

Abelian groups are polynomially stable

Oren Becker ^{*} Jonathan Mosheiff [†]

Abstract

In recent years, there has been a considerable amount of interest in stability of equations and their corresponding groups. Here, we initiate the systematic study of the quantitative aspect of this theory. We develop a novel method, inspired by the Ornstein-Weiss quasi-tiling technique, to prove that abelian groups are polynomially stable with respect to permutations, under the normalized Hamming metrics on the groups $\text{Sym}(n)$. In particular, this means that there exists $D \geq 1$ such that for $A, B \in \text{Sym}(n)$, if AB is δ -close to BA , then A and B are ϵ -close to a commuting pair of permutations, where $\epsilon \leq O(\delta^{1/D})$. We also observe a property-testing reformulation of this result, yielding efficient testers for certain permutation properties.

1 Introduction

We begin with an informal presentation of a general framework that originated in [9]. Fix the normalized Hamming metric over $\text{Sym}(n)$ as a measure of proximity between permutations. Let $(\sigma_1, \dots, \sigma_s)$ be permutations in $\text{Sym}(n)$. Suppose that $(\sigma_1, \dots, \sigma_s)$ “almost” satisfy a given finite system of equations E . Is it necessarily the case that we can “slightly” modify each σ_i to some $\tau_i \in \text{Sym}(n)$, so that (τ_1, \dots, τ_s) satisfy E *exactly*? The answer depends on the system E . In the case where E is comprised of the single equation $XY = YX$, it was shown in [1] that the answer is positive. Simply put, “almost commuting permutations are close to commuting permutations”, where σ_1, σ_2 are said to “almost commute” if the permutations $\sigma_1\sigma_2$ and $\sigma_2\sigma_1$ are close. On the other hand, for E that is the single equation $XY^2 = Y^3X$ the answer is negative [9]. In the language of [9], “ $XY = YX$ is stable, whereas $XY^2 = Y^3X$ is not”.

The main novelty of this paper is that we provide the first results in this area which are *quantitative* and *algorithmic*. Let us illustrate this for the system $E = \{XY = YX\}$. We seek statements of the form “if the distance between $\sigma_1\sigma_2$ and $\sigma_2\sigma_1$ is δ , then there are two commuting permutations τ_1, τ_2 such that

^{*}Department of Mathematics, Hebrew University, Jerusalem, Israel. e-mail: oren.becker@mail.huji.ac.il. Supported by the ERC.

[†]Department of Computer Science, Hebrew University, Jerusalem, Israel. e-mail: yonatanm@cs.huji.ac.il. Supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities.

τ_i is at most $f(\delta)$ away from σ_i for $i = 1, 2$ ". We stress that $f = f_E$, called the *stability rate* of E , depends solely on δ but not σ_1, σ_2 or n . Previously, it was only shown that $\lim_{\delta \rightarrow 0} f(\delta) = 0$, but nothing was said about the rate at which f tends to zero with δ . This is precisely the kind of results that we achieve here. By stating that our results are algorithmic, we mean that they provide an explicit transformation of σ_1, σ_2 into τ_1, τ_2 .

More generally, one may consider the set of equations

$$E_{\text{comm}}^d = \{X_i X_j = X_j X_i \mid 1 \leq i < j \leq d\}$$

and its stability rate function f_d . The paper [1] shows that E_{comm}^d is stable, or, in terms of the framework of the present paper, that $\lim_{\delta \rightarrow 0} f_d(\delta) = 0$ for all $d \in \mathbb{N}$. We give a stronger result by proving, in an algorithmic manner, that

$$f_d(\delta) \leq O\left(\delta^{\frac{1}{D}}\right),$$

where $D = D(d) \geq d$ is an explicitly given constant. We say that " E_{comm}^d is polynomially stable (with degree at most D)". Our proof employs novel elementary methods, and does not rely on previous results concerning stability. Furthermore, we also prove that

$$f_d(\delta) \geq \Omega\left(\delta^{\frac{1}{d}}\right).$$

It remains an open problem to close the gap between the bounds on $f_d(\delta)$.

Following [9], the basic observation is that the stability of a set E of equations is best studied in terms of the group presented by taking E as relators, and that stability is a *group invariant* [1]. We show that the stability rate is a group invariant as well, and then use properties of the group \mathbb{Z}^d to study the quantitative stability of the equations E_{comm}^d . More specifically, our proof of the aforementioned upper bound is based on a tiling procedure, in the spirit of Ornstein-Weiss quasi-tiling for amenable groups [17] (in the setting of group actions which are not necessarily free). The fact that we work with abelian groups (rather than more general amenable groups) enables a highly efficient tiling procedure via an original application of *reduction theory* of lattices in \mathbb{Z}^d . Our use of reduction theory is made through a theorem of Lagarias, Lenstra and Schnorr [14] regarding Korkin-Zolotarev bases [15].

Finally, we examine a connection to the topic of *property testing* in computer science by rephrasing some of the above notions in terms of a certain *canonical testing algorithm* (this connection, in the non-quantitative setting, is the subject of [3]). In particular, we show that our result yields an *efficient* algorithm to test whether a given tuple of permutations in $\text{Sym}(n)$ satisfies the equations E_{comm}^d .

The rest of the introduction is organized as follows. Section 1.1 paraphrases definitions and results from the theory of stability in permutations. In Section 1.2, we define the new, more delicate, notion of quantitative stability, and state our main theorems. Section 1.3 explores the connection between stability and property testing, and Section 1.4 discusses previous work.

1.1 A review of stability in permutations

For a set S , we write \mathbb{F}_S for the free group based on S . We also write $S^{-1} \subseteq \mathbb{F}_S$ for the set of inverses of elements of S and $S^\pm = S \cup S^{-1}$. Fix a finite set of variables $S = \{s_1, \dots, s_m\}$. An S -assignment is a function $\Phi : S \rightarrow \text{Sym}(n)$ for some positive integer n , i.e., we assign permutations to the variables, all of which are in $\text{Sym}(n)$ for the same n . We naturally extend Φ to the domain \mathbb{F}_S via $\Phi(s^{-1}) = \Phi(s)^{-1}$ if $s \in S$, and $\Phi(w_1 \cdots w_t) = \Phi(w_1) \cdots \Phi(w_t)$ if $w_1, \dots, w_t \in S^\pm$. In other words, an S -assignment Φ is also regarded as a homomorphism from \mathbb{F}_S to a finite symmetric group. Hence, Φ naturally describes a group action of \mathbb{F}_S on the finite set $[n] = \{1, \dots, n\}$.

An S -equation-set (or equation-set over the set S of variables) is a finite¹ set $E \subseteq \mathbb{F}_S$. An assignment Φ is said to be an E -solution if $\Phi(w)$ is the identity permutation $1 = 1_{\text{Sym}(n)}$ for every $w \in E$. Equivalently, Φ is an E -solution if, when regarded as a homomorphism $\mathbb{F}_S \rightarrow \text{Sym}(n)$, it factors through the quotient map $\pi : \mathbb{F}_S \rightarrow \mathbb{F}_S / \langle\langle E \rangle\rangle$. Namely, if $\Phi = g \circ \pi$ for some homomorphism $g : \mathbb{F}_S / \langle\langle E \rangle\rangle \rightarrow \text{Sym}(n)$. Here and throughout, $\langle\langle E \rangle\rangle$ denotes the normal closure of a subset E of a group (the containing group will always be clear from context).

Example 1.1. To relate these definitions to the canonical example of almost commuting pairs of permutation, let $S = \{s_1, s_2\}$ and $E = \{s_1 s_2 s_1^{-1} s_2^{-1}\}$, and consider an S -assignment Φ . Note that $\Phi(s_1)$ and $\Phi(s_2)$ commute if and only if Φ is an E -solution. It may be helpful to keep this example in mind when reading up to the end of Section 1.2.

Definition 1.2. For $n \in \mathbb{N}$, the *normalized Hamming distance* d_n on $\text{Sym}(n)$ is defined by $d_n(\sigma_1, \sigma_2) = \frac{1}{n} |\{x \in [n] \mid \sigma_1(x) \neq \sigma_2(x)\}|$ for $\sigma_1, \sigma_2 \in \text{Sym}(n)$. Also, if Φ, Ψ are assignments over some finite set of variables S , we define $d_n(\Phi, \Psi) = \sum_{s \in S} d_n(\Phi(s), \Psi(s))$.

Note that d_n is bi-invariant, namely, $d_n(\tau \cdot \sigma_1 \cdot \nu, \tau \cdot \sigma_2 \cdot \nu) = d_n(\sigma_1, \sigma_2)$ for every $\tau, \nu \in \text{Sym}(n)$. In particular, $d_n(\sigma_1, \sigma_2) = d_n(\sigma_1 \cdot \sigma_2^{-1}, 1)$.

Given an S -equation-set E and an S -assignment Φ , one may ask how close Φ is to being an E -solution. This notion can be interpreted in two different ways. Locally, one can count the points in $[n] = \{1, \dots, n\}$ on which Φ violates the equation-set E . Taking a global view, one measures how much work, that is, change of permutation entries, is needed to “correct” Φ into an E -solution. We proceed to define the *local defect* and the *global defect* of an assignment.

Definition 1.3. Fix a finite set of variables S and an S -equation-set E , and let $\Phi : S \rightarrow \text{Sym}(n)$ be an S -assignment.

- (i) The *local defect* of Φ with respect to E is

$$L_E(\Phi) = \sum_{w \in E} d_n(\Phi(w), 1) = \frac{1}{n} \sum_{w \in E} |\{x \in [n] \mid \Phi(w)(x) \neq x\}|.$$

¹See [4] for an extension of these definitions to infinite equation-sets.

(ii) The *global defect* of Φ with respect to E is

$$G_E(\Phi) = \min \{ |d_n(\Phi, \Psi)| \mid \Psi : S \rightarrow \text{Sym}(n) \text{ is an } E\text{-solution} \}.$$

Note that $L_E(\Phi)$ and $G_E(\Phi)$ are in the ranges $[0, |E|]$ and $[0, |S|]$, respectively. Also, the conditions (i) Φ is an E -solution, (ii) $L_E(\Phi) = 0$ and (iii) $G_E(\Phi) = 0$, are all equivalent.

We wish to study the possible divergence between these two types of defect. It is not hard to see (Lemma A.1) that $G_E(\Phi) \geq \Omega(L_E(\Phi))$. Indeed, changing a single entry in one permutation $\Phi(s)$, resolves, at best, the violation of E at a constant number of points $x \in [n]$. In this work we seek a reverse inequality, i.e., an upper bound on the global defect in terms of the local defect.

Definition 1.4. Let E be an S -equation-set. We define its *stability rate* $\text{SR}_E : (0, |E|] \rightarrow [0, \infty)$ by

$$\text{SR}_E(\delta) = \sup \{ G_E(\Phi) \mid \Phi : S \rightarrow \text{Sym}(n), \text{ where } n \in \mathbb{N} \text{ and } L_E(\Phi) \leq \delta \}.$$

Note that $\text{SR}_E(\delta)$ is a monotone nondecreasing function.

Definition 1.5. [9] If $\lim_{\delta \rightarrow 0} \text{SR}_E(\delta) = 0$ then E is said to be *stable (in permutations)*.

The theory of equation-sets in permutations is related to group theory via a key observation, given below as Proposition 1.7.

Definition 1.6. Let E_1 and E_2 be equation-sets over the respective finite sets of variables S_1 and S_2 . If the groups $\mathbb{F}_{S_1}/\langle\langle E_1 \rangle\rangle$ and $\mathbb{F}_{S_2}/\langle\langle E_2 \rangle\rangle$ are isomorphic then E_1 and E_2 are called *equivalent*.

Proposition 1.7. [1] *Let E_1 and E_2 be equivalent equation-sets. Then E_1 is stable if and only if E_2 is stable.*

Proposition 1.7 allows us to regard stability as a group invariant:

Definition 1.8. [1] Let Γ be a finitely presented group. That is, $\Gamma \cong \mathbb{F}_S/\langle\langle E \rangle\rangle$ for some finite sets S and $E \subseteq \mathbb{F}_S$. Then, Γ is called *stable* if E is a stable equation-set.

This definition enables us to apply the properties of a group in order to study the stability of its defining set of equations. See Section 1.4 for previous results obtained by this method. The following result is of particular interest to us in the context of this work.

Theorem 1.9. [1] *Every finitely generated abelian group is stable.*

The proof of Theorem 1.9 in [1] is not algorithmic. That is, it does not describe an explicit transformation that maps an assignment with small local defect to a nearby E -solution. With some work, it is possible in principle to extract such an algorithm from the idea presented in [1], by looking further into the proof a theorem of Elek and Szabo [7] used there. However, such an approach would perform poorly in the quantitative sense described below.

1.2 Quantitative stability

We turn to discuss the main new definitions introduced in this work, which deals with quantitative stability. The basis for quantitative stability is a stronger version of Proposition 1.7, given in Proposition 1.11 below. Namely, we show that stability can be refined, yet remain a group invariant, by considering the *rate* at which SR_E converges to 0 as $\delta \rightarrow 0$. We now make this claim precise.

Definition 1.10. Let $F_1, F_2 : (0, |E|] \rightarrow [0, \infty)$ be monotone nondecreasing functions. Write $F_1 \sim F_2$ if $F_1(\delta) \leq F_2(C\delta) + C\delta$ and $F_2(\delta) \leq F_1(C\delta) + C\delta$ for some $C > 0$. Let $[F_1]$ denote the class of F_1 with regard to this equivalence relation.

The reason for the introduction of the $C\delta$ summand in the above definition is elaborated upon in Remark 2.4.

Proposition 1.11. *For every equivalent pair of equation-sets E_1 and E_2 we have $\text{SR}_{E_1} \sim \text{SR}_{E_2}$.*

We prove Proposition 1.11 in Section 2. This proposition allows us to define the *stability rate* of a group through the stability rate of a corresponding equation-set (Definition 1.4).

Definition 1.12. Let E be an S -equation-set. The *stability rate* SR_Γ of the group $\Gamma = \mathbb{F}_S/\langle\langle E \rangle\rangle$ is the equivalence class $[\text{SR}_E]$.

Our goal in this paper is to show that for an abelian group Γ , not only does the stability rate converge to 0 as $\delta \rightarrow 0$, but this convergence is *fast*. This claim can be made precise as follows.

Definition 1.13. Let $F : (0, |E|] \rightarrow [0, \infty)$. We define the *degree* of F by

$$\text{deg}(F) = \inf \left\{ k \geq 1 \mid F(\delta) \leq_{\delta \rightarrow 0} O\left(\delta^{\frac{1}{k}}\right) \right\}. \quad (1.1)$$

It is possible that $\text{deg}(F) = \infty$. Also, let $\text{deg}([F]) = \text{deg}(F)$.

Remark 1.14. Note that $\text{deg}([F])$ is well-defined. Indeed, if $F_1 \sim F_2$ and $F_1(\delta) \leq_{\delta \rightarrow 0} O\left(\delta^{\frac{1}{m}}\right)$ for $m \geq 1$, then $F_2(\delta) \leq F_1(C\delta) + C\delta \leq_{\delta \rightarrow 0} O\left(\delta^{\frac{1}{m}}\right)$ as well.

Definition 1.15. In the notation of Definition 1.12, the *degree of polynomial stability* of both E and Γ is defined to be $D = \text{deg}([\text{SR}_E]) = \text{deg}(\text{SR}_\Gamma)$. If $D < \infty$, we say that E and Γ are *polynomially stable*.

We can now state our main theorem.

Theorem 1.16. *Every finitely generated abelian group is polynomially stable.*

Notably, Theorem 1.16 applies to *commutator equation-sets*, namely equation-sets of the form

$$E_{\text{comm}}^d = \{s_i s_j s_i^{-1} s_j^{-1} \mid 1 \leq i < j \leq d\} \quad (1.2)$$

over the variables $\{s_1, \dots, s_d\}$ where $d \geq 0$. Indeed, note that $\mathbb{F}_d/\langle\langle E_{\text{comm}}^d \rangle\rangle \cong \mathbb{Z}^d$, so E_{comm}^d is polynomially stable. In particular, $d = 2$ yields the example of commuting *pairs* of permutations from the beginning of the introduction.

We complement Theorem 1.16 with a lower bound on the degree of polynomial stability.

Theorem 1.17. *For all $d \in \mathbb{N}$, the group \mathbb{Z}^d has degree of polynomial stability at least d .*

We note that, having shown that an abelian group Γ is polynomially stable, we are left with the more delicate question of its *degree of polynomial stability*. Restricting attention to a free abelian group \mathbb{Z}^d , the proof of Theorem 1.16 yields an upper bound on this degree that grows exponentially in d (see Equation (4.2)). A remaining open problem is to close the large gap between this upper bound, and the lower bound of Theorem 1.17.

We diverge from the practice of providing an outline of the proofs in the introduction since such an outline requires a geometric formulation of stability, which is developed in Section 3. Theorem 1.16 is proved in Section 4, and its proof is outlined in Section 4.1. A reader who is interested in the quickest route to understanding this outline may skip Section 2. Theorem 1.17 is the subject of Section 5.

1.3 Application to property testing

The notions of stability and quantitative stability have a natural interpretation in terms of *property testing* (For more on this connection see [3]. For background on property testing see [10, 18]).

Definition 1.18. Fix a nonempty equation-set E over the finite set of variables S . A *tester* for E is an algorithm which takes an assignment $\Phi : S \rightarrow \text{Sym}(n)$ as input, queries the permutations $\{\Phi(s)\}_{s \in S}$ at a constant (in particular, independent of n) number of entries among $1, \dots, n$, and decides whether Φ is an E -solution. The algorithm must satisfy the following:

- (I) If Φ is an E -solution, the algorithm always *accepts*.
- (II) For some fixed function $\delta : (0, \infty) \times \mathbb{N} \rightarrow (0, 1]$, if Φ is not an E -solution, then the algorithm *rejects* with probability at least $\delta(G_E(\Phi), n)$. This function δ is called the *detection probability* of the tester.

The precise term in the literature for this notion is an “adaptive proximity-oblivious tester with one-sided error and constant query complexity” (see, [10] Definition 1.7).

The *canonical tester* \mathcal{N}_E for E samples $x \in [n]$ and $w \in E$ uniformly at random from their respective sets. It accepts if $\Phi(w)(x) = x$ and rejects otherwise. Clearly, if Φ is an E -solution then \mathcal{N}_E always accepts. If Φ is not an E -solution, then \mathcal{N}_E rejects with probability $\frac{1}{|E|} \cdot L_E(\Phi)$. Generally, it is desirable for the detection probability function $\delta(\epsilon, n)$ to be *uniform*, i.e., depend

only on $\epsilon = G_E(\Phi)$ and not on n . In order to show that this is the case for \mathcal{N}_E , we must consider the function

$$\begin{aligned} \delta(\epsilon) &= \inf \left\{ \frac{1}{|E|} \cdot L_E(\Phi) \mid \Phi \text{ is an } S\text{-assignment and } G_E(\Phi) \geq \epsilon \right\} \\ &= \inf \{ \delta \mid \text{SR}_E(\delta) \geq \epsilon \} . \end{aligned}$$

The tester \mathcal{N}_E satisfies Condition 1.18(II) if and only if $\delta(\epsilon)$ is positive for every $\epsilon > 0$. This is equivalent to the condition that $\lim_{\delta \rightarrow 0} \text{SR}_E(\delta) = 0$. Hence, \mathcal{N}_E admits a uniform detection probability function if and only if E is stable.

Furthermore, the smaller the stability rate of E , the larger δ is. In particular, if E is polynomially stable with degree of polynomial stability D , then $\delta(\epsilon)$ is bounded from below by $\Omega(\epsilon^D)$. Therefore, due to Theorem 1.16, the canonical tester \mathcal{N}_E has detection probability polynomial in ϵ and uniform in n , whenever $\mathbb{F}_S/\langle\langle E \rangle\rangle$ is abelian.

In a somewhat weaker formulation of property testing (see [10], Definition 1.6), the tester is only required to distinguish between the cases $G_E(\Phi) = 0$ and $G_E(\Phi) > \epsilon$, where $\epsilon > 0$ is given as input. In this formulation, the detection probability is required to be larger than some constant, say $\frac{1}{2}$, and one seeks to minimize the number of queries. Note that, for every E , the canonical tester \mathcal{N}_E can be used to build a tester $\tilde{\mathcal{N}}_E$, satisfying this weaker formulation: Given $\epsilon > 0$, and an input Φ , the tester $\tilde{\mathcal{N}}_E$ runs \mathcal{N}_E on Φ repeatedly for $\log_{1-\delta(\epsilon,n)} \frac{1}{2} = \Theta\left(\frac{1}{\delta(\epsilon,n)}\right)$ independent iterations, and accepts only if \mathcal{N}_E accepts in all iterations. Hence, the resulting tester $\tilde{\mathcal{N}}_E$ performs $\Theta\left(\frac{1}{\delta(\epsilon,n)}\right)$ queries. In particular, Theorem 1.16 shows that for E such that $\mathbb{F}_S/\langle\langle E \rangle\rangle$ is abelian, the tester $\tilde{\mathcal{N}}_E$ is *efficient*. Namely, it has constant detection probability, and its number of queries is polynomial in $\frac{1}{\epsilon}$ and does not depend on n . No such tester was previously known.

1.4 Previous work

The general question of whether almost-solutions are close to solutions, in various contexts, was suggested by Ulam (see [19], Chapter VI). The most studied question of this sort is whether almost-commuting *matrices* are close to commuting matrices, and the answer depends on the chosen matrix norm and on which kind of matrices is considered (e.g. self-adjoint, unitary, etc.). See the introduction of [1] for a short survey, and [5, 6, 11, 12] for some newer works. In this context, some quantitative results are already known [8, 13]. The question of (non-quantitative) stability in *permutations*, under the normalized Hamming metric, was initiated in [9] and developed further in [1]. The former paper proves that finite groups are stable (see our Proposition A.4 for a quantitative version), and the latter proves that abelian groups are stable. Both papers provide examples of non-stable groups as well, and relate stability in permutations to the notion of sofic groups. These results are generalized in [4], which provides a characterization of stability in permutations, among amenable groups, in terms

of their invariant random subgroups. On the other side of the spectrum (compared to amenable groups), [2] proves that infinite groups with Property (T) are never stable in permutations, and consequently suggests some weaker forms of stability.

2 The stability rate is a group invariant

In this section we prove Proposition 1.11. We start with Lemma 2.1, which formalizes the claim that if two S -assignments Φ and Ψ are close to each other, then so are the permutations $\Phi(w)$ and $\Psi(w)$, provided that $w \in \mathbb{F}_S$ is a short word. It is helpful to consider $w = w_1 \cdots w_t$ ($w_i \in S^\pm$) as a sequence of directions, namely, $\Phi(w)(x)$ is the endpoint of the path that starts at x , moves in the direction w_t to $\Phi(w_t)(x)$, continues to $\Phi(w_{t-1}w_t)(x)$, and so on. The immediate observation behind Lemma 2.1, is that as long as this path moves only along “nice” edges, i.e., edges on which Φ and Ψ agree, it is guaranteed that $\Phi(w)(x) = \Psi(w)(x)$. A similar idea is then used to prove Lemma 2.2 as well.

