniform distribution on G.*

For $i \in I_j$, let σ_i be the second largest singular value of the $(\tilde{Q}_{j-1})_* \mu_i$-random walk on $\ker Q_j$. Then we have

$$||\nu_n - \pi||_{L^2}^2 \leq \sum_{j=1}^{k} \frac{|G_j| - 1}{|G|} \left(\prod_{i \in I_j} \sigma_i^2 \right).$$

To illustrate the use of quotient sequences to apply the results of Theorem 2.5/4.4, we show how a stronger version of Theorem 2.1 follows from Theorem 4.4:

Corollary 4.5. *(Theorem 2.1) Let G be a finite group, and let $\mu_1, \mu_2, \dots, \mu_n$ be probability measures on G. For each subgroup H of G, let $I_H = \{i \mid H = \langle \operatorname{supp} \mu_i \rangle\}$. Let $H_1, \dots, H_k$ be normal subgroups with $G = \left\langle \bigcup_{j=1}^k H_j \right\rangle$. Write $\nu_n = \mu_1 * \cdots * \mu_n$.*

Also, for each i, let σ_i be the second-largest singular value of $*\mu_i$ as an operator on $L^2(\langle \operatorname{supp} \mu_i \rangle)$. Then there are constants $0 < c_j < 1 - \frac{1}{|G|}$ depending only on G and

$H_1, \ldots, H_k$ such that

$$||\nu_n - \pi||_{L^2}^2 \leq \sum_{j=1}^{k} c_j \left(\prod_{i \in I_{H_j}} \sigma_i^2 \right).$$

Proof. By Lemma 4.3 applied to the trivial quotient sequence, the subgroups $H_1, \ldots, H_k$ induce a complete quotient sequence Q for G.

Also, note that if $\langle \operatorname{supp} \mu_i \rangle = H_j$, then $\langle \operatorname{supp}(\tilde{Q}_{j-1})_* \mu_i \rangle = \tilde{Q}_{j-1}(H_j) = \ker Q_j$. Hence, $I_j = I_{H_j}$.

Now we bound the second largest singular values of the $(\tilde{Q}_{j-1})_* \mu_i$-random walks, which we will denote by σ_i'. Recall that σ_i' is the operator norm of $*(\tilde{Q}_{j-1})_* \mu_i$ acting on the subspace of $L^2(\ker Q_j)$ consisting of measures with zero total mass. Let ν be an arbitrary measure on $\ker Q_j$ with zero total mass. Define a measure $\tilde{\nu}$ on H_j by $\tilde{\nu}(h) = \frac{\nu(\tilde{Q}_{j-1}(h))}{|\ker \tilde{Q}_{j-1} \cap H_j|}$. Note that $(\tilde{Q}_{j-1})_* \tilde{\nu} = \nu$, and $\tilde{\nu}$ has zero total mass. In particular, $||\tilde{\nu} * \mu_i||_{L^2(H_j)}^2 \leq \sigma_i^2 ||\tilde{\nu}||_{L^2(H_j)}^2$.

We have

$$\begin{aligned}
||\tilde{\nu}||_{L^2(H_j)}^2 &= \sum_{h \in H_j} \tilde{\nu}(h)^2 \\
&= \sum_{h' \in H_j'} \sum_{h \in \tilde{Q}_{j-1}^{-1}(h') \cap H_j} \tilde{\nu}(h)^2 \\
&= \sum_{h' \in H_j'} \sum_{h \in \tilde{Q}_{j-1}^{-1}(h') \cap H_j} \frac{\nu(h')^2}{|\ker \tilde{Q}_{j-1} \cap H_j|^2} \\
&= \frac{1}{|\ker \tilde{Q}_{j-1} \cap H_j|} ||\nu||_{L^2(\ker Q_j)}^2.
\end{aligned}$$

On the other hand, for any signed measure μ on H_j we have, by Cauchy-Schwarz,

$$\begin{aligned}
||(\tilde{Q}_{j-1})_* \mu||_{L^2(H_j')}^2 &= \sum_{h' \in H_j'} \left(\sum_{h \in \tilde{Q}_{j-1}^{-1}(h') \cap H_j} \mu(h) \right)^2 \\
&\leq |\ker \tilde{Q}_{j-1} \cap H_j| \sum_{h' \in H_j'} \sum_{h \in \tilde{Q}_{j-1}^{-1}(h') \cap H_j} \mu(h)^2 \\
&= |\ker \tilde{Q}_{j-1} \cap H_j| \; ||\mu||_{L^2(H_j)}^2.
\end{aligned}$$