Lemma 2.1. *Let E be an S -equation-set, let $\Phi, \Psi : S \rightarrow \text{Sym}(n)$ be S -assignments, and let $w = w_1 \cdots w_t \in \mathbb{F}_S$ where $w_i \in S^\pm$. Then, $d_n(\Phi(w), \Psi(w)) \leq t \cdot d_n(\Phi, \Psi)$.*

Proof. Let \bar{w}_i denote the suffix $w_i \cdots w_t$. By the bi-invariance of d_n and the triangle inequality,

$$\begin{aligned} d_n(\Phi(w), \Psi(w)) &= d_n((\Psi(w))^{-1} \cdot \Phi(w), 1) \\ &\leq \sum_{i=1}^t d_n((\Psi(\bar{w}_i))^{-1} \cdot \Phi(\bar{w}_i), \Psi(\bar{w}_{i+1}))^{-1} \cdot \Phi(\bar{w}_{i+1})) \\ &= \sum_{i=1}^t d_n((\Psi(w_i))^{-1} \cdot \Phi(w_i), 1) = \sum_{i=1}^t d_n(\Phi(w_i), \Psi(w_i)). \end{aligned}$$

Note that each term of the right hand side is at most $d_n(\Phi, \Psi)$. Here, if w_i is the inverse of a generator, we used the fact that $d_n(\Phi(w_i), \Psi(w_i)) = d_n(\Phi(w_i^{-1}), \Psi(w_i^{-1}))$. The lemma follows. \square

Lemma 2.2. *Fix an S -equation-set $E \subseteq \mathbb{F}_S$, and a word $w \in \langle\langle E \rangle\rangle$, written as $w = u_1 q_1 u_1^{-1} u_2 q_2 u_2^{-1} \cdots u_t q_t u_t^{-1}$, where $u_i \in \mathbb{F}_S$ and each q_i is an element of E or its inverse. Then, for every S -assignment $\Phi : S \rightarrow \text{Sym}(n)$,*

$$d_n(\Phi(w), 1) \leq L_E(\Phi) \cdot t.$$

Proof. We have

$$\begin{aligned}
 d_n(\Phi(w), 1) &= d_n(\Phi(u_1 q_1 u_1^{-1} u_2 q_2 u_2^{-1} \cdots u_t q_t u_t^{-1}), 1) \\
 &\leq \sum_{i=1}^t d_n(\Phi(u_i q_i u_i^{-1} \cdots u_t q_t u_t^{-1}), \Phi(u_{i+1} q_{i+1} u_{i+1}^{-1} \cdots u_t q_t u_t^{-1})) \\
 &= \sum_{i=1}^t d_n(\Phi(u_i q_i u_i^{-1}), 1) \\
 &= \sum_{i=1}^t d_n(\Phi(q_i), 1) \\
 &\leq \sum_{i=1}^t \sum_{w \in E} d_n(\Phi(w), 1) \\
 &= t \cdot L_E(\Phi) .
 \end{aligned}$$

□

The following corollary of Lemma 2.2 will also be useful.

Lemma 2.3. *Fix an S -equation-set E and a homomorphism $\lambda : \mathbb{F}_S \rightarrow \mathbb{F}_S$ such that w and $\lambda(w)$ belong to the same coset in $\mathbb{F}_S / \langle\langle E \rangle\rangle$ for each $w \in \mathbb{F}_S$. Then, there exists a positive $c = c(S, E, \lambda)$ such that for every S -assignment $\Phi : S \rightarrow \text{Sym}(n)$, we have $d_n(\Phi, \Phi \circ \lambda) \leq c \cdot L_E(\Phi)$, where we regard $\Phi \circ \lambda$ as an S -assignment by restricting its domain from \mathbb{F}_S to S .*

Proof. Note that

$$d_n(\Phi, \Phi \circ \lambda) = \sum_{s \in S} d_n(\Phi(s), \Phi(\lambda(s))) = \sum_{s \in S} d_n(1, \Phi(\lambda(s) \cdot s^{-1})) .$$

Since $\lambda(s) \cdot s^{-1} \in \langle\langle E \rangle\rangle$, it follows from Lemma 2.2 that the s -term of this sum is at most $O(L_E(\Phi))$, where the implied constant depends on s , E and λ . Hence, $d_n(\Phi, \Phi \circ \lambda) \leq O(L_E(\Phi))$. □

We turn to prove Proposition 1.11.

Proof of Proposition 1.11. In the course of the proof, when a function whose domain is \mathbb{F}_{S_1} or \mathbb{F}_{S_2} appears where a function whose domain is S_1 or S_2 is expected, the function should be regarded as its respective restriction (for example, when measuring distances between assignments).

Let E_1 and E_2 be equivalent equation-sets over the respective finite sets of variables S_1 and S_2 . By symmetry, it is enough to prove that $\text{SR}_{E_1}(\delta) \leq \text{SR}_{E_2}(C\delta) + C\delta$ for some $C = C(E_1, E_2)$. Equivalently, we need to show that $G_{E_1}(\Phi_1) \leq \text{SR}_{E_2}(C\delta) + C\delta$ for any given $\delta > 0$ and S_1 -assignment $\Phi_1 : S_1 \rightarrow \text{Sym}(n)$ with $L_{E_1}(\Phi_1) \leq \delta$.

Our strategy is to “translate” the S_1 -assignment Φ_1 into an S_2 -assignment Φ_2 , find an E_2 -solution Ψ_2 which is close to Φ_2 , and pull back Ψ_2 to an E_1 -solution Ψ_1 . We apply Lemma 2.2 to bound $L_{E_2}(\Phi_2)$, and then use Lemmas 2.1 and 2.3 to control the distance between Φ_1 and Ψ_1 , yielding an upper bound on $G_{E_1}(\Phi_1)$.

We define the machinery needed to map S_1 -assignments to S_2 -assignments and vice versa. Since E_1 and E_2 are equivalent equation-sets, there exists a group isomorphism $\theta : \mathbb{F}_{S_1}/\langle\langle E_1 \rangle\rangle \rightarrow \mathbb{F}_{S_2}/\langle\langle E_2 \rangle\rangle$. Let π_1 denote the quotient map $\mathbb{F}_{S_1} \rightarrow \mathbb{F}_{S_1}/\langle\langle E_1 \rangle\rangle$, and likewise for π_2 . Fix a homomorphism $\lambda_2 : \mathbb{F}_{S_2} \rightarrow \mathbb{F}_{S_1}$ such that $\pi_2 = \theta \circ \pi_1 \circ \lambda_2$. In other words, we choose λ_2 so that the composition of the following chain of morphisms equals π_2 :

$$\mathbb{F}_{S_2} \xrightarrow{\lambda_2} \mathbb{F}_{S_1} \xrightarrow{\pi_1} \mathbb{F}_{S_1}/\langle\langle E_1 \rangle\rangle \xrightarrow{\theta} \mathbb{F}_{S_2}/\langle\langle E_2 \rangle\rangle .$$

Note that such λ_2 exists since $\theta \circ \pi_1$ is surjective. Similarly, fix a homomorphism $\lambda_1 : \mathbb{F}_{S_1} \rightarrow \mathbb{F}_{S_2}$ satisfying $\pi_1 = \theta^{-1} \circ \pi_2 \circ \lambda_1$. From now on, in our use of asymptotic $O(\cdot)$ -notation, we allow the implied constant to depend on S_1 , S_2 , E_1 and E_2 . Since λ_2 and λ_1 have been fixed solely in terms of these four objects, the implied constant is allowed to depend on them as well.

Let $\delta > 0$ and $n \in \mathbb{N}$. Let $\Phi_1 : S_1 \rightarrow \text{Sym}(n)$ be an S_1 -assignment such that $L_{E_1}(\Phi_1) \leq \delta$. Define the S_2 -assignment $\Phi_2 = \Phi_1 \circ \lambda_2 : S_2 \rightarrow \text{Sym}(n)$. We seek to bound its local defect:

$$L_{E_2}(\Phi_2) = \sum_{w \in E_2} d_n(\Phi_2(w), 1) = \sum_{w \in E_2} d_n(\Phi_1(\lambda_2(w)), 1) .$$

Since $\lambda_2(w) \in \langle\langle E_1 \rangle\rangle$ for every $w \in E_2$, it follows from Lemma 2.2, applied to Φ_1 and the word $\lambda_2(w)$, that the w -term in the above sum is at most $O(L_{E_1}(\Phi_1))$. Consequently,

$$L_{E_2}(\Phi_2) \leq C_1 \cdot L_{E_1}(\Phi_1) \leq C_1 \cdot \delta$$

for some positive $C_1 = C_1(E_1, E_2)$. Hence, there is an E_2 -solution $\Psi_2 : S_2 \rightarrow \text{Sym}(n)$ with

$$d_n(\Phi_2, \Psi_2) \leq \text{SR}_{E_2}(L_{E_2}(\Phi_2)) \leq \text{SR}_{E_2}(C_1 \cdot \delta) .$$

Let Ψ_1 be the S_1 -assignment $\Psi_2 \circ \lambda_1 : S_1 \rightarrow \text{Sym}(n)$. Note that Ψ_1 is an E_1 -solution. Thus, $G_{E_1}(\Phi_1) \leq d_n(\Phi_1, \Psi_1)$. We proceed to bound this distance. By the triangle inequality,

$$\begin{aligned} d_n(\Phi_1, \Psi_1) &\leq d_n(\Phi_1, \Phi_2 \circ \lambda_1) + d_n(\Phi_2 \circ \lambda_1, \Psi_1) \\ &= d_n(\Phi_1, \Phi_1 \circ \lambda_2 \circ \lambda_1) + d_n(\Phi_2 \circ \lambda_1, \Psi_2 \circ \lambda_1) . \end{aligned}$$

We turn to bound both terms of the right hand side. For the first term, note that

$$\pi_1 = \theta^{-1} \circ \pi_2 \circ \lambda_1 = \theta^{-1} \circ \theta \circ \pi_1 \circ \lambda_2 \circ \lambda_1 = \pi_1 \circ \lambda_2 \circ \lambda_1 ,$$

and so $\lambda_2 \circ \lambda_1$ satisfies the requirements of Lemma 2.3. Hence, due to this lemma,

$$d_n(\Phi_1, \Phi_1 \circ \lambda_2 \circ \lambda_1) \leq O(L_{E_1}(\Phi_1)) \leq O(\delta) .$$

Turning to the second term,

$$d_n(\Phi_2 \circ \lambda_1, \Psi_2 \circ \lambda_1) = \sum_{s_2 \in S_2} d_n(\Phi_2(\lambda_1(s_2)), \Psi_2(\lambda_1(s_2))).$$

By Lemma 2.1, applied to Φ_2 , Ψ_2 and the word $\lambda_1(s_2)$, the s_2 -term of this sum is upper bounded by $O(d_n(\Phi_2, \Psi_2))$, and so

$$d_n(\Phi_2 \circ \lambda_1, \Psi_2 \circ \lambda_1) \leq O(d_n(\Phi_2, \Psi_2)) \leq O(\text{SR}_{E_2}(C_1 \cdot \delta)) .$$

We conclude that

$$G_E(\Phi_1) \leq d_n(\Phi_1, \Psi_1) \leq O(\text{SR}_{E_2}(C_1 \cdot \delta) + \delta) .$$

□

Remark 2.4. In Proposition A.5, we show that $\text{SR}_E(\delta) \geq \Omega(\delta)$ for every equation-set E which is not empty and not $\{1\}$. When E is \emptyset or $\{1\}$, however, it is clear that $\text{SR}_E \equiv 0$. This is worth noting, since the free group \mathbb{F}_S can be defined by either a trivial equation-set over S , or by a certain nontrivial equation-set over some larger finite set of variables. Two nuances in Definitions 1.10 and 1.13 ensure that the stability rate and degree of \mathbb{F}_S are well-defined despite this phenomenon. The first is the addition of the term $C\delta$ to the inequalities in Definition 1.10, and the second is the restriction $k \geq 1$ in Equation (1.1) in Definition 1.13.

3 Stability and graphs of actions

This section reformulates the notions of stability and stability rate in terms of group actions, and provides basic tools arising from this point of view. Before we begin, a small clarification regarding terminology is in order: When a group homomorphism $\theta : \Lambda_2 \rightarrow \Lambda_1$ is fixed and understood from the context, we regard any given Λ_1 -set X as a Λ_2 -set as well via θ , i.e., for $g_2 \in \Lambda_2$ and $x \in X$, we let $g_2 \cdot x = \theta(g_2) \cdot x$. In most cases in the sequel, Λ_1 is a group generated by a finite set S , $\Lambda_2 = \mathbb{F}_S$ is a free group on S , and θ is the natural quotient map $\mathbb{F}_S \rightarrow \Lambda_1$. So, when a Λ_1 -set X appears where an \mathbb{F}_S -set is expected, we treat X as an \mathbb{F}_S -set in this manner. In some other cases, the role of Λ_2 is taken by a free abelian group, rather than a free group.

3.1 Stability in terms of group actions

Throughout Section 3.1, we fix an equation-set E over the finite set of variables S , and denote $\Gamma = \mathbb{F}_S / \langle\langle E \rangle\rangle$. As mentioned in Section 1.1, an S -assignment $\Phi : S \rightarrow \text{Sym}(n)$ can also be regarded as a group action of \mathbb{F}_S on $[n]$. We write $\mathbb{F}_S(\Phi)$ for the \mathbb{F}_S -set whose set of points is $[n]$, with the group action given by $s \cdot x = \Phi(s)(x)$. We now expand upon this view, rephrasing the definition of local and global defect in terms of group actions, along the same

lines as in Section 3.2 of [4]. This will enable us to prove our main theorems using a geometric approach, focusing on the geometry of an edge-labeled graph representing $\mathbb{F}_S(\Phi)$.

Definition 3.1. Let X be a finite \mathbb{F}_S -set. The *local defect* of X with respect to E is

$$L_E(X) = \frac{1}{|X|} \cdot |\{(x, w) \in X \times E \mid w \cdot x \neq x\}| .$$

It follows directly from the definitions that $L_E(\Phi) = L_E(\mathbb{F}_S(\Phi))$ for an assignment $\Phi : S \rightarrow \text{Sym}(n)$.

We turn to the global defect. Let us first characterize E -solutions through actions. It is not hard to see that $\Phi : S \rightarrow \text{Sym}(n)$ is an E -solution if and only if the action $\mathbb{F}_S \curvearrowright \mathbb{F}_S(\Phi)$ factors through the group Γ , that is, if every two elements $u, v \in \mathbb{F}_S$, belonging to the same coset in $\mathbb{F}_S/\langle\langle E \rangle\rangle$, act on $[n]$ in the same manner. In this case, Γ itself acts on $\mathbb{F}_S(\Phi)$.

We proceed to define a metric of similarity between \mathbb{F}_S -sets.

Definition 3.2. Let X and Y be finite \mathbb{F}_S -sets, where $|X| = |Y| = n$. For a function $f : X \rightarrow Y$, define

$$\|f\|_S = \sum_{s \in S} \frac{1}{n} \cdot |\{(s, x) \in S \times X \mid f(s \cdot x) \neq s \cdot f(x)\}| .$$

Furthermore, define

$$d_S(X, Y) = \min \{ \|f\|_S \mid f : X \rightarrow Y \text{ is a bijection} \} .$$

Note that $d_S(X, Y) = 0$, if and only if X and Y are isomorphic as \mathbb{F}_S -sets. Also, for S -assignments $\Phi, \Psi : S \rightarrow \text{Sym}(n)$ we have $d_n(\Phi, \Psi) = \|\text{id}_{\Phi, \Psi}\|_S$, where $\text{id}_{\Phi, \Psi} : \mathbb{F}_S(\Phi) \rightarrow \mathbb{F}_S(\Psi)$ is the identity map $[n] \rightarrow [n]$.

We use this metric to express the notion of global defect.

Definition 3.3. Let X be a finite \mathbb{F}_S -set. The *global defect* of X with respect to E is

$$G_E(X) = \min \{ d_S(X, Y) \mid Y \text{ is a } \Gamma\text{-set and } |Y| = |X| \} .$$

Proposition 3.4. Let $\Phi : S \rightarrow \text{Sym}(n)$ be an S -assignment. Then,

$$G_E(\Phi) = G_E(\mathbb{F}_S(\Phi)) .$$

Proof. Let $\Psi : S \rightarrow \text{Sym}(n)$ be an E -solution which minimizes $d_n(\Phi, \Psi)$. Let Y be a Γ -set, $|Y| = n$, which minimizes $d_S(\mathbb{F}_S(\Phi), Y)$. We need to show that $d_n(\Phi, \Psi) = d_S(\mathbb{F}_S(\Phi), Y)$. Indeed, on one hand,

$$d_S(\mathbb{F}_S(\Phi), Y) \leq d_S(\mathbb{F}_S(\Phi), \mathbb{F}_S(\Psi)) \leq \|\text{id}_{\Phi, \Psi}\|_S = d_n(\Phi, \Psi) ,$$

where the first inequality follows from the defining property of Y , and the second from the definition of d_S . On the other hand, take a bijection $f : \mathbb{F}_S(\Phi) \rightarrow Y$ for which $d_S(\mathbb{F}_S(\Phi), Y) = \|f\|_S$, and define an S -assignment $\Theta : S \rightarrow \text{Sym}(n)$

by $\Theta(s)(x) = f^{-1}(s \cdot f(x))$. Then, Θ is an E -solution because Y is a Γ -set, and so

$$d_n(\Phi, \Psi) \leq d_n(\Phi, \Theta) = \|\text{id}_{\Phi, \Theta}\|_S = \|f\|_S = d_S(\mathbb{F}_S(\Phi), Y) .$$

□

The above discussion enables us to define the stability rate of E in terms of group actions, as recorded below:

Proposition 3.5. *The stability rate of E is given by*

$$\text{SR}_E(\delta) = \sup\{G_E(X) \mid X \text{ is a finite } \mathbb{F}_S\text{-set and } L_E(X) \leq \delta\} .$$

Definition 3.6. For an \mathbb{F}_S -set X , define the set of E -abiding points in X as

$$X_E = \{x \in X \mid \forall w \in E \ w \cdot x = x\} .$$

Note that for an \mathbb{F}_S -set X , we have

$$\frac{|X \setminus X_E|}{|X|} \leq L_E(X) \leq |E| \cdot \frac{|X \setminus X_E|}{|X|} . \quad (3.1)$$

3.2 Graphs of actions

As mentioned, it will be useful to represent a group action as a labeled graph. Throughout Section 3.2, we fix a finite set S and a group Λ generated by S . We have a natural surjection $\mathbb{F}_S \rightarrow \Lambda$ which enables us to regard a given action of Λ as an action of \mathbb{F}_S . The *action graph* of a finite Λ -set X (with regard to the set of generators S) is the edge-labeled directed graph over the vertex set X , which has a directed edge labeled s from x to $s \cdot x$ for each $x \in X$ and $s \in S$. Note that the action graph of X remains the same if we choose to treat X as an \mathbb{F}_S -set rather a Λ -set.

In the context of graphs of actions, it is often useful to consider pointed sets (X, x_0) , i.e., a set X together with a distinguished point $x_0 \in X$. The role of X will always be taken by a Λ -set or a subset of a Λ -set. We use the notation $f : (X, x_0) \rightarrow (Y, y_0)$ for a map $f : X \rightarrow Y$ which sends $x_0 \mapsto y_0$.

It is helpful to consider the following definitions with the role of Λ taken by the free group \mathbb{F}_S itself, or with $\Lambda = \mathbb{Z}^{|S|}$. We proceed to define isomorphisms of subgraphs of action graphs, and several related notions.

Definition 3.7. Let X and Y be Λ -sets and let $f : X_0 \rightarrow Y_0$ be a map between subsets $X_0 \subseteq X$ and $Y_0 \subseteq Y$.

- (i) For $s \in S$ and $x \in X_0$, we say that f *preserves* the edge $x \xrightarrow{s}$ if either $s \cdot x \notin X_0$ and $s \cdot f(x) \notin Y_0$, or $s \cdot x \in X_0$, $s \cdot f(x) \in Y_0$ and $f(s \cdot x) = s \cdot f(x)$.
- (ii) If f is bijective, and preserves $x \xrightarrow{s}$ for every $s \in S$ and $x \in X_0$, we say that f is a *subgraph isomorphism* from X_0 to Y_0 .

For a Λ -set X , a point $x \in X$ and a subset $P \subseteq \Lambda$, we write $P \cdot x = \{p \cdot x \mid p \in P\}$. For subsets P_1 and P_2 of Λ , we write $P_1 \cdot P_2 = \{p_1 \cdot p_2 \mid p_1 \in P_1, p_2 \in P_2\}$.

Definition 3.8. Let (X, x) and (Y, y) be pointed Λ -sets and $P \subseteq \Lambda$, a subset. Assume that

$$\text{Stab}_\Lambda(x) \cap (P^{-1}P) \subseteq \text{Stab}_\Lambda(y) \cap (P^{-1}P) . \quad (3.2)$$

Define the function $F_{P,x,y} : P \cdot x \rightarrow P \cdot y$ by $F_{P,x,y}(p \cdot x) = p \cdot y$ for each $p \in P$. Note that this function is well-defined since if $p_1, p_2 \in P$ and $p_1 \cdot x = p_2 \cdot x$, then $p_2^{-1}p_1 \in \text{Stab}_\Lambda(x)$, and so $p_2^{-1}p_1 \in \text{Stab}_\Lambda(y)$, hence $p_1 \cdot y = p_2 \cdot y$.

In the notation of the above definition, the function $F_{P,x,y}$ is injective if and only if the inclusion in (3.2) is in fact an equality. The special case of Definition 3.8 where $X = \Lambda$ gives rise to the following definition:

Definition 3.9. Let Y be a Λ -set, $y \in Y$ and $P \subseteq \Lambda$, a subset. We say that P *injects into* Y at y if the map $F_{P,1_\Lambda,y} : P \rightarrow P \cdot y$ is injective, or, equivalently, if $\text{Stab}_\Lambda(y) \cap (P^{-1}P) = \{1_\Lambda\}$. We say that P *bijjects onto* Y at y if this map is bijective.

Note that Definition 3.9 merely requires the map $F_{P,1_\Lambda,y}$ to be injective, but not necessarily a subgraph isomorphism. We seek sufficient conditions which guarantee that a given map of the form $F_{P,x,y}$, for general \mathbb{F}_S -sets X, Y and points $x \in X, y \in Y$, preserves a given edge, or even that it is a subgraph isomorphism. The following two lemmas provide such conditions by considering short elements of Λ and whether or not they belong to $\text{Stab}_\Lambda(x)$ and $\text{Stab}_\Lambda(y)$.

Lemma 3.10. *Let (X, x) and (Y, y) be pointed Λ -sets and $P \subseteq \Lambda$, a subset, such that Condition (3.2) of Definition 3.8 is satisfied. Let $s \in S$, $p \in P$ and assume that*

$$\text{Stab}_\Lambda(x) \cap (P^{-1} \cdot sp) = \text{Stab}_\Lambda(y) \cap (P^{-1} \cdot sp) .$$

Then, the map $F = F_{P,x,y} : P \cdot x \rightarrow P \cdot y$ preserves $p \cdot x \xrightarrow{s}$.

Proof. We first note that for $p_1 \in P$,

$$sp \cdot x = p_1 \cdot x \quad \text{if and only if} \quad sp \cdot y = p_1 \cdot y . \quad (3.3)$$

Indeed, $sp \cdot x = p_1 \cdot x$ if and only if $p_1^{-1}sp \in \text{Stab}_\Lambda(x)$ if and only if $p_1^{-1}sp \in \text{Stab}_\Lambda(y)$ if and only if $sp \cdot y = p_1 \cdot y$.

It follows from (3.3) that $s \cdot (p \cdot x)$ belongs to $P \cdot x$ if and only if $s \cdot F(p \cdot x) = s \cdot (p \cdot y)$ belongs to $P \cdot y$. Assume that these equivalent conditions hold. It remains to show that in this case, $F(s \cdot (p \cdot x)) = s \cdot F(p \cdot x)$. Write $sp \cdot x = p_1 \cdot x$ for $p_1 \in P$. By (3.3), $sp \cdot y = p_1 \cdot y$. So, $F(s \cdot (p \cdot x)) = F(p_1 \cdot x) = p_1 \cdot y = sp \cdot y = s \cdot F(p \cdot x)$, as required. \square

Lemma 3.11. *Let (X, x) and (Y, y) be pointed Λ -sets and $P \subseteq \Lambda$, a subset. Write $S_1 = S \cup \{1_\Lambda\}$ and assume that*

$$\text{Stab}_\Lambda(x) \cap (P^{-1} \cdot S_1 \cdot P) = \text{Stab}_\Lambda(y) \cap (P^{-1} \cdot S_1 \cdot P) .$$

Then, $F_{P,x,y} : P \cdot x \rightarrow P \cdot y$ is well-defined and is a subgraph isomorphism.

Proof. Since $\text{Stab}_\Lambda(x) \cap (P^{-1} \cdot P) = \text{Stab}_\Lambda(y) \cap (P^{-1} \cdot P)$, the map $F_{P,x,y}$ is well-defined and injective. Furthermore, for each $s \in S$ and $x \in P \cdot x$, since $\text{Stab}_\Lambda(x) \cap (P^{-1} \cdot sp) = \text{Stab}_\Lambda(y) \cap (P^{-1} \cdot sp)$, Lemma 3.10 implies that $F_{P,x,y}$ preserves $x \xrightarrow{s}$. \square

We proceed to define balls in \mathbb{F}_S and \mathbb{F}_S -sets, and give a useful corollary of Lemma 3.11. The word norm $|\cdot|$ on \mathbb{F}_S is defined for a word $w \in \mathbb{F}_S$ as the length of w when written as a reduced word over S^\pm . Let X be an \mathbb{F}_S -set. The word-norm induces a metric d_X on X :

$$\forall x_1, x_2 \in X \quad d_X(x_1, x_2) = \min \{ |w| \mid w \in \mathbb{F}_S, w \cdot x_1 = x_2 \} .$$

Write $B_X(x, r)$ for the ball of radius $r \geq 0$ centered at the point $x \in X$ with respect to d_X . For $A \subseteq X$, let $B(A, r) = \bigcup_{x \in A} B(x, r)$. In the special case $X = \Lambda$, we also write $B_\Lambda(r)$ for $B_\Lambda(1_\Lambda, r)$. This notation will be used often with either $\Lambda = \mathbb{F}_S$ or $\Lambda = \mathbb{Z}^{|S|}$. For $r \geq 0$, plugging in $P = B_\Lambda(r)$ in Lemma 3.11, we deduce the following corollary:

Lemma 3.12. *Let (X, x) and (Y, y) be Λ -sets and $r \geq 0$, an integer. Assume that*

$$\text{Stab}_\Lambda(x) \cap B_\Lambda(2r + 1) = \text{Stab}_\Lambda(y) \cap B_\Lambda(2r + 1) .$$

Then, the map $F_{B_\Lambda(r),x,y} : B_X(x, r) \rightarrow B_Y(y, r)$ is well-defined and is a subgraph isomorphism.

Now, assume that $E \subseteq \mathbb{F}_S$ is a finite set generating the kernel of the surjection $\mathbb{F}_S \rightarrow \Lambda$ as a normal subgroup (hence $\mathbb{F}_S / \langle\langle E \rangle\rangle \cong \Lambda$). We end this section with two definitions and a basic lemma that allows us to bound the global defect $G_E(X)$ of an \mathbb{F}_S -set X with respect to E (see Definition 3.3) in terms of properties of a map between graphs of actions.

Definition 3.13. Let X be a Λ -set and $X_0 \subseteq X$ a subset. A point $x \in X_0$ is *internal* in X_0 if $S \cdot x \subseteq X_0$.

Definition 3.14. Let X and Y be Λ -sets and $X_0 \subseteq X$, a subset. Take a function $f : X_0 \rightarrow Y$. Define the set $\text{Eq}(f) \subseteq X$ of *equivariance points* of f as

$$\text{Eq}(f) = \{ x \in X_0 \mid \forall s \in S \quad s \cdot x \in X_0 \text{ and } f(s \cdot x) = s \cdot f(x) \} .$$

That is, $\text{Eq}(f)$ is the set of internal points $x \in X_0$, for which f preserves $x \xrightarrow{s}$ for all $s \in S$.

Lemma 3.15. *Let Y be a Λ -set, X an \mathbb{F}_S -set and $f : Y \rightarrow X$ an injective map. Then, $G_E(X) \leq |S| \cdot \left(1 - \frac{1}{|X|} \cdot |\text{Eq}(f)| \right)$.*

Proof. Let Z be a trivial Λ -set of cardinality $|Y| - |X|$, i.e., each $g \in \Lambda$ fixes each $z \in Z$. Fix a bijection $f_Z : Z \rightarrow X \setminus \text{Im}(f)$. We consider the disjoint union $Y \amalg Z$ and the bijection $f \amalg f_Z : Y \amalg Z \rightarrow X$. Then,

$$\begin{aligned} G_E(X) &\leq d_S \left(Y \amalg Z, X \right) \\ &\leq \|f \amalg f_Z\|_S \\ &\leq \frac{1}{|X|} \cdot \left(|X| \cdot |S| - |S| \cdot |\text{Eq}(f \amalg f_Z)| \right) \\ &\leq \frac{1}{|X|} \cdot \left(|X| \cdot |S| - |S| \cdot |\text{Eq}(f)| - |S| \cdot |\text{Eq}(f_Z)| \right) \\ &\leq |S| \cdot \left(1 - \frac{1}{|X|} \cdot |\text{Eq}(f)| \right) . \end{aligned}$$

□

4 Abelian groups are polynomially stable

The aim of this section is to prove Theorem 1.16, our main theorem. We begin by defining several objects that shall remain fixed throughout Section 4.

Let Γ be a finitely-generated abelian group. Without loss of generality we can realize Γ as follows: Let $m \geq d \geq 0$, take a basis $\{e_1, \dots, e_m\}$ for \mathbb{Z}^m and let $2 \leq \beta_{m-d+1} \leq \dots \leq \beta_m$ be integers. Define

$$K = \langle \{\beta_i \cdot e_i\}_{i=m-d+1}^m \rangle \leq \mathbb{Z}^m \quad (4.1)$$

and write $\Gamma = \mathbb{Z}^m / K$. Let $\text{Tor}(\Gamma)$ denote the torsion subgroup of Γ . Let

$$\beta_E = \begin{cases} \beta_m & \text{if } m > d \\ 1 & \text{if } m = d . \end{cases}$$

Theorem 1.16 asserts that $\deg(\text{SR}_\Gamma)$ is finite. We will, in fact, provide an explicit upper bound on $\deg(\text{SR}_\Gamma)$. In the case $d = 0$, i.e., if Γ is finite, Proposition A.4 says that $\deg(\text{SR}_\Gamma) = 1$. We proceed assuming that $d \geq 1$. We shall show that

$$\deg(\text{SR}_\Gamma) \leq C_{\text{bound}}(\Gamma) \quad (4.2)$$

for

$$C_{\text{bound}}(\Gamma) = O\left(2^d \cdot d \cdot \max\{d \log d, \log \beta_E, 1\}\right) , \quad (4.3)$$

where the implied constant of the $O(\cdot)$ notation is an absolute constant. We note that, in order to minimize $C_{\text{bound}}(\Gamma)$, one may let $\beta_{m-d+1}, \dots, \beta_m$ correspond to the primary decomposition of Γ , thus making the constant β_E the largest prime power in that decomposition.

Let \mathbb{F}_m be the free group on $S = \{\hat{e}_1, \dots, \hat{e}_m\}$, and consider the surjection $\pi : \mathbb{F}_m \rightarrow \mathbb{Z}^m$, sending $\hat{e}_i \mapsto e_i$. We get a sequence of surjections

$$\mathbb{F}_m \xrightarrow{\pi} \mathbb{Z}^m \longrightarrow \Gamma .$$

We also fix a free group \mathbb{F}_d , generated by $\{\hat{e}_1, \dots, \hat{e}_d\}$. That is, we write \mathbb{F}_d for this fixed copy of a free group of rank d inside our fixed free group \mathbb{F}_m .

By the definition, the stability rate SR_Γ of Γ can be computed through any presentation of Γ (see Proposition 1.11). We proceed to choose the equation-set E , defining Γ , with which we will work. Let

$$E_0 = \{[\hat{e}_i^{\epsilon_1}, \hat{e}_j^{\epsilon_2}] \in \mathbb{F}_m \mid i, j \in [m], \epsilon_1, \epsilon_2 \in \{+1, -1\}\} \subseteq \mathbb{F}_m,$$

where $[x, y]$ denotes the commutator $xyx^{-1}y^{-1}$, and define

$$E = E_0 \cup \{\hat{e}_i^{\beta_i} \mid m-d+1 \leq i \leq m\} \subseteq \mathbb{F}_m.$$

Note that if $d = m$, then E is equivalent to the equation-set E_{comm}^d from the introduction. Since $\Gamma \cong \mathbb{F}_m / \langle\langle E \rangle\rangle$, we have $\text{SR}_\Gamma = [\text{SR}_E]$, and so our goal is to prove that

$$\text{deg}(\text{SR}_E) \leq C_{\text{bound}}(\Gamma).$$

For future reference, we fix the following constants:

$$\begin{aligned} C_d &= \max\{3 \cdot 7^d \cdot d^{2d+2}, \beta_E\} \\ t_E &= \max\{d^{-1} \cdot (m-d) \cdot m \cdot \beta_E, 2\}. \end{aligned}$$

Additionally, we let

$$\hat{T} = \left\{ \prod_{i=m-d+1}^m \hat{e}_i^{\alpha_i} \in \mathbb{F}_m \mid \forall i \ 0 \leq \alpha_i < \beta_i \right\},$$

and

$$T = \left\{ \sum_{i=m-d+1}^m \alpha_i \cdot e_i \in \mathbb{Z}^m \mid \forall i \ 0 \leq \alpha_i < \beta_i \right\}.$$

In the rest of this section, the implied constants in the $O(\cdot)$ notation are allowed to depend on m , d and E .

4.1 Proof plan

In this section we outline our proof of Theorem 1.16, as implemented in Sections 4.2–4.7.

Let X be an \mathbb{F}_m -set, and write $n = |X|$. By Proposition 3.5, in order to bound SR_E , it suffices to bound $G_E(X)$ in terms of $L_E(X)$. To this end, we algorithmically construct a Γ -set Y (Proposition 4.30), together with a certain injection $f : Y \rightarrow X$ with many equivariance points, namely, $|\text{Eq}(f)| \geq n \cdot \left(1 - O\left(L_E(X)^{\frac{1}{C_{\text{bound}}(\Gamma)}}\right)\right)$. Lemma 3.15 then gives the bound $G_E(X) \leq |S| \cdot \left(1 - \frac{|\text{Eq}(f)|}{n}\right) \leq O\left(L_E(X)^{\frac{1}{C_{\text{bound}}(\Gamma)}}\right)$, which yields the claim of Theorem 1.16.

We build Y as the disjoint union of a collection of small Γ -sets $\{Y_x\}_{x \in J}$, each equipped with an injection $f_x : Y_x \rightarrow X$. The images of these injections are

pairwise disjoint and $f : Y \rightarrow X$ is taken to be the disjoint union of the maps f_x . Clearly, due to this construction, $|\text{Eq}(f)|$ is at least $\sum_{x \in J} |\text{Eq}(f_x)|$, so we wish to maximize the latter sum. Towards this end, it is desirable that the images of the injections f_x cover almost all of X , and that each of the injections has a large fraction of equivariance points. We manage to construct the injections f_x so that the equivariance points of f_x are approximately those points in Y_x that are mapped to internal points of $\text{Im}(f_x) \subseteq X$. Hence, we think of the ratio $\frac{|Y_x \setminus \text{Eq}(f_x)|}{|Y_x|}$ as an *isoperimetric ratio*, which we wish to minimize.

The reader may prefer to read both the proof and its outline under the simplifying assumption that Γ is torsion-free, i.e., $m = d$ and, accordingly, $\Gamma = \mathbb{Z}^m$. In fact, the torsion-free case of Theorem 1.16 implies the general case. This follows from Proposition A.3, since every finitely-generated abelian group is a quotient of a torsion-free finitely-generated abelian group by a finitely-generated subgroup. The simplified strategy, of starting with the torsion-free case and then using Proposition A.3, comes at the price of a somewhat worse bound on the degree, compared to (4.2).

We turn to give an outline of our algorithm.

4.1.1 The algorithm constructing Y and f

Our algorithm works iteratively as follows. We first initialize a rather large number t_1 which depends on the ratio $\frac{|X_E|}{|X|}$, which in turn is related to $L_E(X)$ (see Equation (3.1)). In the first iteration, we find, in a greedy manner, a collection of Γ -sets $\{Y_x\}_{x \in J_1}$ and respective injections $\{f_x : Y_x \rightarrow X\}_{x \in J_1}$ with pairwise disjoint images, such that $\frac{|Y_x \setminus \text{Eq}(f_x)|}{|Y_x|} \leq O(\frac{1}{t_1})$ for each $x \in J_1$. We think of the images of the injections $\{f_x\}_{x \in J_1}$ as “tiles” embedded in X , and of t_1 as a parameter used in the construction of these tiles. The set J_1 is maximal in the sense that we cannot add more tiles with parameter t_1 without violating the constraint that they be disjoint. We proceed to tile the remainder of X . We define a new parameter $t_2 < t_1$ which is equal to t_1 divided by some constant, and repeat this process for another iteration, which yields additional Γ -sets $\{Y_x\}_{x \in J_2}$ and corresponding injections, perhaps with a worse isoperimetric ratio. We require that the images of these injections be disjoint from each other, as well as from the images obtained in the previous iteration. We proceed in this manner, tiling a constant fraction of the remainder of X in each iteration, until t_i is below a certain threshold, at which point the iterative algorithm halts. Finally, we set Y to be the disjoint union of the Γ -sets $\{Y_x\}_{x \in J_1 \cup \dots \cup J_s}$ constructed throughout the s iterations, and define $f : Y \rightarrow X$ as the disjoint union of the maps $\{f_x\}_{x \in J_1 \cup \dots \cup J_s}$.

Note that, as the algorithm progresses, our injections become less and less efficient, that is, their images have a larger isoperimetric ratio. After developing the necessary machinery in Sections 4.4–4.6, we conclude the proof of Theorem 1.16 in Section 4.7 by defining the above algorithm, and showing that it produces an injection f with many equivariance points. We turn to discuss the technique by which we build the sets Y_x and the corresponding injections f_x .

4.1.2 Mapping a single Γ -set Y_x into X

As mentioned, the i -th iteration of our algorithm injects a collection of Γ -sets $\{Y_x\}_{x \in J_i}$ into X . The isoperimetric ratio of each of these injections must be bounded by $O(\frac{1}{t_i})$, where t_i is the parameter introduced in Section 4.1.1. We now focus on the main technical challenge, namely, building a single finite Γ -set Y_x and a corresponding injection $f_x : Y_x \rightarrow X$. The efficiency of our construction is reflected in the fact that an “accumulation” of no more than $O(\frac{d}{t_i})$ nearby points of X which belong to X_E suffices to construct an injection with isoperimetric constant bounded by $O(\frac{1}{t_i})$. We begin by describing two essential tools for the construction of Y_x and f_x :

Tool A (Proposition 4.21): This tool requires a point $x \in X_E$ with a large enough neighborhood (called a “box-neighborhood of side-length t_i ”) which is entirely contained in X_E . It provides a radius $r_A \geq \Omega(t_i)$ such that the ball $B_X(x, r_A)$ is contained in the aforementioned neighborhood. It also provides a new (usually infinite) pointed Γ -set (U_A, u_A) and a subgraph isomorphism f_A :

$$\begin{array}{ccc} B_{U_A}(u_A, r_A) & \xrightarrow{f_A} & B_X(x, r_A) \\ \sqcap & & \sqcap \\ (U_A, u_A) & & X \end{array}$$

Tool C (Proposition 4.11): Given a pointed Γ -set (V, v) and $t_C \in \mathbb{N}$, this tool creates an injective map:

$$(Y, y) \xrightarrow{f_C} (V, v)$$

where (Y, y) is a new small *finite* pointed Γ -set and the isoperimetric ratio of $\text{Im}(f_C)$ is bounded by $O(\frac{1}{t_C})$.

One may be tempted to try to construct Y_x and the injective map $f_x : Y_x \rightarrow X$ as follows: Locate a point $x \in X_E$ with the property required by Tool A, and use this tool to create:

$$B_{U_A}(u_A, r_A) \xrightarrow{f_A} X$$

as in the description of Tool A. Then, apply Tool C to $(V, v) = (U_A, u_A)$ with some $t_C \geq \Omega(t_i)$, and create:

$$(Y_x, y_x) \xrightarrow{f_C} (U_A, u_A)$$

as in the description of Tool C. Now, in the very fortunate case where the image of f_C is contained in the domain $B_{U_A}(u_A, r_A)$ of f_A , we can define the map $f_x : Y_x \rightarrow X$ as the following composition:

$$Y_x \xrightarrow{f_C} B_{U_A}(u_A, r_A) \xrightarrow{f_A} X .$$

When this works, the map f_x is injective and its image has a small isoperimetric ratio, as required, because these properties hold for f_C and since $f_A : B_{U_A}(u_A, r_A) \rightarrow B_X(x, r_A)$ is a subgraph isomorphism. However, the image of the map f_C , produced by Tool C, is usually too large to be contained in the domain of f_A . We solve this issue by introducing yet another Γ -set U_B which sits between Y_x and $B_{U_A}(u_A, r_A)$ in the above diagram. As is the case for U_A , the set U_B is usually infinite. It is generated by Tool B (see below), and has a useful combination of properties: (I) locally, U_B looks like U_A in a rather large radius, and (II) when Tool C is applied to U_B , the image in U_B of the resulting injection f_C is relatively small.

Tool B (Proposition 4.13): Given a pointed Γ -set (U_A, u_A) and $t_B \in \mathbb{N}$, this tool creates a subgraph isomorphism:

$$\begin{array}{ccc} (U_B^0, u_B) & \xrightarrow{f_B} & (U_A^0, u_A) \\ \text{!}\cap & & \text{!}\cap \\ (U_B, u_B) & & (U_A, u_A) \end{array}$$

where (U_B, u_B) is a new pointed Γ -set and U_A^0 and U_B^0 are finite sets. The tool guarantees the following properties:

- (I) $U_A^0 \subseteq B_{U_A}(u_A, O(t_B))$, and
- (II) The set U_B^0 is exactly the image of the map f_C that our implementation of Tool C provides when it is applied to $(V, v) = (U_B, u_B)$ with $t_C = t_B$.

Using all three tools, we define f_x as the composition of the following chain of maps:

$$\begin{array}{ccccc} Y_x & \xrightarrow{f_C} & \text{Im}(f_C) & \xrightarrow{f_B} & B_{U_A}(u_A, r_A) & \xrightarrow{f_A} & X \\ & & \text{!}\cap & & \text{!}\cap & & \\ & & U_B & & U_A & & \end{array}$$

The objects and maps in the diagram above are created by first using Tool A to create U_A and f_A , then applying Tool B to U_A to create U_B and f_B , and finally applying Tool C to U_B to create Y_x and f_C . All three maps f_A , f_B and f_C are injective, and so the same is true for f_x . Both f_A and f_B are subgraph isomorphisms onto their respective images, and the image of f_C has isoperimetric ratio at most $O\left(\frac{1}{t_i}\right)$. Hence, the same isoperimetric property is true for f_x , as required.

The proofs for Tools A and B use *basis reduction theory* of sublattices of \mathbb{Z}^m , and the proof for Tool C is also in a related spirit. More specifically, a transitive

Γ -set V is isomorphic to \mathbb{Z}^m/H for some subgroup $H \leq \mathbb{Z}^m$, and we are able to study V by applying reduction theory to H , thought of as a sublattice of \mathbb{Z}^m .

We note that the isoperimetric property possessed by each tile serves two purposes in our proof. The “local purpose” is to ensure that each injection f_x has a large fraction of equivariance points as described above. The “global purpose” is to ensure that we can pack *many* tiles into X .

4.2 Geometric definitions

Section 3.2 introduced the word-norm on a free group, the word-metric on sets which are acted on by a free group, and defined balls with respect to the word-metric. In addition to these, our proof will make use of various norms on \mathbb{Z}^m (specifically, L^1 , L^2 and L^∞), and of “boxes” in free groups. These are described below:

Geometry of \mathbb{Z}^m For $1 \leq p \leq \infty$, let $\|\cdot\|_p$ denote the L^p norm, restricted to \mathbb{Z}^m . Let $B_{\mathbb{Z}^m}^{L^p}(x, r) \subseteq \mathbb{Z}^m$ be the closed ball, with respect to $\|\cdot\|_p$, of radius r , centered at x . Again, we omit x for a ball centered at $0_{\mathbb{Z}^m}$. We note that in the case $p = 1$, $\|\cdot\|_1$ coincides with the word-metric on \mathbb{Z}^m as an \mathbb{F}_m -set, and so $B_{\mathbb{Z}^m}(x, r) = B_{\mathbb{Z}^m}^{L^1}(x, r)$.

Boxes in \mathbb{F}_m and \mathbb{F}_d Let \mathbb{F}_k be the free group on $\{\hat{e}_1, \dots, \hat{e}_k\}$ (we are interested in $k \in \{m, d\}$). We say that a word in \mathbb{F}_k is *sorted* if it is of the form $\prod_{i=1}^k \hat{e}_i^{a_i}$, $a_i \in \mathbb{Z}$. Write $\pi_k : \mathbb{F}_k \rightarrow \mathbb{Z}^k$ for the surjection sending $\hat{e}_i \mapsto e_i$. For each $v \in \mathbb{Z}^k$ we define a canonical representative $\hat{v} \in \mathbb{F}_k$ of the set $\pi_k^{-1}(v)$, namely, \hat{v} is the unique sorted word such that $\pi_k(\hat{v}) = v$. Reusing this notation, for $w \in \mathbb{F}_k$ let \hat{w} denote $\widehat{\pi(w)}$, i.e, the sorted form of w . The following definition introduces a key player in our proof:

Definition 4.1. For $t \in \mathbb{N}$, let

$$\text{Box}_{\mathbb{F}_k}(t) = \left\{ \hat{v} \mid v \in B_{\mathbb{Z}^k}^{L^\infty}(t) \right\} = \left\{ \hat{v} \mid v \in \mathbb{Z}^k, \|v\|_\infty \leq t \right\} \subseteq \mathbb{F}_k.$$

Note that $w \in \text{Box}_{\mathbb{F}_k}(|w|)$ whenever w is sorted. Finally, for a subset A of \mathbb{Z}^k or \mathbb{F}_k , define $\hat{A} = \{\hat{a} \mid a \in A\}$.