Hence,

$$\begin{aligned}||\nu * (\tilde{Q}_{j-1})_*\mu_i||^2_{L^2(\ker Q_j)} &= ||(\tilde{Q}_{j-1})_*(\tilde{\nu} * \mu_i)||^2_{L^2(\ker Q_j)} \\ &\leq |\ker \tilde{Q}_{j-1} \cap H_j|\ ||\tilde{\nu} * \mu_i||^2_{L^2(H_j)} \\ &\leq |\ker \tilde{Q}_{j-1} \cap H_j| \sigma_i^2 ||\tilde{\nu}||^2_{L^2(H_j)} \\ &\leq \sigma_i^2 ||\nu||^2_{L^2(\ker Q_j)},\end{aligned}$$

which means $\sigma_i' \leq \sigma_i$.

Hence, applying Theorem 4.4 yields

$$||\nu_n - \pi||^2_{L^2} \leq \sum_{j=1}^{k} \frac{|G_j| - 1}{|G|} \left(\prod_{i \in I_j} (\sigma_i')^2 \right) \leq \sum_{j=1}^{k} \frac{|G| - 1}{|G|} \left(\prod_{i \in I_j} \sigma_i^2 \right).$$

□

4.2 Equidistribution on Nilpotent Groups

Given a group G, we say $\gamma_1 G := G$, and for integers $i > 1$, $\gamma_i G = [\gamma_{i-1} G, G]$. We denote the i-step commutator $[g_1, [g_2, [\ldots, [g_{n-1}, g_n]]]]$ by $[g_1, \ldots, g_n]$.

We denote by $\mathbb{Z}^n$ the free abelian group on n generators, F_n the free group on n generators, and $N_{c,n} := F_n/\gamma_{c+1}F_n$ the free nilpotent group of class c on n generators. Note that $N_{1,n} = \mathbb{Z}^n$. We say a generating set S of F_n (respectively, $N_{c,n}$) freely generates F_n ($N_{c,n}$) if there is no nontrivial word in S that reduces to 0 (respectively, every nontrivial word in S is a product of $c+1$-step commutators of words in S).

The universality result of [Woo19] for abelian groups relied on the use of coordinates on abelian groups to construct a universality class. While in general groups do not have well-defined coordinate constructions, it is possible to construct a sort of independent coordinate system on torsion-free nilpotent groups.

The free nilpotent group $N_{c,n}$ has a lower central series

$$N_{c,n} = \gamma_1 N_{c,n} \to \gamma_2 N_{c,n} \to \cdots \to \gamma_c N_{c,n} \to \gamma_{c+1} N_{c,n} = \{e\}.$$

Each successive quotient $\gamma_i N_{c,n}/\gamma_{i+1} N_{c,n}$ is isomorphic to a free abelian group $\mathbb{Z}^m$ for some m. Thus, one can choose elements of $\gamma_i N_{c,n}$ whose projections freely generate $\gamma_i N_{c,n}/\gamma_{i+1} N_{c,n}$ as a free abelian group. One can choose these elements to be i-step commutators in the generators of $N_{c,n}$. If $a_1, \ldots, a_n$ freely generate $N_{c,n}$ as a free class-c nilpotent group, then an i-step commutator in $a_1, \ldots, a_n$ is

$[a_{j_1}, \dots, a_{j_i}] \coloneqq [a_{j_1}, [a_{j_2}, [\dots, [a_{j_{i-1}}, a_{j_i}]]]]$. Then these elements $u_1, \dots, u_m$, called a *Maltsev basis* for $N_{c,n}$, generate $N_{c,n}$, and any element of $N_{c,n}$ can be written uniquely in the form $u_1^{x_1} \dots u_m^{x_m}$. The numbers $x_1, \dots, x_m$ are called *Maltsev coordinates*.

For a free nilpotent group $N_{c,n}$ with given Maltsev basis $u_1, \dots, u_m$, denote by $(x_1, \dots, x_m)_M$ the element $u_1^{x_1} \dots u_m^{x_m}$. If $c = 1$, then the standard basis for $N_{1,n} = \mathbb{Z}^n$ is a Maltsev basis. For more on Maltsev bases, see [CMZ17, Chapter 4.2].