4.3 A review of reduction theory (of sublattices of \mathbb{Z}^l)

Let $l \geq 1$. We use the term *lattice* interchangeably with *a subgroup $H \leq \mathbb{Z}^l$* . A *basis* for H is a set of linearly independent vectors $\mathcal{D} \subseteq \mathbb{Z}^l$ such that H is the set of all integer combinations of elements in \mathcal{D} . It is well known that every lattice affords a basis, and that different bases representing the same lattice all have the same cardinality. Hence, we may define the *rank* of H by $\text{rank}H = |\mathcal{D}|$. The goal of *reduction theory* (See [15] for a survey) is to represent H via a *reduced basis*, namely, a basis consisting of relatively short vectors. In this section, we adapt a certain result from reduction theory to our purposes.

Definition 4.2. Let $H \leq \mathbb{Z}^l$ be a lattice and $k = \text{rank} H$. The *successive minima* sequence $\lambda_1(H), \dots, \lambda_k(H)$ is defined as follows: $\lambda_i(H)$ is the minimum radius r for which $\text{rank} \langle \{v \in H \mid \|v\|_2 \leq r\} \rangle \geq i$.

We note that for $l > 4$, the vectors that yield the successive minima are not necessarily a basis for H ([16] p. 51). However, as we now elaborate, H does accommodate a basis consisting of vectors within the same order of magnitude as the successive minima. It is known that every lattice has a basis of a certain type called *Korkin-Zolotarev reduced* [15]. The following proposition about such bases is a part of Theorem 2.1 in [14].

Proposition 4.3. *If $H \leq \mathbb{Z}^l$ is a lattice and $B = \{b_1, \dots, b_k\}$ is a Korkin-Zolotarev reduced basis for H , then $\|b_i\|_2 \leq \frac{1}{2} \cdot \sqrt{i+3} \cdot \lambda_i(H)$ for each $1 \leq i \leq k$.*

Our use of Proposition 4.3 will always be mediated through Proposition 4.5, below.

Definition 4.4. For a finite subset $A \subseteq \mathbb{Z}^l$, write $\|A\|_1 = \sum_{v \in A} \|v\|_1$.

Proposition 4.5. *Let $H \leq \mathbb{Z}^l$ be a lattice of rank k and $t \in \mathbb{N}$, such that $\text{rank}(H \cap B_{\mathbb{Z}^l}^{L^2}(t)) = k$. Then, H has a basis \mathcal{D} such that:*

(i) $\mathcal{D} \subseteq B_{\mathbb{Z}^l}(l \cdot t)$.

(ii) $\|\mathcal{D}\|_1 \leq l^2 \cdot t$.

Proof. Let $\mathcal{D} = \{b_1, \dots, b_k\}$ be a Korkin-Zolotarev reduced basis for H . By Proposition 4.3,

$$\|b_i\|_2 \leq \frac{1}{2} \sqrt{i+3} \cdot \lambda_i(H) \leq \frac{1}{2} \sqrt{l+3} \cdot t \leq \frac{1}{2} \sqrt{4l} \cdot t = \sqrt{l} \cdot t,$$

and so

$$\|b_i\|_1 \leq l \cdot t,$$

which yields the first claim. The second claim follows since

$$\|\mathcal{D}\|_1 \leq |\mathcal{D}| \cdot \max_{1 \leq i \leq k} \|b_i\|_1 \leq k \cdot l \cdot t \leq l^2 \cdot r.$$

□

4.4 Tool C: The standard-completion isoperimetric method

In this section we develop *Tool C*. Let (V, v_0) be a pointed Γ -set and $t \in \mathbb{N}$. Our goal is to build a small finite pointed Γ -set (Y, y_0) , and an injective map $f_C : (Y, y_0) \rightarrow (V, v_0)$ such that

$$|\text{Eq}(f_C)| \geq |Y| \cdot \left(1 - \mathcal{O}\left(\frac{1}{t}\right)\right). \tag{4.4}$$

We may assume that V is transitive, since otherwise we may consider only the component containing v_0 . Hence, V is realizable as \mathbb{Z}^m/H for some $K \leq H \leq \mathbb{Z}^m$ (see Equation (4.1)), with v_0 corresponding to the coset $0 + H$. Let \mathcal{D}_0 be a lattice basis for H . We first consider the simple case where \mathcal{D}_0 consists of axis-parallel vectors, namely, $\mathcal{D}_0 = \{\alpha_i \cdot e_i \mid i \in [m] \setminus I\}$ for some $I \subseteq [m]$, where $\{\alpha_i\}_{i \in [m] \setminus I}$ are positive integers. We complete \mathcal{D}_0 to a full-rank basis $\mathcal{T} = \mathcal{D}_0 \cup \{2t \cdot e_i\}_{i \in I}$. Define $(Y, y_0) = (\mathbb{Z}^m / \langle \mathcal{T} \rangle, 0 + \langle \mathcal{T} \rangle)$, and note that Y is a Γ -set since $K \subseteq H \subseteq \langle \mathcal{T} \rangle$.

Each point $y \in Y$ has a unique representation as

$$y = \sum_{f \in \mathcal{D}_0} a_f \cdot f + \sum_{i \in I} b_i \cdot e_i + \langle \mathcal{T} \rangle \quad 0 \leq a_f < 1, \quad -t \leq b_i < t.$$

Let $f_C : (Y, y_0) \rightarrow (V, v_0)$ map such a point y to the point

$$\sum_{f \in \mathcal{D}_0} a_f \cdot f + \sum_{i \in I} b_i \cdot e_i + H$$

of V . This map f_C is clearly injective. Now, consider a point $y \in Y$ as above. The map f_C necessarily preserves the edge $y \xrightarrow{e_i}$ for each $i \in [m] \setminus I$. If $i \in I$, then f_C preserves this edge **unless** $b_i = t - 1$. It follows that

$$\text{Eq}(f_C) = \left\{ \sum_{f \in \mathcal{D}_0} a_f \cdot f + \sum_{i \in I} b_i \cdot e_i + \langle \mathcal{T} \rangle \mid 0 \leq a_f < 1 \text{ and } -t \leq b_i < t - 1 \right\},$$

so Equation (4.4) is satisfied. As mentioned, it is desirable that Y be small, i.e., that \mathcal{T} be a short basis. While we cannot control \mathcal{D}_0 , we have chosen the additional vectors $2t \cdot e_i$ to be as short as possible, that is, just long enough to guarantee Equation (4.4).

We turn to the general case, where the basis \mathcal{D}_0 of H might not consist of axis-parallel vectors. Again, we augment \mathcal{D}_0 with axis-parallel vectors to form a full-rank basis, namely, $\mathcal{T} = \mathcal{D}_0 \cup \{2t \cdot e_i \mid i \in I\}$ for some carefully chosen $I \subseteq [m]$, and define (Y, y_0) and f_C as above. However, it is now possible that an edge labeled e_j ($j \in [m] \setminus I$) is not preserved by f_C . In order to analyze the behavior of the generator e_j , we apply a linear transformation that maps \mathcal{T} to an axis-parallel basis. The image of e_j under this transformation may be large in its I coordinates, which means that many e_j -labeled edges are not preserved by f_C . In order to control the effect of these generators e_j , we need to choose I so that the vectors $\{e_i\}_{i \in I}$ are *nearly orthogonal* to \mathcal{D}_0 . Consequently, \mathcal{T} is already close to being an orthogonal basis, thereby bounding the distortion of the linear transformation that “fixes” it to being orthogonal. We proceed to formally define these notions.

Given a finite ordered set of vectors $A \subseteq \mathbb{R}^m$, we write \mathcal{M}_A for the matrix whose columns are the elements of A in the standard basis. Note that if A is a basis for \mathbb{R}^m then \mathcal{M}_A^{-1} is the change of basis matrix transforming a standard-coefficients vector to an A -coefficients vector (by multiplying column vectors of

coefficients from the left). Given a set $I \subseteq [m]$, let \mathcal{M}_A^I denote the matrix \mathcal{M}_A with the i -th row stricken out for every $i \in I$. Notice that if $|I| = m - |\mathcal{D}_0|$, then

$$|\det \mathcal{M}_{\mathcal{D}_0}^I| = |\det \mathcal{M}_{\mathcal{D}_0 \cup \mathcal{D}_1}|, \quad (4.5)$$

where $\mathcal{D}_1 = \{e_i \mid i \in I\}$.

Definition 4.6. Let $\mathcal{D}_0 \subseteq \mathbb{Z}^m$ be an ordered set of linearly independent vectors.

- (i) Write $k = |\mathcal{D}_0|$. For a set of coordinates $I \subseteq [m]$ with $|I| = m - k$, the *strength* of I with respect to \mathcal{D}_0 is $|\det(\mathcal{M}_{\mathcal{D}_0}^I)|$, that is, the unsigned volume of the parallelotope generated by $\mathcal{D}_0 \cup \{e_i\}_{i \in I}$ (see Equation (4.5)).
- (ii) We say that I , as above, is a set of *strongest coordinates* for \mathcal{D}_0 if $|\det(\mathcal{M}_{\mathcal{D}_0}^I)| \geq |\det(\mathcal{M}_{\mathcal{D}_0}^J)|$ for every $J \subseteq [m]$ satisfying $|J| = m - k$. Note that in this case, $|\det(\mathcal{M}_{\mathcal{D}_0}^I)|$ is strictly positive.
- (iii) For a set $\mathcal{D}_1 \subseteq \mathbb{Z}^m$, we say the \mathcal{D}_1 is a *standard complement* for \mathcal{D}_0 if $\mathcal{D}_0 \cup \mathcal{D}_1$ is a basis for \mathbb{R}^m and $\mathcal{D}_1 = \{e_i\}_{i \in I}$ for some $I \subseteq [m]$. A standard complement \mathcal{D}_1 for \mathcal{D}_0 as above is *strong* if I is a set of strongest coordinates with respect to \mathcal{D}_0 .

Note that for every set of strongest coordinates $I \subseteq [m]$ for \mathcal{D}_0 , the set $\mathcal{D}_1 = \{e_i\}_{i \in I}$ is a strong standard complement for \mathcal{D}_0 . The following lemma formulates the near orthogonality condition mentioned above.

Lemma 4.7. *Let $\mathcal{D}_0 \subseteq \mathbb{Z}^m$ be a linearly independent set. Take a strong standard complement \mathcal{D}_1 for \mathcal{D}_0 . Write $\mathcal{D}_0 = \{h_1, \dots, h_k\}$ and $\mathcal{D}_1 = \{e_i\}_{i \in I}$, $I = \{i_1, \dots, i_{m-k}\}$, and let $\mathcal{D} = (h_1, \dots, h_k, e_{i_1}, \dots, e_{i_{m-k}})$. Then, all entries of the lower $(m-k) \times m$ block of the matrix $\mathcal{M}_{\mathcal{D}}^{-1}$ are in the range $[-1, 1]$*

Proof. Permuting the rows of $\mathcal{M}_{\mathcal{D}}$ maintains the property that the right $m-k$ columns of $\mathcal{M}_{\mathcal{D}}$ are a strong standard complement for the left k columns, and affects $\mathcal{M}_{\mathcal{D}}^{-1}$ by a permutation of its columns (in particular, it permutes the lower $(m-k) \times m$ block of $\mathcal{M}_{\mathcal{D}}^{-1}$). Therefore, we assume that $I = \{k+1, \dots, m\}$ without loss of generality. Accordingly, $\mathcal{D} = (h_1, \dots, h_k, e_{k+1}, \dots, e_m)$. Let $k+1 \leq i \leq m$ and $1 \leq j \leq m$. Our task is to show that $|(\mathcal{M}_{\mathcal{D}}^{-1})_{i,j}| \leq 1$. Note that the lower-right block of $\mathcal{M}_{\mathcal{D}}^{-1}$ is an $(m-k) \times (m-k)$ identity matrix. Thus, the claim holds whenever $j > k$. Henceforth, assume that $1 \leq j \leq k$. Let $\mathcal{M}_{\mathcal{D}}^{j,i}$ denote the matrix $\mathcal{M}_{\mathcal{D}}$ with the j -th row and i -th column removed. By Cramer's rule, the (i,j) -entry of $\mathcal{M}_{\mathcal{D}}^{-1}$ is either

$$\frac{\det \mathcal{M}_{\mathcal{D}}^{j,i}}{\det \mathcal{M}_{\mathcal{D}}}$$

or its negation. Hence, it suffices to show that $|\det \mathcal{M}_{\mathcal{D}}^{j,i}| \leq |\det \mathcal{M}_{\mathcal{D}}|$. By Equation (4.5), $|\det \mathcal{M}_{\mathcal{D}}| = |\det \mathcal{M}_{\mathcal{D}_0}^I|$. Similarly, $|\det \mathcal{M}_{\mathcal{D}}^{j,i}| = |\det \mathcal{M}_{\mathcal{D}_0}^{(I \setminus \{i\}) \cup \{j\}}|$. The claim follows since I is a set of strongest coordinates for \mathcal{D}_0 , and so $|\det \mathcal{M}_{\mathcal{D}_0}^I| \geq |\det \mathcal{M}_{\mathcal{D}_0}^{(I \setminus \{i\}) \cup \{j\}}|$. \square

We need to fix, once and for all, a strong standard complement $\mathcal{C}(\mathcal{D}_0)$ for each linearly independent $\mathcal{D}_0 \subseteq \mathbb{Z}^m$. For our purposes, it does not matter how this is done. However, for concreteness, we do it as follows:

Definition 4.8. For an ordered linearly independent subset $\mathcal{D}_0 \subseteq \mathbb{Z}^m$, let $\mathcal{C}(\mathcal{D}_0) \subseteq \mathbb{Z}^m$ be the lexicographically least among all strong standard complements for \mathcal{D}_0 . Let $I(\mathcal{D}_0)$ denote the subset of $[m]$ for which $\mathcal{C}(\mathcal{D}_0) = \{e_i\}_{i \in I(\mathcal{D}_0)}$.

Definition 4.9. Let $\mathcal{D}_0 \subseteq \mathbb{Z}^m$ be a linearly independent set. Write $k = |\mathcal{D}_0|$, and order $I(\mathcal{D}_0)$ by writing $I(\mathcal{D}_0) = \{i_{k+1}, \dots, i_m\}$, $i_{k+1} < \dots < i_m$. Define an ordered basis $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{C}(\mathcal{D}_0) = (h_1, \dots, h_k, e_{i_{k+1}}, \dots, e_{i_m})$ for \mathbb{R}^m . For $t \in \mathbb{N}$, define a *discrete parallelotope*

$$P_t^{\mathcal{D}_0} = \left\{ x \in \mathbb{Z}^m \mid \mathcal{M}_{\mathcal{D}}^{-1}(x) \in [0, 1)^k \times [-t, t)^{m-k} \right\},$$

and write $\hat{P}_t^{\mathcal{D}_0}$ for its canonical lift to \mathbb{F}_m (see Section 4.2), to wit,

$$\hat{P}_t^{\mathcal{D}_0} = \left\{ \hat{x} \mid x \in P_t^{\mathcal{D}_0} \right\}.$$

Lemma 4.10. *Let $\mathcal{D}_0 \subseteq \mathbb{Z}^m$ be a linearly independent set and $t \in \mathbb{N}$. Then,*

(i) $P_t^{\mathcal{D}_0} \subseteq B_{\mathbb{Z}^m}(\|\mathcal{D}_0\|_1 + (m - |\mathcal{D}_0|) \cdot t)$.

(ii) $|P_t^{\mathcal{D}_0}| = |P_1^{\mathcal{D}_0}| \cdot t^{m-|\mathcal{D}_0|}$.

(iii) *Let $U = \mathbb{Z}^m / \langle \mathcal{D}_0 \rangle$ and $u_0 \in U$. Then, $(-P_t^{\mathcal{D}_0} + P_t^{\mathcal{D}_0}) \cdot u_0 \subseteq P_{2t}^{\mathcal{D}_0} \cdot u_0$.*

Proof. The first statement follows from the triangle inequality of the L^1 -norm in \mathbb{Z}^m . The second statement holds since $P_t^{\mathcal{D}_0}$ is the disjoint union of $t^{m-|\mathcal{D}_0|}$ translates of $P_1^{\mathcal{D}_0}$. We turn to prove the third statement.

Define $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{C}(\mathcal{D}_0)$ as in Definition 4.9. We have

$$\begin{aligned} (-P_t^{\mathcal{D}_0} + P_t^{\mathcal{D}_0}) \cdot u_0 &= \left\{ x \in \mathbb{Z}^m \mid \mathcal{M}_{\mathcal{D}}^{-1}(x) \in (-1, 1)^k \times (-2t, 2t)^{m-k} \right\} \cdot u_0 \\ &= \left\{ x \in \mathbb{Z}^m \mid \mathcal{M}_{\mathcal{D}}^{-1}(x) \in [0, 1)^k \times (-2t, 2t)^{m-k} \right\} \cdot u_0 \\ &\subseteq P_{2t}^{\mathcal{D}} \cdot u_0, \end{aligned}$$

where the second equality above follows from the fact that $\mathcal{D}_0 \subseteq \text{Stab}_{\mathbb{Z}^m}(u_0)$. \square

The next proposition yields Tool C, namely, we show that for a pointed Γ -set (V, v_0) , there is an injection $f_C : (Y, y_0) \rightarrow (V, v_0)$, where Y is a finite Γ -set and $\text{Eq}(f_C)$ is large. We also provide some control over word-metric neighborhoods of $\text{Im}(f_C)$ in V .

Proposition 4.11. *Let $H \leq \mathbb{Z}^m$, $V = \mathbb{Z}^m / H$, $v_0 = 0 + H$ and $t \in \mathbb{N}$. Take a basis \mathcal{D}_0 for H . Then, there is a finite quotient Y of V , and an injective map $f_C : Y \rightarrow P_t^{\mathcal{D}_0} \cdot v_0$, such that:*

$$(i) |\text{Eq}(f_C)| \geq \left(1 - \frac{m - |\mathcal{D}_0|}{t}\right) \cdot |Y|.$$

$$(ii) \text{ For every integer } \tilde{t} \geq 0, |B_V(\text{Im}(f_C), \tilde{t})| \leq \left(1 + \frac{\tilde{t}}{t}\right)^{m - |\mathcal{D}_0|} \cdot |\text{Im}(f_C)|.$$

Proof. Let $k = |\mathcal{D}_0|$ and write $P_r = P_r^{\mathcal{D}_0}$ for every $r \in \mathbb{N}$. Let $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{C}(\mathcal{D}_0)$ and $\mathcal{T} = \mathcal{D}_0 \cup 2t \cdot \mathcal{C}(\mathcal{D}_0)$. Note that

$$H = \langle \mathcal{D}_0 \rangle = \left\{ x \in \mathbb{Z}^m \mid \mathcal{M}_{\mathcal{D}}^{-1}(x) \in \mathbb{Z}^k \times \{0\}^{m-k} \right\} \quad (4.6)$$

and

$$\langle \mathcal{T} \rangle = \left\{ x \in \mathbb{Z}^m \mid \mathcal{M}_{\mathcal{D}}^{-1}(x) \in \mathbb{Z}^k \times (2t\mathbb{Z})^{m-k} \right\}. \quad (4.7)$$

Define $Y = \mathbb{Z}^m / \langle \mathcal{T} \rangle$ and $y_0 = 0 + \langle \mathcal{T} \rangle$. Let $f_C = F_{P_t, y_0, v_0} : P_t \cdot y_0 \rightarrow P_t \cdot v_0$. First, we show that f_C is well-defined and injective. By the remark after Definition 3.8, it is enough to show that $\langle \mathcal{T} \rangle \cap (-P_t + P_t) = \langle \mathcal{D}_0 \rangle \cap (-P_t + P_t)$. Clearly, the right-hand side is contained in the left-hand side, so we only need to show the reverse inclusion. Let $x \in -P_t + P_t$. Then, for $k+1 \leq i \leq m$, we have $(\mathcal{M}_{\mathcal{D}}^{-1}(x))_i \in [-t, t] - [-t, t] = (-2t, 2t)$. If, further, $x \in \langle \mathcal{T} \rangle$, then $(\mathcal{M}_{\mathcal{D}}^{-1}(x))_i = 0$, and so $x \in \langle \mathcal{D}_0 \rangle$, as required.

For $v \in V = \mathbb{Z}^m / H$ and $k+1 \leq i \leq m$, define $(\mathcal{M}_{\mathcal{D}}^{-1}(v))_i = (\mathcal{M}_{\mathcal{D}}^{-1}(x))_i$ using any $x \in \mathbb{Z}^m$ such that $v = x + H$. This is well-defined due to Equation (4.6). By Lemma 4.7, for every $k+1 \leq i \leq m$ and $1 \leq j \leq m$,

$$|(\mathcal{M}_{\mathcal{D}}^{-1}(e_j + x) - \mathcal{M}_{\mathcal{D}}^{-1}(x))_i| \leq 1. \quad (4.8)$$

We turn to proving (i) and (ii).

(i) First, we show that

$$P_{t-1} \cdot y_0 \subseteq \text{Eq}(f_C). \quad (4.9)$$

Write $S = \{e_1, \dots, e_m\} \subseteq \mathbb{Z}^m$. Since the domain $P_t \cdot y_0$ of f_C is equal to the entire set Y , all of its points are internal in Y . Therefore, a point $y \in P_{t-1} \cdot y_0$ belongs to $\text{Eq}(f_C)$ if and only if f_C preserves $y \xrightarrow{s}$ for every $s \in S$. Hence, by Lemma 3.10, it is sufficient to prove that $\langle \mathcal{T} \rangle \cap (-P_t + S + P_{t-1}) = \langle \mathcal{D}_0 \rangle \cap (-P_t + S + P_{t-1})$. Let $x \in -P_t + S + P_{t-1}$. By Equation (4.8), $(\mathcal{M}_{\mathcal{D}}^{-1}(x))_i \in (-2t, 2t)$ for all $k+1 \leq i \leq m$. Using (4.6) and (4.7) as before, we see that $x \in \langle \mathcal{D}_0 \rangle$ if and only if $x \in \langle \mathcal{T} \rangle$, which proves Inclusion (4.9). Finally, using Lemma 4.10(ii), we get

$$\begin{aligned} |\text{Eq}(f_C)| &\geq |P_{t-1} \cdot y_0| = |P_{t-1}| \\ &= |P_1| \cdot (t-1)^{m - |\mathcal{D}_0|} \\ &= |P_t| \cdot \left(1 - \frac{1}{t}\right)^{m - |\mathcal{D}_0|} \\ &\geq |Y| \cdot \left(1 - \frac{m - |\mathcal{D}_0|}{t}\right). \end{aligned}$$

- (ii) Take an integer $\tilde{t} \geq 0$. First, we show that $B_V(P_t \cdot v_0, \tilde{t}) \subseteq P_{t+\tilde{t}} \cdot v_0$. Note that $P_r \cdot v_0 = \{v \in V \mid (\mathcal{M}_{\mathcal{D}}^{-1}(v))_i \in [-r, r) \text{ for } k+1 \leq i \leq m\}$ for every $r \in \mathbb{N}$. Let $v \in B_V(P_t \cdot v_0, \tilde{t})$. By Equation (4.8), $(\mathcal{M}_{\mathcal{D}}^{-1}(v))_i \in [-(t+\tilde{t}), t+\tilde{t})$, so $v \in P_{t+\tilde{t}} \cdot v_0$. Finally, since $\text{Im}(f_C) = P_t \cdot v_0$, using Lemma 4.10(ii) again, we get

$$\begin{aligned} |B_V(\text{Im}(f_C), \tilde{t})| &\leq |P_{t+\tilde{t}} \cdot v_0| \\ &= |P_t| \cdot \left(1 + \frac{\tilde{t}}{t}\right)^{m-|\mathcal{D}_0|} \\ &= |\text{Im}(f_C)| \cdot \left(1 + \frac{\tilde{t}}{t}\right)^{m-|\mathcal{D}_0|}. \end{aligned}$$

□

4.5 Tool B: The method of far-reaching short sets

We turn to *Tool B*, in which we are given a pointed, usually infinite, Γ -set (V, v_0) and a positive integer t . Our goal is to create a pointed Γ -set $(U, u_0) = (\mathbb{Z}^m / \langle \mathcal{D} \rangle, 0 + \langle \mathcal{D} \rangle)$ and a subgraph isomorphism $f_B : (P_t^{\mathcal{D}} \cdot u_0, u_0) \rightarrow (P_t^{\mathcal{D}} \cdot v_0, v_0)$. Here, $\mathcal{D} \subseteq \mathbb{Z}^m$ is a linearly independent set of our choice, and $P_t^{\mathcal{D}} \cdot v_0$ must be bounded within a ball of radius $O(t)$. By Lemma 4.10(i), a sufficient condition for this bound is that $\|\mathcal{D}\|_1 \leq O(t)$. Throughout this section, the reader should keep the discussion in the beginning of Section 4 in mind, and, in particular, the realization of Γ as the quotient \mathbb{Z}^m/K , and the constants C_d and t_E .

As before, we may assume that V is transitive, so it can be regarded as \mathbb{Z}^m/H for some $K \leq H \leq \mathbb{Z}^m$, with $v_0 = 0 + H$. Our strategy is to find a linearly independent set $\mathcal{D} \subseteq H$ of *short vectors*, such that the restriction f_B of the quotient map $U = \mathbb{Z}^m / \langle \mathcal{D} \rangle \rightarrow \mathbb{Z}^m / H = V$ to $P_t^{\mathcal{D}} \cdot u_0$ is a subgraph isomorphism. To show that such a set \mathcal{D} exists, we first prove Lemma 4.12, which asserts the existence of a subset $\mathcal{D} \subseteq H$, consisting of short vectors, such that all vectors in H up to a certain length $R = R(\|\mathcal{D}\|_1, t)$ are spanned by \mathcal{D} . Proposition 4.13 then shows that this set \mathcal{D} is suitable for our purposes.

Lemma 4.12. *Let $K \leq H \leq \mathbb{Z}^m$ and $t_E \leq t \in \mathbb{N}$. Then, there is a linearly independent subset $\mathcal{D} \subseteq H$, such that $K \leq \langle \mathcal{D} \rangle$ and the following holds:*

(i) $\|\mathcal{D}\|_1 \leq 2 \cdot 7^d \cdot d^{2d+2} \cdot t,$

(ii) $\langle \mathcal{D} \rangle \cap B_{\mathbb{Z}^m}(R) = H \cap B_{\mathbb{Z}^m}(R)$ for $R = 2 \cdot (\|\mathcal{D}\|_1 + dt) + 1.$

Proof. Let $\tau : \mathbb{Z}^m \rightarrow \mathbb{Z}^d \oplus \{0\}^{m-d}$ denote the projection onto the first d coordinates, and let $\bar{H} = \tau(H)$. For each $0 \leq i \leq d+1$, define $t_i = t \cdot (7d^2)^i$. Consider the monotone nondecreasing sequence of integers

$$\{\text{rank}(\bar{H} \cap B_{\mathbb{Z}^m}(t_i))\}_{i=0}^{d+1}.$$

Since the last element is at most d , there must be an integer $0 \leq i \leq d$ such that

$$\text{rank}\langle \bar{H} \cap B_{\mathbb{Z}^m}(t_i) \rangle = \text{rank}\langle \bar{H} \cap B_{\mathbb{Z}^m}(t_{i+1}) \rangle .$$

Fix such an integer i , and define $G = \langle \bar{H} \cap B_{\mathbb{Z}^m}(t_{i+1}) \rangle$. In particular, $G \cap B_{\mathbb{Z}^m}(t_i) = \bar{H} \cap B_{\mathbb{Z}^m}(t_i)$, and so,

$$\text{rank}\langle G \cap B_{\mathbb{Z}^m}(t_i) \rangle = \text{rank}\langle \bar{H} \cap B_{\mathbb{Z}^m}(t_i) \rangle = \text{rank}\langle \bar{H} \cap B_{\mathbb{Z}^m}(t_{i+1}) \rangle = \text{rank}G .$$

Hence, $\text{rank}\langle G \cap B_{\mathbb{Z}^m}^{L^2}(t_i) \rangle = \text{rank}G$, and so Proposition 4.5 yields a basis $\bar{\mathcal{D}}_0$ for G satisfying

$$\|\bar{\mathcal{D}}_0\|_1 \leq d^2 \cdot t_i . \quad (4.10)$$

Recall the definition of the set $T \subseteq \mathbb{Z}^m$ from the beginning of Section 4. Construct a set $\mathcal{D}_0 \subseteq H$ consisting of one preimage $h \in H$ for each $\bar{h} \in \bar{\mathcal{D}}_0$ (i.e., $\tau(h) = \bar{h}$), such that $h - \bar{h} \in T$ (this is possible since $K \leq H$). So, each of the last $m - d$ coordinates of each $h \in \mathcal{D}_0$ is in the range $0, \dots, \beta_m - 1$. Now, define $H_T = H \cap (\{0\}^d \oplus \mathbb{Z}^{m-d})$. Then, $K \leq H_T$, and so $\text{rank}H_T = m - d = \text{rank}K$. Since K has a basis $\beta_{d+1} \cdot e_{d+1}, \dots, \beta_m \cdot e_m$, which is contained in $B_{\{0\}^d \oplus \mathbb{Z}^{m-d}}^{L^2}(\beta_E)$, we conclude from Proposition 4.5 that H_T has a basis \mathcal{D}_1 , with $\|\mathcal{D}_1\|_1 \leq (m - d)^2 \cdot \beta_E$. Then, the set $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{D}_1$ is linearly independent, and

$$\begin{aligned} \|\mathcal{D}\|_1 &= \|\mathcal{D}_0\|_1 + \|\mathcal{D}_1\|_1 \\ &\leq \sum_{x \in \bar{\mathcal{D}}_0} (\|x\|_1 + (m - d) \cdot \beta_E) + (m - d)^2 \cdot \beta_E \\ &= \|\bar{\mathcal{D}}_0\|_1 + (|\bar{\mathcal{D}}_0| \cdot (m - d) + (m - d)^2) \cdot \beta_E \\ &\leq \|\bar{\mathcal{D}}_0\|_1 + (d + (m - d)) \cdot (m - d) \cdot \beta_E \\ &\leq d^2 \cdot t_i + m \cdot (m - d) \cdot \beta_E , \end{aligned}$$

where the last inequality follows from Equation (4.10). Now, $t \geq t_E$ implies that $m \cdot (m - d) \cdot \beta_E \leq d \cdot t$, and so,

$$\|\mathcal{D}\|_1 \leq d^2 \cdot t_i + d \cdot t \leq d^2 \cdot (7d^2)^d \cdot t + d \cdot t \leq 2 \cdot 7^d \cdot d^{2d+2} \cdot t ,$$

proving (i). Furthermore, for $R = 2 \cdot (\|\mathcal{D}\|_1 + dt) + 1$,

$$\begin{aligned} R &\leq 2 \cdot (d^2 \cdot t_i + m \cdot (m - d) \cdot \beta_E + dt) + 1 \\ &\leq 2 \cdot (3 \cdot d^2 \cdot t_i) + 1 \leq 7 \cdot d^2 \cdot t_i = t_{i+1} . \end{aligned}$$

Hence, to prove (ii) it suffices to show that

$$\langle \mathcal{D} \rangle \cap B_{\mathbb{Z}^m}(t_{i+1}) = H \cap B_{\mathbb{Z}^m}(t_{i+1}) .$$

Let $h \in H \cap B_{\mathbb{Z}^m}(t_{i+1})$ and write $\bar{h}_0 = \tau(h)$. So,

$$\bar{h}_0 \in \bar{H} \cap B_{\mathbb{Z}^m}(t_{i+1}) \subseteq G = \langle \bar{\mathcal{D}}_0 \rangle .$$

Hence, there exists $h_0 \in \langle \bar{\mathcal{D}}_0 \rangle$ such that $\tau(h_0) = \bar{h}_0$, and so $h - h_0 \in H_T = \langle \mathcal{D}_1 \rangle$. Hence, $h = h_0 + (h - h_0) \in \langle \mathcal{D} \rangle$, proving the claim. \square

The following proposition yields Tool B.

Proposition 4.13. *Let $K \leq H \leq \mathbb{Z}^m$ and $t_E \leq t \in \mathbb{N}$. Write $V = \mathbb{Z}^m/H$ and $v_0 = 0 + H$. Then, there is a linearly independent set $\mathcal{D} \subseteq H$, $K \leq \langle \mathcal{D} \rangle$, satisfying the following:*

- (i) $P_t^{\mathcal{D}} \subseteq B_{\mathbb{Z}^m}(C_d \cdot t)$.
- (ii) For $U = \mathbb{Z}^m/\langle \mathcal{D} \rangle$ and $u_0 = 0 + \langle \mathcal{D} \rangle$, there is a subgraph isomorphism $f_B : P_t^{\mathcal{D}} \cdot u_0 \rightarrow P_t^{\mathcal{D}} \cdot v_0$ (given by $f_B = F_{P_t^{\mathcal{D}}, u_0, v_0}$).

Proof. Apply Lemma 4.12 to H and t to obtain a linearly independent set \mathcal{D} with $K \leq \langle \mathcal{D} \rangle$ such that

$$\|\mathcal{D}\|_1 \leq 2 \cdot 7^d \cdot d^{2d+2} \cdot t$$

and

$$\langle \mathcal{D} \rangle \cap B_{\mathbb{Z}^m}(R) = H \cap B_{\mathbb{Z}^m}(R) , \quad (4.11)$$

where $R = 2 \cdot (\|\mathcal{D}\|_1 + dt) + 1$. Write $P = P_t^{\mathcal{D}}$. The first claim follows from Lemma 4.10(i) since

$$P \subseteq B_{\mathbb{Z}^m}(\|\mathcal{D}\|_1 + (m - |\mathcal{D}|) \cdot t) \subseteq B_{\mathbb{Z}^m}(\|\mathcal{D}\|_1 + d \cdot t) \quad (4.12)$$

and

$$\|\mathcal{D}\|_1 + d \cdot t \leq 2 \cdot 7^d \cdot d^{2d+2} \cdot t + d \cdot t \leq C_d \cdot t .$$

We turn to proving that $F_{P_t^{\mathcal{D}}, u_0, v_0}$ is a subgraph isomorphism. By Lemma 3.11, it suffices to show that, for $S_1 = \{e_1, \dots, e_m\} \cup \{0\} \subseteq \mathbb{Z}^m$,

$$H \cap (-P + S_1 + P) = \langle \mathcal{D} \rangle \cap (-P + S_1 + P) .$$

Equation (4.12) yields $-P + S_1 + P \subseteq B_{\mathbb{Z}^m}(2 \cdot (\|\mathcal{D}\|_1 + dt) + 1) = B_{\mathbb{Z}^m}(R)$, hence the claim follows from Equation (4.11). \square

4.6 Tool A: The bounded-addition method

We turn to *Tool A*. Here, we are given a pointed \mathbb{F}_m -set (X, x_0) and $t \in \mathbb{N}$ such that

$$\text{Box}_{\mathbb{F}_m}(t) \cdot x_0 \subseteq X_E , \quad (4.13)$$

i.e., X abides by the set of equations E within a certain neighborhood of x_0 . Our goal is to construct a pointed Γ -set (U, u_0) and a subgraph isomorphism $f_A : B_U(u_0, r) \rightarrow B_X(x_0, r)$ for some $r \geq \Omega(t)$. A key notion in our proof is “bounded addition”, formalized in the following definition:

Definition 4.14. Let $\mathcal{D} \subseteq B_{\mathbb{Z}^m}^L(R)$, $R \in \mathbb{N}$. Define $[\mathcal{D}]_R \subseteq B_{\mathbb{Z}^m}^L(R)$ as the minimal subset of \mathbb{Z}^m satisfying the following conditions:

- (I) $\mathcal{D} \cup \{0\} \subset [\mathcal{D}]_R$.

(II) If $x \in [\mathcal{D}]_R$, then $-x \in [\mathcal{D}]_R$.

(III) If $x_1, x_2 \in [\mathcal{D}]_R$ and $\|x_1 + x_2\|_\infty \leq R$, then $x_1 + x_2 \in [\mathcal{D}]_R$.

Informally, the relevance of Definition 4.14 to the problem Tool A aims to solve is the following: Inclusion (4.13) yields a guarantee on the behavior of X in a neighborhood of x_0 (see Lemma 4.20). Therefore, we would like to know which elements of \mathbb{Z}^m can be generated from a given finite subset $\mathcal{D} \subseteq \mathbb{Z}^m$ without straying far from the origin.

We proceed to develop *bounded addition* in order to prove Lemma 4.19 below, and then use it to provide Tool A. We begin with two immediate observations about $[\mathcal{D}]_R$ without proof :

Lemma 4.15. *Let $\mathcal{D} \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R)$, $R \in \mathbb{N}$. Then:*

(i) *A vector $y \in \mathbb{Z}^m$ belongs to $[\mathcal{D}]_R$ if and only if there is a sequence $\{y_i\}_{i=1}^k \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R)$ with $y = y_k$, in which every element y_i satisfies at least one of the following:*

(a) $y_i \in \mathcal{D}$, or

(b) y_i is the negation of a previous element, or

(c) y_i is the sum of two previous elements.

(ii) *For $\mathcal{E} \subseteq \mathbb{Z}^m$, if $\mathcal{E} \subseteq [\mathcal{D}]_R$ then $[\mathcal{E}]_R \subseteq [\mathcal{D}]_R$.*

The sequence $\{y_i\}_{i=1}^k$ in Lemma 4.15(i) is called a (\mathcal{D}, R) -*generation-sequence* for y .

Lemma 4.16 provides a technical result, which is then used by Lemmas 4.17, 4.18 and 4.19 to give sufficient conditions for membership in $[\mathcal{D}]_R$. The proof of Lemma 4.16 is inspired by *Thomas Jefferson's method* for proportional allocation of seats in the United States House of Representatives – a method used in many countries to this day. Recall that \mathcal{D}^\pm is the set of elements of \mathcal{D} and their negations.

Lemma 4.16. *Let $\mathcal{D} \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R)$, $R \in \mathbb{N}$, and let $y \in \langle \mathcal{D} \rangle$. Then, there is a finite sequence x_1, \dots, x_N of elements of \mathcal{D}^\pm , such that $y = \sum_{j=1}^N x_j$, and for every $0 \leq k \leq N$,*

$$\left\| \sum_{j=1}^k x_j - \frac{k}{N} \cdot y \right\|_\infty \leq 2 \cdot |\mathcal{D}| \cdot R . \quad (4.14)$$

Proof. Let M be a multiset of elements of \mathcal{D}^\pm whose sum is y . Assume, without loss of generality, that M does not contain both x and $-x$ for any $x \in \mathcal{D}$. Write $\bar{x}_1, \dots, \bar{x}_p$ ($p \leq |\mathcal{D}|$) for the distinct elements of M , and a_i for the multiplicity of \bar{x}_i in M . Thus, $y = \sum_{i=1}^p a_i \cdot \bar{x}_i$. Let $N = |M| = \sum_{i=1}^p a_i$. We seek to order the elements of M in a sequence x_1, \dots, x_N satisfying the claim.

Let

$$F = \left\{ \left(N \cdot \frac{b}{a_i}, i \right) \in [0, N] \times [p] \mid i \in [p] \text{ and } 1 \leq b \leq a_i \right\} .$$

It is helpful to think of an element $(z, i) \in F$ as a flag of color i , located at $z \in [0, N]$. So, for every $i \in [p]$, we have a_i flags of color i , positioned evenly from $\frac{N}{a_i}$ to N . Let $(z_1, i_1), \dots, (z_N, i_N)$ be a sequence, consisting of all elements of F , ordered so that $z_1 \leq z_2 \leq \dots \leq z_N$. For $1 \leq j \leq N$, we set $x_j = \bar{x}_{i_j}$. Note that $y = \sum_{j=1}^N x_j$, so we only need to show that Equation (4.14) holds.

First, we would like to show that

$$\forall j \in [N] \quad z_j \geq j. \quad (4.15)$$

For $z \in [0, N]$, let $w_z = |\{j \in [N] \mid z_j \leq z\}|$, i.e., the number of flags up to location z . We claim that $w_z \leq z$. Indeed, $|\{j \in [N] \mid z_j \leq z \text{ and } i_j = i\}| = \lfloor z \cdot \frac{a_i}{N} \rfloor$ for $i \in [p]$. Summing over i yields

$$w_z = \sum_{i=1}^p \left\lfloor z \cdot \frac{a_i}{N} \right\rfloor \leq \sum_{i=1}^p z \cdot \frac{a_i}{N} = z.$$

In particular, taking $z = z_j$, we have $w_{z_j} \leq z_j$, but clearly $w_{z_j} \geq j$, and so Inequality (4.15) follows.

For $i \in [p]$ and $l \in [N]$, define $b_{i,l} = |\{1 \leq j \leq l \mid i_j = i\}|$, that is, the number of i -colored flags among the first l flags. Let $k \in [N]$. We seek to bound the L^1 -distance Δ between the vectors

$$(b_{i,k})_{i=1}^p \in \mathbb{R}^p \quad \text{and} \quad \left(\frac{k}{N} \cdot a_i \right)_{i=1}^p \in \mathbb{R}^p. \quad (4.16)$$

Let $i \in [p]$. We would like to show that

$$b_{i,k} \geq \frac{k}{N} \cdot a_i - 1. \quad (4.17)$$

This clearly holds if $k < \frac{N}{a_i}$. Assume that $k \geq \frac{N}{a_i}$ and define $k' = \lfloor \frac{k}{N} \cdot a_i \rfloor$. Let j be minimal such that $b_{i,j} = k'$. In other words, the k' -th i -colored flag is j -th among all flags. Note that $z_j = \frac{N}{a_i} \cdot k'$. By Inequality (4.15), $j \leq z_j = \frac{N}{a_i} \cdot k' \leq k$. Equation (4.17) follows as

$$b_{i,k} \geq b_{i,j} = k' \geq \frac{k}{N} \cdot a_i - 1.$$

Since the two vectors in (4.16) have the same sum k , the distance between them is

$$\Delta = \sum_{i=1}^p \left| \frac{k}{N} \cdot a_i - b_{i,k} \right| = 2 \cdot \sum_{i=1}^p \max \left\{ 0, \frac{k}{N} \cdot a_i - b_{i,k} \right\},$$

so $\Delta \leq 2p$ due to Equation (4.17). Now,

$$\begin{aligned} \left\| \sum_{j=1}^k x_j - \frac{k}{N} \cdot y \right\|_{\infty} &= \left\| \sum_{i=1}^p b_{i,k} \cdot x_i - \frac{k}{N} \cdot \sum_{i=1}^p a_i \cdot x_i \right\|_{\infty} = \left\| \sum_{i=1}^p \left(b_{i,k} - \frac{k}{N} \cdot a_i \right) \cdot x_i \right\|_{\infty} \\ &\leq \sum_{i=1}^p \left| b_{i,k} - \frac{k}{N} \cdot a_i \right| \cdot R \\ &= \Delta \cdot R \leq 2p \cdot R \leq 2 \cdot |\mathcal{D}| \cdot R. \end{aligned}$$

□

Lemma 4.17. *Let $\mathcal{D} \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R)$, $R \in \mathbb{N}$, and $y \in \langle \mathcal{D} \rangle$. Then, $y \in [\mathcal{D}]_{2 \cdot |\mathcal{D}| \cdot R + \|y\|_\infty}$.*

Proof. By Lemma 4.16, there is a sequence x_1, \dots, x_m of elements of \mathcal{D}^\pm , such that $y = \sum_{j=1}^m x_j$, and for every $0 \leq k \leq m$, we have $\|\sum_{j=1}^k x_j - \frac{k}{m} \cdot y\|_\infty \leq 2 \cdot |\mathcal{D}| \cdot R$. Consequently, $\|\sum_{j=1}^k x_j\|_\infty \leq 2 \cdot |\mathcal{D}| \cdot R + \|y\|_\infty$, so by induction on k , each of the partial sums $\sum_{j=1}^k x_j$ belongs to $[\mathcal{D}]_{2 \cdot |\mathcal{D}| \cdot R + \|y\|_\infty}$. In particular, $k = m$ yields the claim. □

Proposition 4.5 says that a lattice in \mathbb{Z}^m which is generated by short elements has a short basis. The following lemma is an analogue statement in the context of bounded addition.

Lemma 4.18. *Let $\mathcal{D} \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R)$, $R \in \mathbb{N}$. The lattice $\langle \mathcal{D} \rangle$ admits a basis contained in $[\mathcal{D}]_{5m^2 \cdot R} \cap B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$.*

Proof. Write $\mathcal{D} = \{x_1, \dots, x_p\}$. For $0 \leq i \leq p$, let $\mathcal{D}_i = \{x_1, \dots, x_i\}$. We prove by induction on i that $\langle \mathcal{D}_i \rangle$ has a basis \mathcal{E}_i contained in $[\mathcal{D}]_{5m^2 \cdot R} \cap B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$. Taking $\mathcal{E}_0 = \emptyset$, the base case $i = 0$ is immediate. Assume that $i \geq 1$ and that $\langle \mathcal{D}_{i-1} \rangle$ has a basis \mathcal{E}_{i-1} contained in $[\mathcal{D}]_{5m^2 \cdot R} \cap B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$. We consider two generating sets for $\langle \mathcal{D}_i \rangle$:

- (I) The set \mathcal{D}_i , which is contained in $B_{\mathbb{Z}^m}^{L^\infty}(R)$.
- (II) The set $\mathcal{E}_{i-1} \cup \{x_i\}$, which is contained in $B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$, and has at most $m + 1$ elements.