The idea is to replace random integer vectors in [Woo19] with random vectors in the Maltsev coordinates. For this, we will need another definition of codes that works in non-abelian groups:

Definition 4.6. Let $V = \langle S \rangle$ be a group with a distinguished generating set S, and let G be a group. We say that $f \in \mathrm{Hom}(V, G)$ is an *S-code* of distance w (or a code of distance w with respect to S) if for any $T \subseteq S$ with $\#T < w$, the restriction $f|_{\langle S \setminus T \rangle} \colon \langle S \setminus T \rangle \to G$ is surjective.

If f is a code with respect to the free generators of $N_{c,n}$, then it is a code for each successive quotient $\gamma_i N_{c,n} / \gamma_{i+1} N_{c,n}$ in the following sense:

Lemma 4.7. *Let V be a free nilpotent group of class c on $\{u_1, \dots, u_n\}$. Let G be a finite nilpotent group of class at most c. Let $f \colon N_{c,n} \to H$ be a code of distance w with respect to $\{a_1, \dots, a_n\}$. Then for each i, the map f induces a map $f_i \colon \gamma_i N_{c,n} / \gamma_{i+1} N_{c,n} \to \gamma_i G / \gamma_{i+1} G$, which is a code of distance w/i with respect to the i-step commutators in $a_1, \dots, a_n$ that freely generate $\gamma_i N_{c,n} / \gamma_{i+1} N_{c,n}$ as an abelian group.*

Proof. Let S_i be the set of i-step commutators in $a_1, \dots, a_n$ that freely generate $\gamma_i N_{c,n} / \gamma_{i+1} N_{c,n}$ as an abelian group. Let $T_i \subset S_i$ with $\#T_i < w/i$. Let T be the set of generators of $N_{c,n}$ that appear in any i-step commutator in T_i. Each of these commutators can include at most i distinct elements of S, so fewer than w/i commutators include fewer than w distinct elements of S and $\#T < w$.

Consider the subgroup W of $N_{c,n}$ generated by $S \setminus T$. This is also a free nilpotent group of class c. Also, $\gamma_i W / \gamma_{i+1} W$ is free abelian on the i-step basic commutators of elements of $S \setminus T$, so we can identify it with the subgroup of $\gamma_i N_{c,n} / \gamma_{i+1} N_{c,n}$ generated by images of the i-step basic commutators of elements of S. In particular, $\gamma_i W / \gamma_{i+1} W$ is a subgroup of $\langle S_i \setminus T_i \rangle \subseteq \gamma_i V / \gamma_{i+1} V$.

Since f is an S-code of distance w, $f|_W$ is surjective. In particular, the restriction $f|_{\gamma_i W} \colon \gamma_i W \to \gamma_i G \cap H$ is surjective, since every element of $\gamma_i G$ is an i-step commutator of elements of G, hence an i-step commutator of elements in $f(W)$, and hence itself in $f(\gamma_i W)$. This means $(f|_W)_i$, the map $\gamma_i W/\gamma_{i+1} W \to \gamma_i G/\gamma_{i+1} G$ induced by $f|_W$, is surjective. But since $\gamma_i W/\gamma_{i+1} W$ is a subgroup of $\langle S_i \setminus T_i \rangle \subseteq \gamma_i N_{c,n}/\gamma_{i+1} N_{c,n}$, the induced map $f|_{\langle S_i \setminus T_i \rangle} \colon \langle S_i \setminus T_i \rangle \subseteq \gamma_i N_{c,n}/\gamma_{i+1} N_{c,n} \to \gamma_i G/\gamma_{i+1} G$ is also surjective.

Hence, f_i is a code of distance w/i. □

Finally, we use the quotient sequences formalism to prove that a random vector in Maltsev coordinates maps to something close to uniform under a code.