By virtue of the former generating set and by Proposition 4.5, $\langle \mathcal{D}_i \rangle$ has a basis $\mathcal{E}_i \subseteq B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R) \subseteq B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$. It is now enough to show that $\mathcal{E}_i \subseteq [\mathcal{D}]_{5m^2 \cdot R}$. Let $y \in \mathcal{E}_i$. Then, $\|y\|_\infty \leq m \cdot R$, and since $y \in \mathcal{E}_i \subseteq \langle \mathcal{D}_i \rangle$, we have $y \in \langle \mathcal{E}_{i-1} \cup \{x_i\} \rangle$. As $\langle \mathcal{E}_i \cup \{x_i\} \rangle \subseteq B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$ and $|\mathcal{E}_{i-1} \cup \{x_i\}| \leq m + 1$, Lemma 4.17 implies that

$$y \in [\mathcal{E}_{i-1} \cup \{x_i\}]_{2(m+1) \cdot mR + mR} \subseteq [\mathcal{E}_{i-1} \cup \{x_i\}]_{5m^2 \cdot R}.$$

As $\mathcal{E}_{i-1} \subseteq [\mathcal{D}]_{5m^2 \cdot R}$ by the induction hypothesis, and since $\{x_i\} \subseteq [\mathcal{D}]_{5m^2 \cdot R}$, it follows from Lemma 4.15(ii) that $y \in [\mathcal{D}]_{5m^2 \cdot R}$. □

Lemma 4.19. *Let $\mathcal{D} \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R)$, $R \in \mathbb{N}$. Then, $\langle \mathcal{D} \rangle \cap B_{\mathbb{Z}^m}^{L^\infty}(R) = [\mathcal{D}]_{5m^2 \cdot R} \cap B_{\mathbb{Z}^m}^{L^\infty}(R)$.*

Proof. Given $x \in \langle \mathcal{D} \rangle \cap B_{\mathbb{Z}^m}^{L^\infty}(R)$, we need to show that $x \in [\mathcal{D}]_{5m^2 \cdot R}$. Lemma 4.18 yields a basis $\mathcal{E} \subseteq [\mathcal{D}]_{5m^2 \cdot R} \cap B_{\mathbb{Z}^m}^{L^\infty}(m \cdot R)$ for $\langle \mathcal{D} \rangle$. Since $x \in \langle \mathcal{E} \rangle$, Lemma 4.17 implies that $x \in [\mathcal{E}]_{2|\mathcal{E}| \cdot mR + R} \subseteq [\mathcal{E}]_{5m^2 \cdot R}$. Lemma 4.15(ii) yields the claim, since $\mathcal{E} \subseteq [\mathcal{D}]_{5m^2 \cdot R}$. □

At this point, we have established the required groundwork concerning bounded addition. Before proving our main proposition about Tool A, we also need the following lemma. It states that, within X_E , an \mathbb{F}_m -set X behaves in a certain sense like a \mathbb{Z}^m -set. The reader should recall, from Section 4.2, the definition of a *sorted word* and the notation \hat{w} for a given $w \in \mathbb{F}_m$.

Lemma 4.20. *Let X be an \mathbb{F}_m -set, $x \in X$ and $R \in \mathbb{N}$. Assume that $\text{Box}_{\mathbb{F}_m}(R) \cdot x \subseteq X_E$. Then, $w \cdot x = \hat{w} \cdot x$ for every $w \in B_{\mathbb{F}_m}(R)$.*

Proof. First, we fix some notation: For a word $w \in \mathbb{F}_m$, whose reduced form is $w = \hat{e}_{i_1}^{\epsilon_1} \cdots \hat{e}_{i_k}^{\epsilon_k}$, $i_1, \dots, i_k \in [m]$ and $\epsilon_1, \dots, \epsilon_k \in \{+1, -1\}$, write $\iota(w)$ for the number of *inversions* in w , namely,

$$\iota(w) = |\{(j_1, j_2) \mid 1 \leq j_1 < j_2 \leq k \text{ and } i_{j_1} > i_{j_2}\}| .$$

Let $w \in B_{\mathbb{F}_m}(R)$ and write $w = \hat{e}_{i_1}^{\epsilon_1} \cdots \hat{e}_{i_k}^{\epsilon_k}$ as above ($k \leq R$). If $\iota(w) = 0$, then w is sorted, i.e., $w = \hat{w}$, and we are done. Otherwise, take the maximal $1 \leq l \leq k - 1$ for which $i_l > i_{l+1}$. Let w_{l+1} denote the suffix $\hat{e}_{i_{l+1}}^{\epsilon_{l+1}} \cdots \hat{e}_{i_k}^{\epsilon_k}$ of w . Then w_{l+1} is a sorted word by the definition of l , and so $w_{l+1} \in \text{Box}_{\mathbb{F}_m}(R)$, implying that $w_{l+1} \cdot x \in X_E$. Hence, $\left[\hat{e}_{i_l}^{-\epsilon_l}, \hat{e}_{i_{l+1}}^{\epsilon_{l+1}}\right] \cdot w_{l+1} \cdot x = w_{l+1} \cdot x$. Define $w' \in \mathbb{F}_m$ by $w' = \hat{e}_{i_1}^{\epsilon_1} \cdots \hat{e}_{i_l}^{\epsilon_l} \cdot \left[\hat{e}_{i_l}^{-\epsilon_l}, \hat{e}_{i_{l+1}}^{\epsilon_{l+1}}\right] \cdot w_{l+1} = \hat{e}_{i_1}^{\epsilon_1} \cdots \hat{e}_{i_{l-1}}^{\epsilon_{l-1}} \cdot \hat{e}_{i_{l+1}}^{\epsilon_{l+1}} \cdot \hat{e}_{i_l}^{\epsilon_l} \cdot \hat{e}_{i_{l+2}}^{\epsilon_{l+2}} \cdots \hat{e}_{i_k}^{\epsilon_k}$. Then, $\widehat{w'} = \hat{w}$, $\iota(w') < \iota(w)$ and

$$\begin{aligned} w' \cdot x &= \hat{e}_{i_1}^{\epsilon_1} \cdots \hat{e}_{i_l}^{\epsilon_l} \cdot \left(\left[\hat{e}_{i_l}^{-\epsilon_l}, \hat{e}_{i_{l+1}}^{\epsilon_{l+1}}\right] \cdot w_{l+1} \cdot x\right) \\ &= \hat{e}_{i_1}^{\epsilon_1} \cdots \hat{e}_{i_l}^{\epsilon_l} \cdot (w_{l+1} \cdot x) \\ &= w \cdot x . \end{aligned}$$

Therefore, the claim follows by induction. \square

We turn to our main statement in this section, which yields Tool A.

Proposition 4.21. *Let X be an \mathbb{F}_m -set, $x_0 \in X$ and $r \in \mathbb{N}$. Assume that $\text{Box}_{\mathbb{F}_m}(30m^3 \cdot r) \cdot x_0 \subseteq X_E$. Let $H = \langle \pi(\text{Stab}_{\mathbb{F}_m}(x_0) \cap B_{\mathbb{F}_m}(2r+1)) \rangle \leq \mathbb{Z}^m$ and define the pointed \mathbb{Z}^m -set $(U, u_0) = (\mathbb{Z}^m/H, 0 + H)$. Then, there is a subgraph isomorphism $f_A : B_U(u_0, r) \rightarrow B_X(x_0, r)$ (given by $f_A = F_{B_{\mathbb{F}_m}(r), u_0, x_0}$).*

Proof. Write $R = 2r + 1$, and note that

$$\text{Box}_{\mathbb{F}_m}(10m^3 \cdot R) \cdot x_0 \subseteq X_E .$$

Recall the natural surjection $\pi : \mathbb{F}_m \rightarrow \mathbb{Z}^m$, defined in the beginning of Section 4. Note that $\text{Stab}_{\mathbb{F}_m}(u_0) = \pi^{-1}(H)$. Hence, by Lemma 3.12, it suffices to show that

$$\text{Stab}_{\mathbb{F}_m}(x_0) \cap B_{\mathbb{F}_m}(R) = \pi^{-1}(H) \cap B_{\mathbb{F}_m}(R) . \quad (4.18)$$

The \subseteq inclusion in Equation (4.18) is clear from the definition of H . We proceed to prove the \supseteq inclusion.

Define $\mathcal{D} = \pi(\text{Stab}_{\mathbb{F}_m}(x_0) \cap B_{\mathbb{F}_m}(R))$, implying $H = \langle \mathcal{D} \rangle \leq \mathbb{Z}^m$. Note that

$$\mathcal{D} \subseteq B_{\mathbb{Z}^m}^{L^\infty}(R). \quad (4.19)$$

Let $w \in \pi^{-1}(H) \cap B_{\mathbb{F}_m}(R)$. It suffices to prove that $w \cdot x_0 = x_0$. Now,

$$\pi(w) \in H \cap B_{\mathbb{Z}^m}^{L^\infty}(R) = \langle \mathcal{D} \rangle \cap B_{\mathbb{Z}^m}^{L^\infty}(R). \quad (4.20)$$

By Lemma 4.19, Equations (4.19) and (4.20) imply that $\pi(w) \in [\mathcal{D}]_{5m^2 \cdot R}$. Let $v_1, \dots, v_k \in B_{\mathbb{Z}^m}^{L^\infty}(5m^2 \cdot R)$ be a $(\mathcal{D}, 5m^2 \cdot R)$ -generation-sequence for $\pi(w)$. In particular, $v_k = \pi(w)$. We shall construct, inductively, a sequence of *sorted* words $w_1, \dots, w_k \in \mathbb{F}_m$, such that for each $1 \leq i \leq k$,

$$w_i \in B_{\mathbb{F}_m}(5m^3 \cdot R), \quad \pi(w_i) = v_i \quad \text{and} \quad w_i \cdot x_0 = x_0. \quad (4.21)$$

Let $1 \leq i \leq k$, and assume that a sequence of sorted words $w_1, \dots, w_{i-1} \in \mathbb{F}_m$, satisfying (4.21), has been constructed. We define w_i by considering three separate cases:

- (I) Assume that $v_i \in \mathcal{D}$. By the definition of \mathcal{D} , there is $u \in \text{Stab}_{\mathbb{F}_m}(x_0) \cap B_{\mathbb{F}_m}(R)$ satisfying $\pi(u) = v_i$. Since $\text{Box}_{\mathbb{F}_m}(R) \cdot x_0 \subseteq X_E$, Lemma 4.20 implies that $\hat{u} \cdot x_0 = u \cdot x_0$. Define $w_i = \hat{u}$. Then, $w_i \cdot x_0 = x_0$ since $u \in \text{Stab}_{\mathbb{F}_m}(x_0)$, and $w_i \in B_{\mathbb{F}_m}(R)$ since $u \in B_{\mathbb{F}_m}(R)$.
- (II) Otherwise, assume that there is $1 \leq j < i$ for which $v_i = -v_j$. By the induction hypothesis, $w_j \in B_{\mathbb{F}_m}(5m^3 \cdot R)$ and $w_j \cdot x_0 = x_0$. So, the same holds for w_j^{-1} . Since $\text{Box}_{\mathbb{F}_m}(5m^3 \cdot R) \cdot x_0 \subseteq X_E$, Lemma 4.20 implies that $\widehat{w_j^{-1}} \cdot x_0 = w_j^{-1} \cdot x_0 = x_0$. Define $w_i = \widehat{w_j^{-1}}$.
- (III) Otherwise, there must be integers $j_1, j_2 \in [i-1]$ such that $v_i = v_{j_1} + v_{j_2}$. Define $u = w_{j_1} \cdot w_{j_2} \in B_{\mathbb{F}_m}(10m^3 \cdot R)$. Then,

$$\|\pi(u)\|_1 = \|v_{j_1} + v_{j_2}\|_1 = \|v_i\|_1 \leq m \cdot \|v_i\|_\infty \leq 5m^3 \cdot R.$$

This means that, although the length of $u \in \mathbb{F}_m$ is merely bounded by $10m^3 \cdot R$, the length of the sorted word $\hat{u} \in \mathbb{F}_m$ is bounded by $5m^3 \cdot R$. Define $w_i = \hat{u}$. Since $\text{Box}_{\mathbb{F}_m}(10m^3 \cdot R) \cdot x_0 \subseteq X_E$, Lemma 4.20 implies that, $w_i \cdot x_0 = u \cdot x_0 = w_{j_1} \cdot w_{j_2} \cdot x_0 = x_0$.

Finally, $\hat{w} = w_k$ since w_k is sorted and $\pi(w_k) = v_k = \pi(w)$. As $w \in B_{\mathbb{F}_m}(R)$ and $\text{Box}_{\mathbb{F}_m}(R) \cdot x_0 \subseteq X_E$, Lemma 4.20 implies that $w \cdot x_0 = \hat{w} \cdot x_0$. Thus, $w \cdot x_0 = w_k \cdot x_0 = x_0$, as claimed. \square

4.7 The tiling algorithm - proof of the main theorem

We can now implement the algorithm outlined in Section 4.1.1. The reader should recall the definitions and objects fixed in the beginning of Section 4. In particular we shall refer to the equation-set E , the constants C_d and t_E , the integers $\{\beta_i\}_{i=d+1}^m$, the generators $\{\hat{e}_1, \dots, \hat{e}_m\}$ of \mathbb{F}_m and the set $\hat{T} \subseteq \mathbb{F}_m$. Given an \mathbb{F}_m -set X , we first discuss the injection of a single tile into X by means of Tools A, B and C.

Definition 4.22. Let X be an \mathbb{F}_m -set, $x \in X$, and $t \in \mathbb{N}$.

- (i) We say that x admits a t -parallelotope if there is a linearly independent set $\mathcal{D} \subseteq \mathbb{Z}^m$, where $K \leq \langle \mathcal{D} \rangle$, such that the pointed subset $(P_t^{\mathcal{D}} \cdot u_0, u_0)$ of $(U, u_0) = (\mathbb{Z}^m / \langle \mathcal{D} \rangle, 0 + \langle \mathcal{D} \rangle)$ is subgraph isomorphic to $(\hat{P}_t^{\mathcal{D}} \cdot x, x)$, and

$$\hat{P}_t^{\mathcal{D}} \subseteq B_{\mathbb{F}_m}(C_d \cdot t) . \quad (4.22)$$

More elaborately, we also say that x admits the t -parallelotope $P_t^{\mathcal{D}}$.

- (ii) Fix some arbitrary well-ordering \prec on finite subsets of \mathbb{Z}^m , e.g., let \prec order \mathbb{Z}^m -subsets lexicographically with regard to some well-ordering on \mathbb{Z}^m itself. If x admits a t -parallelotope, we denote by $\mathcal{D}_{x,t}$ the \prec -minimal linearly independent set \mathcal{D} for which x admits the t -parallelotope $P_t^{\mathcal{D}}$.

Due to Equation (4.22), the collection of sets \mathcal{D} satisfying condition (i) above depends only on the ball of radius $C_d \cdot t$ centered at x . Since we take $\mathcal{D}_{x,t}$ to be the \prec -minimum of this collection, the following simple fact follows.

Lemma 4.23. Let (X, x) and (\tilde{X}, \tilde{x}) be pointed \mathbb{F}_m -sets. Let $t \in \mathbb{N}$, and assume that $B_X(x, C_d \cdot t)$ and $B_{\tilde{X}}(\tilde{x}, C_d \cdot t)$ are subgraph isomorphic as pointed sets, and that x admits a t -parallelotope. Then, \tilde{x} also admits a t -parallelotope and $\mathcal{D}_{x,t} = \mathcal{D}_{\tilde{x},t}$.

In light of Definition 4.22, Proposition 4.13 can be rephrased as “For every $t \geq t_E$, every point in a Γ -set admits a t -parallelotope”.

Definition 4.24. Let X be an \mathbb{F}_m -set and $t \in \mathbb{N}$. A t -tile in X is a pair (x, f) such that:

- (i) $x \in X$ admits a $2t$ -parallelotope, and
(ii) f is a bijection from some finite Γ -set onto $\hat{P}_t^{\mathcal{D}} \cdot x \subseteq X$, where $\mathcal{D} = \mathcal{D}_{x,2t}$.

Note that, in the above definition we take \mathcal{D} to be $\mathcal{D}_{x,2t}$, rather than $\mathcal{D}_{x,t}$. This “extra length” will be useful later in controlling the amount of interference between tiles.

As explained in Section 4.1.1, for a tile (x, f) , we want the set $\text{Eq}(f)$ to be large. Also, we are interested in choosing, from among all tiles, a large collection of *pairwise disjoint* tiles. To this end, we seek to minimize the interference between tiles. The sets defined below are used to measure this interference:

Definition 4.25. Let $A \subseteq X$ for some \mathbb{F}_m -set X . For $t \in \mathbb{N}$, let

$$\eta_t(A) = A \cup \{x \in X \mid \text{There exists a } t\text{-tile } (x, f) \text{ such that } \text{Im}f \cap A \neq \emptyset\} .$$

We turn to prove the existence of tiles with good parameters (Proposition 4.27). We require the following observation.

Lemma 4.26. *Let X be an \mathbb{F}_m -set, $x \in X$ and $t \in \mathbb{N}$. Assume that*

$$\text{Box}_{\mathbb{F}_d}(t) \cdot \hat{T} \cdot x \subseteq X_E . \quad (4.23)$$

Then,

$$\text{Box}_{\mathbb{F}_m}(t) \subseteq X_E .$$

Proof. For $d+1 \leq i \leq m+1$, let

$$\hat{T}_i = \left\{ \prod_{j=i}^m \hat{e}_j^{\alpha_j} \mid 0 \leq \alpha_j < \beta_j \right\}$$

$$\hat{T}_i^\infty = \left\{ \prod_{j=i}^m \hat{e}_j^{\alpha_j} \mid \alpha_j \in \mathbb{Z} \right\}$$

Clearly, $\hat{T}_i \cdot x \subseteq \hat{T}_i^\infty \cdot x$ for every $d+1 \leq i \leq m+1$. We show, by induction on $i = m+1, \dots, d+1$, that this is in fact an equality. The lemma then follows since

$$\text{Box}_{\mathbb{F}_m}(t) \cdot x \subseteq \text{Box}_{\mathbb{F}_d}(t) \cdot \hat{T}_{d+1}^\infty \cdot x = \text{Box}_{\mathbb{F}_d}(t) \cdot \hat{T}_{d+1} \cdot x = \text{Box}_{\mathbb{F}_d}(t) \cdot \hat{T} \cdot x .$$

The base case $i = m+1$ is trivial, as $\hat{T}_{m+1} = \hat{T}_{m+1}^\infty = \{1_{\mathbb{F}_m}\}$. Let $d+1 \leq i \leq m$ and assume that $\hat{T}_{i+1} \cdot x = \hat{T}_{i+1}^\infty \cdot x$. Let $w \in \hat{T}_i^\infty$. Note that we can write $w = \hat{e}_i^{\alpha_i} \cdot v$, where $\alpha_i \in \mathbb{Z}$ and $v \in \hat{T}_{i+1}^\infty$. By the induction hypothesis $v \cdot x \in \hat{T}_{i+1} \cdot x$. Since $\hat{T}_{i+1} \subseteq \hat{T}$, it follows from Equation (4.23) that $v \cdot x \in X_E$. In particular, $\hat{e}_i^{\beta_i} \cdot (v \cdot x) = v \cdot x$. Hence,

$$w \cdot x = \hat{e}_i^{\alpha_i} \cdot (v \cdot x) = \hat{e}_i^{\bar{\alpha}_i} \cdot (v \cdot x) \in \hat{e}_i^{\bar{\alpha}_i} \cdot \hat{T}_{i+1} \cdot x$$

where $0 \leq \bar{\alpha}_i < \beta_i$. As $\hat{e}_i^{\bar{\alpha}_i} \cdot \hat{T}_{i+1} \subseteq \hat{T}_i$, we have $w \cdot x \in \hat{T}_i \cdot x$. \square

Define the constant

$$C_{\text{Box}} = 180 \cdot m^3 \cdot C_d .$$

Proposition 4.27. *Let $t \geq t_E$ be an integer. Let X be an \mathbb{F}_m -set, and $x \in X$ such that*

$$\text{Box}_{\mathbb{F}_d}(C_{\text{Box}} \cdot t) \cdot \hat{T} \cdot x \subseteq X_E . \quad (4.24)$$

Then, there is a t -tile (x, f) with

$$|\text{Eq}(f)| \geq \left(1 - \frac{d}{t}\right) \cdot |\text{Im}f| \quad (4.25)$$

and for every $t_E \leq \tilde{t} \leq t$,

$$|\eta_{\tilde{t}}(\text{Im}f)| \leq \left(1 + 2C_d \cdot \frac{\tilde{t}}{t}\right)^d \cdot |\text{Im}f| . \quad (4.26)$$

Furthermore,

$$|\eta_t(\text{Im}f)| \leq 2^d \cdot |\text{Im}f| . \quad (4.27)$$

Proof. By Assumption (4.24) and Lemma 4.26, we have

$$\text{Box}_{\mathbb{F}_m}(C_{\text{Box}} \cdot t) \cdot x \subseteq X_E . \quad (4.28)$$

Define

$$H = \langle \pi(\text{Stab}_{\mathbb{F}_m}(x) \cap B_{\mathbb{F}_m}(12 \cdot C_d \cdot t + 1)) \rangle , \text{ and} \\ (U_A, u_A) = (\mathbb{Z}^m / H, 0 + H) .$$

Note that $\{\hat{e}_{d+1}^{\beta_{d+1}}, \dots, \hat{e}_m^{\beta_m}\} \subseteq \text{Stab}_{\mathbb{F}_m}(x) \cap B_{\mathbb{F}_m}(12 \cdot C_d \cdot t + 1)$, since $x \in X_E$ and $12 \cdot C_d \cdot t + 1 \geq C_d \geq \beta_E$. Thus, $K \leq H$, so U_A is a Γ -set. By virtue of Inclusion (4.28), Proposition 4.21 (Tool A) yields a subgraph isomorphism

$$f_A : B_{U_A}(u_A, 6C_d \cdot t) \rightarrow B_X(x, 6C_d \cdot t) .$$

By applying Proposition 4.13 (Tool B) to H and $2t$, it follows that u_A admits the $2t$ -parallelopete $P_{2t}^{\mathcal{D}}$ where $\mathcal{D} = \mathcal{D}_{u_A, 2t}$ and $K \leq \langle \mathcal{D} \rangle$. Hence, for

$$(U_B, u_B) = (\mathbb{Z}^m / \langle \mathcal{D} \rangle, 0 + \langle \mathcal{D} \rangle) ,$$

we have a subgraph isomorphism $f_B : P_{2t}^{\mathcal{D}} \cdot u_B \rightarrow P_{2t}^{\mathcal{D}} \cdot u_A$. Since $P_{2t}^{\mathcal{D}} \cdot u_A \subseteq B_{U_A}(u_A, C_d \cdot 2t)$ (due to Equation (4.22)), we can define $f_{AB} : P_{2t}^{\mathcal{D}} \cdot u_B \rightarrow \hat{P}_{2t}^{\mathcal{D}} \cdot x$ by $f_{AB} = f_A \circ f_B$, and f_{AB} is a subgraph isomorphism since both f_A and f_B are. In particular, x admits a $2t$ -parallelopete. By virtue of the subgraph isomorphism f_A , the balls $B_{U_A}(u_A, C_d \cdot 2t)$ and $B_X(x, C_d \cdot 2t)$ are isomorphic, so Lemma 4.23 implies that $\mathcal{D}_{u_A, 2t} = \mathcal{D} = \mathcal{D}_{x, 2t}$.

We now apply Proposition 4.11 (Tool C) to $\langle \mathcal{D} \rangle$ with the basis \mathcal{D} , and to the parameter t . This yields a finite Γ -set Y and a bijection f_C from Y onto $P_t^{\mathcal{D}} \cdot u_B$, with

$$|\text{Eq}(f_C)| \geq \left(1 - \frac{m - |\mathcal{D}|}{t}\right) \cdot |Y| \geq \left(1 - \frac{d}{t}\right) \cdot |Y| .$$

Note that the restriction of f_{AB} to $\text{Im} f_C = P_t^{\mathcal{D}} \cdot u_B$ is a subgraph isomorphism onto $\hat{P}_t^{\mathcal{D}} \cdot x$. Thus, $f = f_{AB} \circ f_C$ is a bijection from Y onto $\hat{P}_t^{\mathcal{D}} \cdot x$, so (x, f) is a t -tile. Also, $\text{Eq}(f) = |\text{Eq}(f_C)| \geq \left(1 - \frac{d}{t}\right) \cdot |\text{Im} f|$, proving Equation (4.25). We turn to prove Equations (4.26) and (4.27).

Let $t_E \leq \tilde{t} \leq t$, and consider a \tilde{t} -tile (\tilde{x}, \tilde{f}) in X . Denote $\tilde{\mathcal{D}} = \mathcal{D}_{\tilde{x}, 2\tilde{t}}$. Assume that $\text{Im} \tilde{f} \cap \text{Im} f \neq \emptyset$, i.e., $\hat{P}_{\tilde{t}}^{\tilde{\mathcal{D}}} \cdot \tilde{x} \cap \hat{P}_t^{\mathcal{D}} \cdot x \neq \emptyset$. Then, $\tilde{x} \in W \cdot x$, where

$$W = \left(\hat{P}_{\tilde{t}}^{\tilde{\mathcal{D}}}\right)^{-1} \cdot \hat{P}_t^{\mathcal{D}} .$$

In other words, $\eta_{\tilde{t}}(\text{Im} \tilde{f}) \subseteq W \cdot x$. Since $W \subseteq \left(\hat{P}_{2\tilde{t}}^{\tilde{\mathcal{D}}}\right)^{-1} \cdot \hat{P}_{2t}^{\mathcal{D}}$, Equation (4.22) implies that

$$W \subseteq B_{\mathbb{F}_m}(2C_d \cdot \tilde{t} + 2C_d \cdot t) \subseteq B_{\mathbb{F}_m}(4C_d \cdot t) .$$

Hence, we may consider the restriction of $f_A : B_{U_A}(u_A, 6C_d \cdot t) \rightarrow B_X(x, 6C_d \cdot t)$ to $W \cdot u_A$, which is a subgraph isomorphism $W \cdot u_A \rightarrow W \cdot x$. Since $|\eta_{\tilde{t}}(\text{Im}f)| \leq |W \cdot x|$, it follows that

$$|\eta_{\tilde{t}}(\text{Im}f)| \leq |W \cdot u_A| .$$

Now,

$$\begin{aligned} |W \cdot u_A| &= \left| \left(P_{\tilde{t}}^{\tilde{\mathcal{D}}} \right)^{-1} \cdot P_{\tilde{t}}^{\mathcal{D}} \cdot u_A \right| \leq \left| \left(P_{2\tilde{t}}^{\tilde{\mathcal{D}}} \right)^{-1} \cdot P_{\tilde{t}}^{\mathcal{D}} \cdot u_A \right| \\ &\leq |B_{U_A}(P_{\tilde{t}}^{\mathcal{D}} \cdot u_A, 2C_d \cdot \tilde{t})| = |B_{U_B}(P_{\tilde{t}}^{\mathcal{D}} \cdot u_B, 2C_d \cdot \tilde{t})| \\ &\leq \left(1 + \frac{2C_d \cdot \tilde{t}}{t} \right)^d \cdot |\text{Im}(f_C)| , \end{aligned}$$

where the second inequality follows from Equation (4.22), and the third from Proposition 4.11(ii). Equation (4.26) follows since $|\text{Im}(f_C)| = |\text{Im}f|$.

We turn to prove Equation (4.27). Observe that Equation (4.27) is a tighter version of (4.26) in the special case $\tilde{t} = t$. Hence, we continue with the existing notation, and assume further that $\tilde{t} = t$. Note that, since Γ is abelian, all transitive Γ -sets are Cayley graphs of quotients of Γ , and so balls of the same radius in the Γ -set U_A are isomorphic.

We have

$$B_X(\tilde{x}, 2C_d \cdot t) \subseteq B_X(W \cdot x, 2C_d \cdot t) \subseteq B_X(x, 6C_d \cdot t) = \text{Im}f_A .$$

Thus, by virtue of the subgraph isomorphism f_A , the balls of radius $2C_d \cdot t$ centered at x and at \tilde{x} are both isomorphic, as subgraphs of X , to balls of the same radius in the Γ -set U_A . Hence, these two balls are also subgraph isomorphic to each other. Consequently, $\mathcal{D} = \tilde{\mathcal{D}}$ due to Lemma 4.23, and so $W = \left(\hat{P}_t^{\mathcal{D}} \right)^{-1} \cdot \hat{P}_t^{\mathcal{D}}$.

Now, Lemma 4.10(iii) implies that $W \cdot u_B = \left(\hat{P}_t^{\mathcal{D}} \right)^{-1} \cdot \hat{P}_t^{\mathcal{D}} \cdot u_B \subseteq \hat{P}_{2t}^{\mathcal{D}} \cdot u_B$. Hence, the subgraph isomorphism $f_{AB} : P_{2t}^{\mathcal{D}} \cdot u_B \rightarrow \hat{P}_{2t}^{\mathcal{D}} \cdot x$ restricts to a subgraph isomorphism $W \cdot u_B \rightarrow W \cdot x$. Thus,

$$|\eta_t(\text{Im}f)| \leq |W \cdot x| = |W \cdot u_B| \leq \left| \hat{P}_{2t}^{\mathcal{D}} \cdot u_B \right| \leq |P_{2t}^{\mathcal{D}}| \leq 2^{m-|\mathcal{D}|} \cdot |P_t^{\mathcal{D}}| ,$$

where the last inequality follows from Lemma 4.10(ii). As $|\mathcal{D}| \geq m - d$ and $|\text{Im}f| = |P_t^{\mathcal{D}}|$, Equation (4.27) follows. \square

We turn to discuss an iteration of our algorithm. We require the following observation.

Lemma 4.28. *Let $\mathcal{C} = (A_i)_{i \in I}$ be a finite collection of finite sets. Let $c > 0$, and assume that for each $i \in I$, at most $c \cdot |A_i|$ sets A_j ($j \in I$) intersect A_i (including A_i itself). Say that $J \subseteq I$ is intersection-free if $A_{j_1} \cap A_{j_2} = \emptyset$ for all distinct $j_1, j_2 \in J$. Then, I has an intersection-free subset J such that $\left| \bigcup_{j \in J} A_j \right| \geq \frac{|I|}{c}$.*

Proof. Let J be a maximal intersection-free subset of I , and let $M = \bigcup_{j \in J} A_j$. Note that M intersects at most $c \cdot |M|$ sets out of $(A_i)_{i \in I}$. By maximality, each of the $|I|$ elements of \mathcal{C} intersects M , and so $|I| \leq c \cdot |M|$. \square

In the following proposition, which describes a single iteration of the tiling algorithm, we think of the set A as the image of the tiles already injected into X in previous iterations.

Proposition 4.29. *Let X be a finite \mathbb{F}_m -set, $A \subseteq X$ and $t_E \leq t \in \mathbb{N}$. Then, there is a finite Γ -set Y and an injection $f : Y \rightarrow X \setminus A$, with the following properties:*

$$(i) \quad |\text{Im}f| \geq \frac{1}{2^d} \cdot \left(|X| - (3C_{\text{Box}} \cdot t)^d \cdot |\text{Tor}(\Gamma)| \cdot |X \setminus X_E| - |\eta_t(A)| \right).$$

$$(ii) \quad |\text{Eq}f| \geq \left(1 - \frac{d}{t}\right) \cdot |\text{Im}f|.$$

$$(iii) \quad |\eta_{\tilde{t}}(\text{Im}f)| \leq \left(1 + 2C_d \cdot \frac{\tilde{t}}{t}\right)^d \cdot |\text{Im}f| \text{ for every } t_E \leq \tilde{t} \leq t.$$

Proof. Let

$$M = \left(\text{Box}_{\mathbb{F}_d}(C_{\text{Box}} \cdot t) \cdot \hat{T} \right)^{-1} \cdot (X \setminus X_E)$$

and $\tilde{X} = X \setminus (M \cup \eta_t(A))$. Note that $\text{Box}_{\mathbb{F}_d}(C_{\text{Box}} \cdot t) \cdot \hat{T} \cdot \tilde{X} \subseteq X_E$, so Proposition 4.27, applied to each $x \in \tilde{X}$ separately, yields a set of t -tiles $(x, f_x)_{x \in \tilde{X}}$.

Let $x \in \tilde{X}$. By Equation (4.27), $\eta_t(\text{Im}f_x) \leq 2^d \cdot |\text{Im}f_x|$, so $\text{Im}f_x$ intersects at most $2^d \cdot |\text{Im}f_x|$ of the sets $\text{Im}f_{x'}$ ($x' \in \tilde{X}$). In other words, the collection $(\text{Im}f_x)_{x \in \tilde{X}}$ satisfies the requirements of Lemma 4.28, with $c = 2^d$. Hence, there is a subset $J \subseteq \tilde{X}$ such that the sets $(\text{Im}f_x)_{x \in J}$ are pairwise disjoint, and their union is of size at least $2^{-d} \cdot |\tilde{X}|$. We create a Γ -set $Y = \coprod_{x \in J} \text{domain}(f_x)$ and an injection $f = \coprod_{x \in J} f_x : Y \rightarrow X$. By the definition of \tilde{X} , we have $\tilde{X} \cap \eta_t(A) = \emptyset$. This means that if (x, f_x) is a t -tile and $x \in \tilde{X}$, then $\text{Im}f_x \subseteq X \setminus A$. By the definition of f , it follows that $\text{Im}f \subseteq X \setminus A$. We proceed to prove that f has the stated properties. First,

$$|\text{Im}f| \geq 2^{-d} \cdot |\tilde{X}| \geq 2^{-d} (|X| - |M| - |\eta_t(A)|),$$

and so Property (i) follows since

$$\begin{aligned} |M| &\leq \left| \text{Box}_{\mathbb{F}_d}(C_{\text{Box}} \cdot t) \cdot \hat{T} \right| \cdot |X \setminus X_E| = (2C_{\text{Box}} \cdot t + 1)^d \cdot |\text{Tor}(\Gamma)| \cdot |X \setminus X_E| \\ &\leq (3C_{\text{Box}} \cdot t)^d \cdot |\text{Tor}(\Gamma)| \cdot |X \setminus X_E|. \end{aligned}$$

Property (ii) holds since

$$|\text{Eq}f| \geq \sum_{x \in J} |\text{Eq}f_x| \geq \sum_{x \in J} \left(1 - \frac{d}{t}\right) \cdot |\text{Im}f_x| = \left(1 - \frac{d}{t}\right) \cdot |\text{Im}f|,$$

where the second inequality is guaranteed by Equation (4.25).

Property (iii) follows since, for $t_E \leq \tilde{t} \leq t$, Equation (4.26) yields

$$|\eta_{\tilde{t}}(\text{Im}f)| \leq \sum_{x \in J} |\eta_{\tilde{t}}(\text{Im}f_x)| \leq \sum_{x \in J} \left(1 + 2C_d \cdot \frac{\tilde{t}}{t}\right)^d \cdot |\text{Im}f_x| = \left(1 + 2C_d \cdot \frac{\tilde{t}}{t}\right)^d \cdot |\text{Im}f| .$$

□

We turn to describe and analyze the tiling algorithm itself.

Proposition 4.30. *Let X be an \mathbb{F}_m -set, and denote $n = |X|$. Let $\delta > 0$, and assume that $\frac{|X \setminus X_E|}{n} \leq \delta$. Then, there is a Γ -set Y and an injection $f : Y \rightarrow X$, with*

$$|\text{Eq}(f)| \geq \left(1 - C \cdot \delta^{\frac{1}{Q}}\right) \cdot n , \quad (4.29)$$

where $C > 0$ is a constant which depends only on the equation-set E , and $Q \leq O(2^d \cdot d \cdot \max\{d \log d, \log \beta_E, 1\})$ with an absolute implied constant.

Proof. We inductively define a sequence f_1, \dots, f_s of injections into X . The domain of each f_i is a finite Γ -set.

Let $i \geq 1$. Assume that the injections f_1, \dots, f_{i-1} have already been defined. Let $A_i = \bigcup_{j=1}^{i-1} \text{Im}f_j$, and define

$$\begin{aligned} h &= 16d \cdot C_d + 1 \\ H_i &= 24C_{\text{Box}} \cdot |\text{Tor}(\Gamma)|^{\frac{1}{d}} \cdot \delta^{\frac{1}{d}} \cdot n \cdot h^{i-1} \\ b_i &= n - |A_i| \\ t_i &= \frac{b_i}{H_i} . \end{aligned}$$

It is helpful to remember that H_1, H_2, \dots is a geometric sequence. To generate the injection f_i , we apply Proposition 4.29 to the set X , with $t = \lfloor t_i \rfloor$ and $A = A_i$. We continue this process as long as $t_i \geq 2t_E$ (hence, $\lfloor t_i \rfloor \geq t_E$), and denote the obtained injections by f_1, \dots, f_s .

Proposition 4.29 guarantees that $\text{Im}f_i \cap A_i = \emptyset$, so the sets $(\text{Im}f_i)_{i=1}^s$ are pairwise disjoint. Hence, we may take $Y = \coprod_{i=1}^s \text{domain}(f_i)$ and $f = \coprod_{i=1}^s f_i$, yielding an injection from the finite Γ -set Y into X . We turn to prove that f satisfies Equation (4.29).

Let $1 \leq j \leq s$. We seek to control the interference of an injection f_j with subsequent iterations. Let $j < i \leq s$. By Proposition 4.29,

$$|\eta_{\lfloor t_i \rfloor}(\text{Im}f_j) \setminus \text{Im}f_j| \leq \left(\left(1 + \frac{2C_d \cdot \lfloor t_i \rfloor}{\lfloor t_j \rfloor}\right)^d - 1 \right) \cdot |\text{Im}f_j| .$$

Note that, as $t_j \geq t_E \geq 2$, we have $\lfloor t_j \rfloor \geq \frac{t_j}{2}$. Hence,

$$2 \cdot \frac{C_d \cdot \lfloor t_i \rfloor}{\lfloor t_j \rfloor} \leq 4 \cdot \frac{C_d \cdot t_i}{t_j} = 4C_d \cdot \frac{b_i H_j}{b_j H_i} \leq 4C_d \cdot \frac{H_j}{H_i} \leq 4C_d \cdot \frac{1}{h} \leq \frac{1}{4d} .$$

In general, $(1+x)^d \leq 1+2d \cdot x$ for $0 \leq x \leq \frac{1}{2d}$, and so,

$$\begin{aligned} |\eta_{\lfloor t_i \rfloor}(\text{Im}f_j) \setminus \text{Im}f_j| &\leq \frac{4d \cdot C_d \cdot \lfloor t_i \rfloor}{\lfloor t_j \rfloor} \cdot |\text{Im}f_j| \leq \frac{8d \cdot C_d \cdot t_i}{t_j} \cdot |\text{Im}f_j| \\ &\leq \frac{8d \cdot C_d \cdot t_i}{t_j} \cdot b_j = 8d \cdot C_d \cdot t_i \cdot H_j. \end{aligned} \quad (4.30)$$

We next give a lower bound on the number of points covered in the i -th iteration. For $1 \leq i \leq s$, Equation (4.30) yields

$$\begin{aligned} |\eta_{\lfloor t_i \rfloor}(A_i) \setminus A_i| &\leq \sum_{j=1}^{i-1} |\eta_{\lfloor t_i \rfloor}(\text{Im}f_j) \setminus \text{Im}f_j| \leq 8d \cdot C_d \cdot t_i \cdot \sum_{j=1}^{i-1} H_j \\ &= 8d \cdot C_d \cdot t_i \cdot \frac{H_i - H_1}{h-1} \leq \frac{8d \cdot C_d}{h-1} \cdot t_i \cdot H_i \leq \frac{t_i \cdot H_i}{2} = \frac{b_i}{2}. \end{aligned} \quad (4.31)$$

By Proposition 4.29(i),

$$|\text{Im}f_i| \geq 2^{-d} \cdot \left(n - (3C_{\text{Box}} \cdot \lfloor t_i \rfloor)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n - |\eta_{\lfloor t_i \rfloor}(A_i)| \right).$$

Equation (4.31) yields

$$\begin{aligned} |\text{Im}f_i| &\geq 2^{-d} \cdot \left(n - (3C_{\text{Box}} \cdot \lfloor t_i \rfloor)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n - \frac{b_i}{2} - |A_i| \right) \\ &= 2^{-d} \cdot \left(b_i - (3C_{\text{Box}} \cdot \lfloor t_i \rfloor)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n - \frac{b_i}{2} \right) \\ &= 2^{-d} \cdot \left(\frac{b_i}{2} - (3C_{\text{Box}} \cdot \lfloor t_i \rfloor)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n \right) \\ &\geq 2^{-d} \cdot \left(\frac{b_i}{2} - (3C_{\text{Box}} \cdot t_i)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n \right). \end{aligned}$$

Now,

$$\begin{aligned} (3C_{\text{Box}} \cdot t_i)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n &= \left(\frac{3C_{\text{Box}} \cdot b_i}{H_i} \right)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n \\ &\leq \left(\frac{3C_{\text{Box}} \cdot b_i}{H_1} \right)^d \cdot |\text{Tor}(\Gamma)| \cdot \delta n \\ &= \left(\frac{b_i}{8n} \right)^d \cdot n \leq \frac{b_i}{8n} \cdot n = \frac{b_i}{8} \leq \frac{b_i}{4}, \end{aligned}$$

and consequently,

$$|\text{Im}f_i| \geq 2^{-(d+2)} \cdot b_i.$$

Let $\gamma = 1 - 2^{-(d+2)}$. Then, $b_i = b_{i-1} - |\text{Im}f_{i-1}| \leq b_{i-1} \cdot \gamma$ for every $2 \leq i \leq s$, so $b_i \leq n \cdot \gamma^{i-1}$.

Finally, we turn to bound $|\text{Eq}f|$ from below. Proposition 4.29(ii) gives

$$\begin{aligned}
 |\text{Im}f| - |\text{Eq}f| &= \left(\sum_{i=1}^s |\text{Im}f_i| \right) - |\text{Eq}f| \leq \sum_{i=1}^s (|\text{Im}f_i| - |\text{Eq}f_i|) \\
 &\leq \sum_{i=1}^s \frac{d}{[t_i]} \cdot |\text{Im}f_i| \leq \sum_{i=1}^s \frac{2d}{t_i} \cdot |\text{Im}f_i| = \sum_{i=1}^s 2d \cdot \frac{|\text{Im}f_i|}{b_i} \cdot H_i \\
 &\leq \sum_{i=1}^s 2d \cdot H_i \leq 2d \cdot H_{s+1} .
 \end{aligned} \tag{4.32}$$

By the termination condition of the algorithm $t_{s+1} \leq 2t_E$, but $t_{s+1} = \frac{b_{s+1}}{H_{s+1}}$ and $b_{s+1} = n - |\text{Im}f|$. Consequently, $|\text{Im}f| \geq n - 2t_E \cdot H_{s+1}$. Together with Equation (4.32), this implies

$$|\text{Eq}f| = |\text{Im}f| - |\text{Im}f \setminus \text{Eq}f| \geq n - 2(t_E + d) \cdot H_{s+1} . \tag{4.33}$$

In order to bound H_{s+1} , we first need to bound s . Appealing again to the termination condition, we have

$$2t_E \leq t_s = \frac{b_s}{H_s} \leq \frac{n \cdot \gamma^{s-1}}{H_1 \cdot h^{s-1}} ,$$

so

$$s \leq \log_{\gamma/h} \left(\frac{2t_E \cdot H_1}{n} \right) + 1 \leq \frac{\log \left(48t_E \cdot C_{\text{Box}} \cdot |\text{Tor}(\Gamma)|^{\frac{1}{d}} \cdot \delta^{\frac{1}{d}} \right)}{\log \gamma - \log h} + 1 .$$

Hence,

$$h^s \leq C' \cdot \delta^{\frac{\log h}{d(\log \gamma - \log h)}}$$

for some $C' > 0$ which depends only on the equation-set E . Thus,

$$H_{s+1} = 24C_{\text{Box}} \cdot |\text{Tor}(\Gamma)|^{\frac{1}{d}} \cdot n \cdot \delta^{\frac{1}{d}} \cdot h^s \leq C'' \cdot n \cdot \delta^{\frac{1}{d} + \frac{\log h}{d(\log \gamma - \log h)}} = C'' \cdot n \cdot \delta^{\frac{1}{Q}} . \tag{4.34}$$

where $C'' = 24C_{\text{Box}} \cdot |\text{Tor}(\Gamma)|^{\frac{1}{d}} \cdot C'$ and $Q = d \cdot \left(1 - \frac{\log h}{\log \gamma} \right)$. Now,

$$\begin{aligned}
 Q &= d \cdot \left(1 - \frac{\log h}{\log \gamma} \right) \leq d \cdot \left(1 + \frac{\log h}{2^{-(d+2)}} \right) \leq O \left(2^d \cdot d \cdot \log(d \cdot C_d) \right) \\
 &\leq O \left(2^d \cdot d \cdot \max \{ d \log d, \log \beta_E, 1 \} \right) ,
 \end{aligned}$$

where the implied constant is absolute. Therefore, the proposition follows from Equations (4.33) and (4.34). \square

Finally, we turn to proving the main theorem.

Proof of Theorem 1.16. By Lemma 3.15 and Proposition 4.30, for any finite Γ -set X we have

$$G_E(X) \leq |S| \cdot C \delta^{\frac{1}{Q}} ,$$

where C and Q are as in Proposition 4.30. This yields Equation (4.2) from the beginning of Section 4, which is just a more elaborate form of Theorem 1.16. \square

5 A lower bound on the polynomial stability degree of \mathbb{Z}^d

We turn to prove Theorem 1.17. We rely here on the formulation of stability in terms of group actions and labeled graphs, which was introduced in Section 3. The current section is independent of Section 4, except for the notion of a *sorted word* and the corresponding notation \hat{r} for $r \in \mathbb{Z}^d$ (see Section 4.2).