Proposition 4.8. *Let G be a finite nilpotent group of nilpotency class c. Let $\varepsilon > 0$ and $\delta > 0$ be real numbers. Let n be a positive integer. Let $S = S_1 \sqcup S_2 \sqcup \cdots = \{u_1, \ldots, u_m\}$ be a Maltsev basis for $N_{c,n}$ consisting of commutators in the free generators $u_1, \ldots, u_n$, where S_i consists of i-step commutators. Let $X_1, \ldots, X_m$ be ε-balanced random integers valued in $\mathbb{Z}/k\mathbb{Z}$, where k is a multiple of the exponent of G, and write $X = (X_1, \ldots, X_m)_M$. Let $f \in \mathrm{Hom}(N_{c,n}, G)$ be a code of distance δn with respect to $S_1 = \{u_1, \ldots, u_n\}$, and let g be an element of G.*

$$\left|\mathbb{P}(f(X) = g) - \frac{1}{|G|}\right| \leq |G| \exp(-\varepsilon\delta n/2c|G|^4).$$

Proof. Let μ_i be the distribution of $u_i^{X_i}$ in $N_{c,n}$, then X has distribution $\mu_1 * \ldots \mu_m$, and $f(X)$ has distribution $f_*\mu_1 * \cdots * f_*\mu_m$. The goal is to apply Theorem 4.4 to this convolution. The bulk of the work for this proof goes into constructing an appropriate complete quotient sequence to apply Theorem 4.4.

There is a natural complete quotient sequence for G of length c given by

$$Q \colon G = G_0 \to G/\gamma_c G \to G/\gamma_{c-1} G \to \cdots \to G/\gamma_2 G \to G/\gamma_1 G = \{e\},$$

since $\gamma_i G$ is a normal subgroup of G containing $\gamma_{i+1} G$ and therefore projects to a normal subgroup of the quotient $G/\gamma_{i+1} G$.

The image in $G/\gamma_{i+1} G$ of the i-step commutator subgroup of G is central, since by taking the quotient by all $i+1$-step commutators, we force i-step commutators to commute with everything in $G/\gamma_{i+1} G$. In particular, for any i-step commutator x in G, the cyclic subgroup generated by its image in $G/\gamma_{i+1} G$ is normal.

For each i, let $T_i = \{g \in \gamma_i G \mid \#(f^{-1}(g) \cap S_i) \geq \delta n/i|G|\}$. Note that $\#f^{-1}(T_i) > |S_i| - \delta n/i$, since there are fewer than $\delta n/i$ elements of S_i whose image is not in T_i. Since f is a code of distance δn, the induced map $f_i \colon \gamma_i N_{c,n}/\gamma_{i+1} N_{c,n} \to \gamma_i G/\gamma_{i+1} G$ is a code of width $\delta n/i$. Hence, the restriction of f_i to the image of $f^{-1}(T_i)$ is still surjective. In other words, the images of the elements of T_i generate $\gamma_i G/\gamma_{i+1} G$ inside $G/\gamma_{i+1} G$.

Hence, by Lemma 4.3, we can refine Q using the cyclic subgroups $\langle g \rangle \trianglelefteq G/\gamma_{i+1} G$ for $g \in T_i$ for each i. This gives a quotient sequence Q' for G. By construction of the T_i, we have $|I_j| \geq \delta n/i|G|$ for each j.

Now applying Theorem 4.4, we get

$$\begin{aligned}\left|\mathbb{P}(f(X) = g) - \frac{1}{|G|}\right| &\leq ||f_*\mu_1 * \cdots * f_*\mu_m - \pi||_{L^2} \\ &\leq \sqrt{\sum_{j=1}^{|T_1 \cup \cdots \cup T_c|} \left(\prod_{i \in I_j} \sigma_i^2\right)} \leq \sum_{j=1}^{|T_1 \cup \cdots \cup T_c|} \left(\prod_{i \in I_j} \sigma_i\right).\end{aligned}$$

By Lemma 3.7 and the fact that $|I_j| \geq \delta_n/i|G| \geq \delta n/c|G|$, we get

$$\begin{aligned}\sum_{j=1}^{|T_1 \cup \cdots \cup T_c|} \left(\prod_{i \in I_j} \sigma_i\right) &\leq \sum_{j=1}^{|T_1 \cup \cdots \cup T_c|} \exp(-\varepsilon\delta n/2c|G|^4) \\ &\leq |G| \exp(-\varepsilon\delta n/2c|G|^4),\end{aligned}$$

and the result follows. □

4.3 Future Work

A major direction in our future work is to complete the universality results for nilpotent matrices from the previous section. We are also interested in using the machinery of codes to prove similar results about shorter random walks on the free group. This will probably require some extension of Theorem 2.5 that allows fewer subgroups to be normal. On the side of universality for cokernels of random matrices, future work will involve extending Theorem 1.3 to strictly generalize the result of [NW22], including eliminating the assumption of identical distributions in [NW22].