Fix a constant $d \in \mathbb{N}$ and fix generator sets $S = \{e_1, \dots, e_d\}$ and $\hat{S} = \{\hat{e}_1, \dots, \hat{e}_d\}$ for \mathbb{Z}^d and \mathbb{F}_d , respectively. Note that we have a natural homomorphism $\pi: \mathbb{F}_d \rightarrow \mathbb{Z}^d$ which maps \hat{e}_i to e_i . Let E denote the *commutator equation-set* $E_{\text{comm}}^d = \{\hat{e}_i \hat{e}_j \hat{e}_i^{-1} \hat{e}_j^{-1} \mid 1 \leq i < j \leq d\}$ from Equation (1.2) and note that $\mathbb{F}_d / \langle\langle E_{\text{comm}}^d \rangle\rangle \cong \mathbb{Z}^d$. We seek to construct an infinite sequence of \mathbb{F}_d -sets $\{X_t\}_{t=1}^\infty$, with $\lim_{t \rightarrow \infty} L_E(X_t) = 0$ and

$$G_E(X_t) \geq \Omega_{t \rightarrow \infty} \left(L_E(X_t)^{\frac{1}{d}} \right). \quad (5.1)$$

By virtue of Proposition 3.5, Inequality (5.1) implies that $\deg(\text{SR}_E) \geq d$, yielding Theorem 1.17.

Fix a positive integer t . Given $x \in \mathbb{Z}^d$, let $[x]$ denote its coset in the quotient group $\mathbb{Z}^d / (t \cdot \mathbb{Z}^d)$. We build the set X_t by taking the natural action of \mathbb{F}_d on $\mathbb{Z}^d / (t \cdot \mathbb{Z}^d)$, removing a single point, and slightly fixing the result so that it remains an \mathbb{F}_d -set. Concretely, the ground-set for X_t is $(\mathbb{Z}^d / (t \cdot \mathbb{Z}^d)) \setminus \{[0]\}$. Each generator \hat{e}_i acts by taking $[(t-1) \cdot e_i]$ to $[e_i]$, and mapping any other $\left[\sum_{j=1}^d a_j \cdot e_j \right]$ to $\left[\left(\sum_{j=1}^d a_j \cdot e_j \right) + e_i \right]$, as usual.

We wish to compute $L_E(X_t)$. Let $1 \leq i < j \leq d$, and consider the set of points in which the relation $[\hat{e}_i, \hat{e}_j]$ is violated, namely $\{x \in X_t \mid \hat{e}_i \hat{e}_j \hat{e}_i^{-1} \hat{e}_j^{-1} \cdot x \neq x\}$. It is not hard to verify that these are exactly the three points

$$\{[e_i], [e_j], [e_i + e_j]\}.$$

Summing over the $\binom{d}{2}$ elements of E , Definition 3.1 gives

$$L_E(X_t) = \frac{1}{|X_t|} \cdot \binom{d}{2} \cdot 3.$$

Since $|X_t| = t^d - 1$, it follows that $L_E(X_t) \leq O_{t \rightarrow \infty}(t^{-d})$. Hence, the following proposition implies Equation (5.1), which yields Theorem 1.17.

Proposition 5.1. $G_E(X_t) \geq \Omega_{t \rightarrow \infty}(t^{-1})$

Proof. Let Y be a \mathbb{Z}^d -set of size $t^d - 1$ and let $f: Y \rightarrow X_t$ be a bijection. We need to prove that $\|f\|_S \geq \Omega_{t \rightarrow \infty}(\frac{1}{t})$.

Define the set

$$X^0 = X_t \setminus \left(\bigcup_{i=1}^d \{[k \cdot e_i] \mid 0 \leq k \leq t-1\} \right),$$

and let $Y^0 = f^{-1}(X^0)$.

Let U be a connected component of Y , i.e., an \mathbb{F}_d -orbit, and let $u \in U$. Note that the subset $A = \left\{ \sum_{i=1}^d \alpha_i \cdot e_i \mid 0 \leq \alpha_i \leq t-1 \text{ for every } 1 \leq i \leq d \right\}$ of \mathbb{Z}^d does not inject into U at u , since $|A| = t^d > |Y| \geq |U|$. Hence, there are $p \neq q \in A$ such that $p \cdot u = q \cdot u$. Let $0 \neq r = p - q$, so $r \in \text{Stab}_{\mathbb{Z}^d}(u)$. Hence, $r \in \text{Stab}_{\mathbb{Z}^d}(y)$ for every $y \in U$. Write $r = \sum_{i=1}^d \alpha_i \cdot e_i$ where $-(t-1) \leq \alpha_i \leq t-1$ and define $w \in \mathbb{F}_d$ by $w = \hat{r}$ (see Section 4.2). We also write $w = w_1 \cdots w_k$ as a reduced word, with $w_i \in \hat{S}^\pm$. Note that $w \cdot y = y$ for every $y \in U$. However, it is not hard to see that $w \cdot x \neq x$ for each $x \in X^0$. Indeed, a non-trivial sorted word in $\text{Stab}_{\mathbb{F}_d}(x)$ must contain some generator or its inverse at least t times.

For $y \in U \cap Y^0$, let P_y denote the set of edges in the path that starts at y and proceeds as directed by w . Namely, the first such edge is $y \xrightarrow{w_k} w_k \cdot y$, then $w_k \cdot y \xrightarrow{w_{k-1}} w_{k-1} w_k \cdot y$, and so on (if w_i is the inverse of a generator, we walk along an edge backwards). Assume that f preserves all of the edges in P_y . Then $f(w \cdot y) = w \cdot f(y)$, but this is absurd, since $w \cdot y = y$, while $w \cdot f(y) \neq f(y)$, as $f(y) \in X^0$. Hence, P_y must contain a non-preserved edge.

We claim that each edge of Y is contained in at most $t-1$ of the paths $\{P_y\}_{y \in U \cap Y^0}$. Note that for $1 \leq j \leq k$, the map that takes $y \in U \cap Y^0$ to the j -th edge in the path P_y is an injection. Thus, an edge labeled \hat{e}_i ($i \in [d]$) is contained in at most $|\alpha_i| \leq t-1$ of these paths, yielding the claim. Since every such path contains an edge which is not preserved by f , it follows that U contains at least $\frac{|U \cap Y^0|}{t-1}$ non-preserved edges. Summing over all connected components U yields

$$\|f\|_S \geq \frac{1}{|Y|} \sum_U \frac{|U \cap Y^0|}{t-1} = \frac{1}{t^d - 1} \cdot \frac{|Y^0|}{t-1} = \frac{t^d - 1 - d(t-1)}{(t^d - 1) \cdot (t-1)} \geq \Omega_{t \rightarrow \infty} \left(\frac{1}{t} \right).$$

□

6 Discussion

In this work we proved that finitely generated abelian groups are polynomially stable, and bounded their degree of polynomial stability. For the free abelian group \mathbb{Z}^d , our lower and upper bounds on the degree are, respectively, d and an exponential expression in d . It would be interesting to close this gap. We note that the exponential term in our upper bound comes from the 2^d -factor in the right-hand side of Equation (4.27). More precisely, replacing this factor in Equation (4.27) by some smaller term, would yield the same replacement in Equation (4.3).

In particular, when d is small, some of our lemmas can be simplified. For instance, in a lattice of degree ≤ 4 , the vectors yielding the local minima form a basis (see p. 51 of [16], cf. Proposition 4.3). This fact may be helpful in computing the exact degree of polynomial stability for certain constant values of d , starting with $d = 2$.

Another open question that suggests itself is whether polynomial stability holds for a larger class of groups, for example, groups of polynomial growth.

One may also consider a more flexible notion of stability (see Section 4 of [2]): In our definition of the global defect of an \mathbb{F}_m -set X , we do not allow adding points to X . It is also natural to consider a model where adding points is allowed. More precisely, given two finite \mathbb{F}_m -sets X and Y , $|X| \leq |Y|$, we allow making X isomorphic to Y by first adding $|Y| - |X|$ points, and then adding and modifying edges. We set the cost of edge addition and edge modification to $\frac{1}{|X|}$ per edge. This generalizes Definition 3.2.

Note that our proof of Theorem 1.17 does not hold under the above model, since one can transform X_t to a \mathbb{Z}^d -set by augmenting it with a single point, and then changing a constant number of edges. We do have a proof (not included in the present paper) applicable for this model, that the degree of polynomial stability of \mathbb{Z}^d ($d \geq 2$) is at least 2. We do not know of a better bound.

In regard to the applications to property testing (Section 1.3), our observation that an equation-set is polynomially stable if and only if its *canonical* tester is efficient, raises the question of which sets of equations admit *any* efficient tester. The subject of [3] is a similar question in the non-quantitative setting.

Finally, we note that the proof of Proposition 4.30 gives, in fact, a stronger statement: Each orbit of the constructed set Y is of size at most $O(\frac{1}{\delta})$. So, for every finite \mathbb{F}_m -set X and $\delta > \frac{|X \setminus X_E|}{|X|}$, the set X is $O(\delta^{1/Q})$ -close to a Γ -set Y whose orbits are of size at most $O(\frac{1}{\delta})$.

Acknowledgments

We, the authors, would like express our gratitude to Alex Lubotzky for his patient and devoted support, for his encouragement, and for many important suggestions regarding the presentation of the ideas in this work. We are also indebted to Nati Linial for his support and help.

A Stability of certain quotients and finite groups

The first goal of this appendix is to prove Proposition A.3, which relates the stability rate of a group Γ and some of its quotients. This proposition is of independent interest, and also enables a somewhat simplified approach to the proof of our main theorem (Theorem 1.16), as described in the beginning of Section 4.

Lemma A.1. *Let \mathbb{F} be a free group on a finite set S , and $E \subseteq \mathbb{F}$ a finite subset. Let X and Y be finite \mathbb{F} -sets, $|X| = |Y|$. Then,*

$$L_E(Y) \leq L_E(X) + \left(\sum_{w \in E} |w| \right) \cdot d_S(X, Y) . \quad (\text{A.1})$$

Consequently,

$$L_E(Y) \leq \left(\sum_{w \in E} |w| \right) \cdot G_E(Y) . \quad (\text{A.2})$$

Proof. Write $n = |X| = |Y|$. Recalling the notation of Section 3.1, let $\Phi_X, \Phi_Y : S \rightarrow \text{Sym}(n)$ be S -assignments such that $\mathbb{F}(\Phi_X)$ is isomorphic to X , $\mathbb{F}(\Phi_Y)$ is isomorphic to Y , and $d_S(X, Y) = d_n(\Phi_X, \Phi_Y)$. Then, using the triangle inequality and Lemma 2.1, we see that

$$\begin{aligned} L_E(Y) &= L_E(\Phi_Y) \\ &= \sum_{w \in E} d_n(\Phi_Y(w), 1) \\ &\leq \sum_{w \in E} d_n(\Phi_Y(w), \Phi_X(w)) + \sum_{w \in E} d_n(\Phi_X(w), 1) \\ &\leq \sum_{w \in E} |w| \cdot d_n(\Phi_Y, \Phi_X) + L_E(\Phi_X) \\ &= \left(\sum_{w \in E} |w| \right) \cdot d_S(X, Y) + L_E(X) . \end{aligned}$$

We turn to the last assertion (A.2). Assume, for the sake of contradiction, that $L_E(Y) > \left(\sum_{w \in E} |w| \right) \cdot G_E(Y)$. Write $\Gamma = \mathbb{F}/\langle\langle E \rangle\rangle$. Then, there is a Γ -set Z such that $L_E(Y) > \left(\sum_{w \in E} |w| \right) \cdot d_S(Z, Y)$. But $L_E(Z) = 0$, and so, applying Inequality (A.1) with $X = Z$, we get a contradiction. \square

Definition A.2. Let $F_1, F_2 : (0, |E|] \rightarrow [0, \infty)$ be monotone nondecreasing functions. Write $F_1 \lesssim F_2$ if $F_1(\delta) \leq F_2(C\delta) + C\delta$ for some $C > 0$. Recalling the equivalence relation \sim from Definition 1.10, we have $F_1 \sim F_2$ if and only if $F_1 \lesssim F_2$ and $F_2 \lesssim F_1$. The partial order \lesssim on functions enables us to define a partial order on \sim -equivalence classes, namely, $[F_1] \lesssim [F_2]$ if and only if $F_1 \lesssim F_2$.

The following proposition relates the stability rate (see Definition 1.12) of a group Γ to the stability rate of certain quotients Γ/N .

Proposition A.3. *Let Γ be a finitely-presented group. Let N be a normal subgroup of Γ . Assume that N is a finitely-generated group. Then, $\text{SR}_{\Gamma/N} \lesssim C \cdot \text{SR}_{\Gamma}$ for some $C = C(\Gamma, N) > 0$. In particular: (i) if Γ is stable, then so is Γ/N , and (ii) $\text{deg}(\text{SR}_{\Gamma/N}) \leq \text{deg}(\text{SR}_{\Gamma})$.*

Proof. Let $\pi : \mathbb{F} \rightarrow \Gamma$ be a presentation of Γ as a quotient of a finitely-generated free group $\mathbb{F} = \mathbb{F}_S$, $|S| < \infty$. Then, $\pi^{-1}(N)$ is a normal subgroup of \mathbb{F} , $\pi^{-1}(N)/\text{Ker}\pi \cong N$ is finitely-generated and $\mathbb{F}/\pi^{-1}(N) \cong \Gamma/N$. Let $E_1 \subseteq \text{Ker}\pi$ be a finite set which generates $\text{Ker}\pi$ as a normal subgroup of \mathbb{F} . Let $E_2 \subseteq \pi^{-1}(N)$ be a finite set whose image in $\pi^{-1}(N)/\text{Ker}\pi$ generates the group $\pi^{-1}(N)/\text{Ker}\pi$. Then, $E = E_1 \cup E_2$ generates $\pi^{-1}(N)$ as a normal subgroup of \mathbb{F} . So, $\mathbb{F}/\langle\langle E_1 \rangle\rangle \cong \Gamma$ and $\mathbb{F}/\langle\langle E \rangle\rangle \cong \Gamma/N$. Thus, it suffices to show that $\text{SR}_E(\delta) \leq O(\text{SR}_{E_1}(\delta) + \delta)$.

Let X be a finite \mathbb{F} -set. Write $n = |X|$ and $\delta = L_E(X)$. We need to show that $G_E(X) \leq O(\text{SR}_{E_1}(\delta) + \delta)$. Let Z be a Γ -set, $|Z| = |X| = n$, for which $d_S(X, Z) \leq \text{SR}_{E_1}(L_{E_1}(X))$. Since $E_1 \subseteq E$, we have $L_{E_1}(X) \leq L_E(X)$, and so

$$d_S(X, Z) \leq \text{SR}_{E_1}(L_E(X)) = \text{SR}_{E_1}(\delta) .$$

Using Lemma A.1, we deduce

$$\begin{aligned} L_{E_2}(Z) &\leq L_E(Z) \leq L_E(X) + \left(\sum_{w \in E} |w| \right) \cdot d_S(X, Z) \\ &\leq \delta + \left(\sum_{w \in E} |w| \right) \cdot \text{SR}_{E_1}(\delta) . \end{aligned}$$

This allows us to bound the size of Z_{E_2} (see Definition 3.6) from below. Indeed, there are $n \cdot L_{E_2}(Z)$ pairs $(z, w) \in Z \times E$ for which $w \cdot z \neq z$. Clearly, the number of distinct values of z in those pairs is at most $n \cdot L_{E_2}(Z)$, and so

$$|Z_{E_2}| \geq n - n \cdot L_{E_2}(Z) \geq n - n \cdot \left(\delta + \left(\sum_{w \in E} |w| \right) \cdot \text{SR}_{E_1}(\delta) \right) .$$

Now, as $N = \langle \pi(E_2) \rangle$, each $z \in Z_{E_2}$ is fixed by N . In fact, since N is normal in Γ , every $z \in Z_{E_2}$ belongs to an orbit $\Gamma \cdot z \subseteq Z$ where all points are fixed by N . Therefore, Z_{E_2} is a union of Γ -orbits of Z , and the action of Γ on Z_{E_2} factors through Γ/N . Consider the inclusion map $\iota : Z_{E_2} \rightarrow Z$. All points of Z_{E_2} are equivariance points of ι . So, applying Lemma 3.15 to ι , we get

$$\begin{aligned} G_E(Z) &\leq |S| \cdot \left(1 - \frac{1}{n} \cdot |Z_{E_2}| \right) \\ &\leq |S| \cdot \left(\delta + \left(\sum_{w \in E} |w| \right) \cdot \text{SR}_{E_1}(\delta) \right) \\ &\leq O(\text{SR}_{E_1}(\delta) + \delta) . \end{aligned}$$

Hence,

$$\begin{aligned} G_E(X) &\leq G_E(Z) + d_S(X, Z) \\ &\leq O(\text{SR}_{E_1}(\delta) + \delta) + \text{SR}_{E_1}(\delta) \\ &\leq O(\text{SR}_{E_1}(\delta) + \delta) , \end{aligned}$$

as required. \square

Theorem 2 of [9] states that all finite groups are stable. The following is the quantitative version:

Proposition A.4. *Let Γ be a finite group. Then, the degree of polynomial stability $\text{deg}(\text{SR}_\Gamma)$ of Γ is 1. In particular, Γ is stable.*

Proof. Write $\delta \mapsto \delta$ for the inclusion map $(0, |E|] \rightarrow (0, \infty]$, and $[\delta \mapsto \delta]$ for its \sim -class (Definition 1.10). Our task is to show that $\text{SR}_\Gamma = [\delta \mapsto \delta]$. By the definition of the relation \preceq , we have $[\delta \mapsto \delta] \preceq \text{SR}_\Gamma$, and we need to prove the reverse inequality. Let $\pi : \mathbb{F} \rightarrow \Gamma$ be a presentation of Γ as a quotient of a finitely-generated free group \mathbb{F} . Then, $[\mathbb{F} : \text{Ker}\pi] < \infty$, and so $\text{Ker}\pi$ itself is a finitely-generated group. Therefore, by Proposition A.3, for some $C > 0$,

$$\text{SR}_\Gamma \preceq C \cdot \text{SR}_{\mathbb{F}} \sim [\delta \mapsto \delta] .$$

□

The following proposition shows that for every equation-set, except for the trivial equation-sets \emptyset and $\{1\}$, the stability rate is at least linear. This motivates the requirement $k \geq 1$ in Definition 1.13.

Proposition A.5. *Let \mathbb{F} be a free group on a finite set S , and $E \subseteq \mathbb{F}$ a finite subset. Assume that $E \neq \emptyset$ and $E \neq \{1_{\mathbb{F}}\}$. Then,*

$$\text{SR}_E(\delta) \geq C \cdot \delta \tag{A.3}$$

for some $C > 0$.

Proof. Since SR_E is monotone non-decreasing, it suffices to prove Inequality (A.3) for $\delta \in (0, \delta_0] \cap \mathbb{Q}$ for some $\delta_0 > 0$. As E contains a non-trivial element of \mathbb{F} , there is a finite \mathbb{F} -set X such that $\delta_0 \stackrel{\text{def}}{=} L_E(X)$ is positive (and rational). Write $\delta_0 = \frac{m_0}{n_0}$ for integers $0 < m_0 \leq n_0$. Take $\delta \in (0, \delta_0] \cap \mathbb{Q}$, and write $\delta = \frac{p}{q} \cdot \delta_0$, where $0 < p \leq q$ are integers. Define an \mathbb{F} -set Y as the disjoint union of p copies of X and $(q - p) \cdot |X|$ additional fixed points. Then,

$$L_E(Y) = \frac{p \cdot |X| \cdot L_E(X)}{p \cdot |X| + (q - p) \cdot |X|} = \frac{p}{q} \cdot L_E(X) = \delta .$$

Finally, by Inequality (A.2) in Lemma A.1, $G_E(Y) \geq (\sum_{w \in E} |w|)^{-1} \cdot L_E(Y)$, and so $\text{SR}_E(\delta) \geq (\sum_{w \in E} |w|)^{-1} \cdot \delta$, as required. □

References

- [1] Goulnara Arzhantseva and Liviu Păunescu, *Almost commuting permutations are near commuting permutations*, J. Funct. Anal. **269** (2015), no. 3, 745–757.
- [2] Oren Becker and Alexander Lubotzky, *Group stability and Property (T)*, available at <https://arxiv.org/abs/1809.00632>.
- [3] ———, *Testability of permutation equations and group stability* (In preparation).
- [4] Oren Becker, Alexander Lubotzky, and Andreas Thom, *Stability and Invariant Random Subgroups*, available at <https://arxiv.org/abs/1801.08381>.
- [5] Marcus De Chiffre, Lev Glebsky, Alexander Lubotzky, and Andreas Thom, *Stability, cohomology vanishing, and non-approximable groups*, available at <https://arxiv.org/abs/1711.10238>.

- [6] S. Eilers, T. Shulman, and A. P. W. Sørensen, *C*-stability of discrete groups*, available at <https://arxiv.org/abs/1808.06793>.
- [7] Gábor Elek and Endre Szabó, *Sofic representations of amenable groups*, Proc. Amer. Math. Soc. **139** (2011), no. 12, 4285–4291.
- [8] Nikolay Filonov and Ilya Kachkovskiy, *A Hilbert-Schmidt analog of Huaxin Lin's Theorem*, available at <https://arxiv.org/abs/1008.4002>.
- [9] Lev Glebsky and Luis Manuel Rivera, *Almost solutions of equations in permutations*, Taiwanese J. Math. **13** (2009), no. 2A, 493–500.
- [10] Oded Goldreich, *Introduction to property testing*, Cambridge University Press, 2017.
- [11] Don Hadwin and Tatiana Shulman, *Stability of group relations under small Hilbert-Schmidt perturbations*, J. Funct. Anal. **275** (2018), no. 4, 761–792.
- [12] ———, *Tracial stability for C*-algebras*, Integral Equations Operator Theory **90** (2018), no. 1, Art. 1, 35.
- [13] M. B. Hastings, *Making almost commuting matrices commute*, Comm. Math. Phys. **291** (2009), no. 2, 321–345.
- [14] J. C. Lagarias, H. W. Lenstra Jr., and C.-P. Schnorr, *Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice*, Combinatorica **10** (1990), no. 4, 333–348.
- [15] Jeffrey C. Lagarias, *Point lattices*, Handbook of combinatorics, Vol. 1, 2, 1995, pp. 919–966.
- [16] Jacques Martinet, *Perfect lattices in Euclidean spaces*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 327, Springer-Verlag, Berlin, 2003.
- [17] Donald S. Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, J. Analyse Math. **48** (1987), 1–141.
- [18] Dana Ron, *Algorithmic and analysis techniques in property testing*, Found. Trends Theor. Comput. Sci. **5** (2009), no. 2, front matter, 73–205.
- [19] S. M. Ulam, *A collection of mathematical problems*, Interscience Tracts in Pure and Applied Mathematics, no. 8, Interscience Publishers, New York-London, 1960.