BIBLIOGRAPHY

[BBH21] Nigel Boston, Michael R. Bush, and Farshid Hajir. “Heuristics for p-Class Towers of Real Quadratic Fields”. In: *Journal of the Institute of Mathematics of Jussieu. JIMJ. Journal de l’Institut de Mathématiques de Jussieu* 20.4 (2021), pp. 1429–1452. ISSN: 1474-7480. DOI: 10.1017/S1474748019000641. arXiv: 1803.04047.

[BE11] Nigel Boston and Jordan S. Ellenberg. “Random pro-*p* groups, braid groups, and random tame Galois groups”. In: *Groups, Geometry, and Dynamics* 5 (2011), pp. 265–280.

[CL83] H. Cohen and H. W. Lenstra, Jr. “Heuristics on class groups of number fields”. In: *Number theory*. Vol. 1068. Lecture Notes in Math. Noordwijkerhout, 1983, pp. 33–62.

[CMZ17] Anthony E. Clement, Stephen Majewicz, and Marcos Zyman. *The theory of nilpotent groups*. Cham, Switzerland: Birkhäuser, 2017. ISBN: 978-3-319-66211-4.

[FW89] Eduardo Friedman and Lawrence C. Washington. “On the distribution of divisor class groups of curves over a finite field”. In: *Proceedings of the International Number Theory Conference held at Université Laval, July 5-18, 1987*. Ed. by Jean M. de Koninck and Claude Levesque. Berlin, New York: De Gruyter, 1989, pp. 227–239. ISBN: 9783110852790. DOI: doi:10.1515/9783110852790.227. URL: https://doi.org/10.1515/9783110852790.227.

[LW20] Yuan Liu and Melanie Matchett Wood. “The free group on n generators modulo n + u random relations as n goes to infinity”. In: *Journal für die reine und angewandte Mathematik (Crelles Journal)* 2020.762 (2020), pp. 123–166. DOI: doi:10.1515/crelle-2018-0025.

[LWZ19] Yuan Liu, Melanie Matchett Wood, and David Zureick-Brown. “A Predicted Distribution for Galois Groups of Maximal Unramified Extensions”. In: *arXiv:1907.05002 [math]* (July 2019). arXiv: 1907.05002 [math].

[NW22] Hoi H. Nguyen and Melanie Matchett Wood. “Random Integral Matrices: Universality of Surjectivity and the Cokernel”. In: *Inventiones mathematicae* 228.1 (Apr. 2022), pp. 1–76. ISSN: 1432-1297. DOI: 10.1007/s00222-021-01082-w.

[Sal04] Laurent Saloff-Coste. “Random Walks on Finite Groups”. en. In: *Probability on Discrete Structures*. Ed. by A.-S. Sznitman, S. R. S. Varadhan, and Harry Kesten. Vol. 110. Series Title: Encyclopaedia of Mathematical Sciences. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 263–346. DOI: 10.1007/978-3-662-09444-0_5. URL: http:

//link.springer.com/10.1007/978-3-662-09444-0_5 (visited on 08/22/2022).

[Saw20] Will Sawin. “Identifying measures on non-abelian groups and modules by their moments via reduction to a local problem”. In: *arXiv: Number Theory* (2020).

[SW22] Will Sawin and Melanie Matchett Wood. *The moment problem for random objects in a category*. 2022. arXiv: 2210.06279 [math.PR].

[SZ07] Laurent Saloff-Coste and Jesse Zúñiga. “Convergence of some time inhomogeneous Markov chains via spectral techniques”. In: *Stochastic Processes and their Applications* 117.8 (2007), pp. 961–979. ISSN: 0304-4149. DOI: https://doi.org/10.1016/j.spa.2006.11.004. URL: https://www.sciencedirect.com/science/article/pii/S0304414906001736.

[Woo14] Melanie Matchett Wood. “The distribution of sandpile groups of random graphs”. In: *Journal of the American Mathematical Society* 30.4 (2014).

[Woo19] Melanie Matchett Wood. “Random integral matrices and the Cohen-Lenstra heuristics”. In: *American Journal of Mathematics* 141.2 (2019), pp. 383–398.

[Woo23] Melanie Matchett Wood. *Probability Theory for Random Groups Arising in Number Theory*. Jan. 2023. DOI: 10.48550/arXiv.2301.09687. arXiv: arXiv:2301.09687. (Visited on 04/16/2023).