Houghton's group H_3 has superpolynomial sofic profile

Seth Douglas Nidhal Mghirbi

November 2026 (version 1.2.0)

Abstract

How many points does a permutation model of a group need? Cornulier's sofic profile measures the size needed to imitate a fixed finite piece at a given accuracy. We show that Houghton's finitely presented elementary amenable group on three rays has a finite piece with superpolynomial profile. The unconditional lower exponent is one divided by one plus the base-two logarithm of 107; the upper bound is exponential in the square root of the accuracy parameter times its logarithm.

The lower bound uses entropy carried by approximately commuting copies of the symmetric group on three letters, with polynomial commutator fillings certified by a doubling self-similarity. A stronger sweep criterion charges a whole ordered suffix at once and rounds permutations in Hamming distance. Conditionally, a quadratic Dehn function, or a quadratic bound for block coherence with at most three letters per position, gives a square-root lower exponent, matching the upper exponent up to a logarithm. We prove an all-scale quadratic lower bound for block coherence; its quadratic upper bound remains open.

Every homomorphism from a Houghton group on at least three rays to a group of polynomial sofic profile kills the finitary alternating subgroup. We locate the error in Cornulier's extension estimate and prove a corrected bound that depends on scale. We also give unitary lower bounds, unitary models smaller than our permutation models, a polynomial Dehn-function upper bound, and a direct proof that every finite piece of a Baumslag-Solitar group has at most linear profile.

Contents

1 Introduction **2**

2 Houghton's group, chunks and profiles **6**
2.1 Images and finite-dimensional representations 7

3 Polynomial area of far commutators **9**

4 Stability and counting **11**
4.1 An entropic sector bound 11
4.2 A phase-robust form 12
4.3 From an approximate presentation to dimension 13

5 Lower bounds **14**

6 Upper bounds **15**
6.1 Permutation models 15
6.2 Unitary models 17

7 The correction **20**

8 Other groups **22**
8.1 A criterion from finitely many relations 22
8.2 Whole suffix sweeps 23
8.3 Block coherence and the conditional square-root bound 25
8.4 Baumslag–Solitar groups 27

9 Open problems **29**

Use of AI **30**

A Fillings for the far-commutator bound **30**

B From far commutators to the full Dehn function **34**

1 Introduction

Sofic groups are the groups whose finite pieces can be modelled by permutations up to a small error, and hyperlinear groups those whose finite pieces can be modelled by unitary matrices. Quantitative versions ask, for groups that can be approximated, how large the models must be. Cornulier introduced the *sofic profile* $\sigma_E(r)$ of a finite piece (a chunk) E of a group: the least n for which E maps to the symmetric group Sym_n with multiplication defect $1/r$ and with distinct elements at normalized Hamming distance at least $1 - 1/r$ [8]. Amenable groups are sofic. The first non-sofic group, the unit group of the binary Leavitt algebra, was announced by OpenAI on 1 August 2026 [33], building on work of Kun and of Kun and Thom; further examples are given in [16, 26], and [16] recalls this history. Every non-sofic group has a chunk with infinite profile. Between bounded and linear profiles there is a gap: Cornulier showed that the profile of a chunk is either bounded or at least r for every $r > 1$ [8, Fact 3.5], and Lemma 2.1 sharpens r to $2r$. He asked for methods to bound the sofic profile of explicit groups from below, whether some group has sofic profile that is unbounded and not linear in r, and whether such a group can be sofic [8, Problem 3.18].

Houghton's group H_3 [24] is finitely presented [5], with an explicit presentation due to Johnson [25], and elementary amenable: it is an extension of the locally finite group $\mathrm{FSym}(\mathbb{N})$ of finitary permutations by $\mathbb{Z}^2$. Houghton's groups H_m, $m \geq 3$, defined in the same way with m rays, are infinite isolated groups [9, Section 5.3], one of the classes Cornulier proposed as test cases for his Problem 3.18 [8, Example 3.19]. Section 2 defines H_3 and a chunk E_H. Here and below logarithms are to base two, and $\log 107 \approx 6.7415$.

Theorem 1.1. *The chunk E_H of H_3 has sofic profile*

$$2^{r^{1/(\log 107+1)}/1280-1/16} \;\leq\; \sigma_{E_H}(r) \;\leq\; 2^{O(\sqrt{r}\log r)} \qquad (r \geq 2).$$

The unweighted bound obtained from certified finite fillings already gives $\sigma_{E_H}(r) \geq 2^{r^{1/12}/270-1/16}$.

The displayed lower bound exceeds the trivial value one only for $r > 80^{\log 107+1} \approx 5.40\cdot 10^{14}$; its significance is asymptotic, not a prediction of a practical crossover.

Corollary 1.2. *The elementary amenable group H_3, an extension $1 \to \mathrm{FSym}(\mathbb{N}) \to H_3 \to \mathbb{Z}^2 \to 1$ of a locally finite group by $\mathbb{Z}^2$, has a chunk E_H whose sofic profile is not bounded by any polynomial.*

This contradicts [8, Theorem 3.15], which for this extension would give $\sigma_{E_H}(r) = O(r^2)$ (Section 7), and the conclusion drawn from it in [8, Example 3.16], that every elementary amenable group has polynomial sofic profile. The contradiction needs only superpolynomial

growth, so the unweighted bound of Theorem 1.1, whose fillings the original-relator certificate archive checks without the auxiliary optimisation behind the exponent $1/(\log 107+1)$, already suffices.

Corollary 1.3 (Homomorphisms to polynomial-profile groups)**.** *Let H_m denote the group of permutations of m rays that are eventually translations on each ray. Suppose every finite chunk of Γ has polynomial sofic profile. Every homomorphism $H_m \to \Gamma$, $m \geq 3$, kills the finitary alternating subgroup. Its image has commutator subgroup of order at most two, nilpotency class at most two, and an abelian subgroup of index at most $2^{\lfloor (m-1)/2 \rfloor}$; for H_3 this index is at most two.*

For any field K and absolutely irreducible finite-dimensional variety X over K, every chunk of the birational group $\mathrm{Bir}_K(X)$ has polynomial sofic profile by [8, Corollary 4.5], which does not depend on Cornulier's Theorem 3.15. So the conclusions of Corollary 1.3 hold for $\Gamma = \mathrm{Bir}_K(X)$, and no group containing H_3 embeds into such a birational group. In characteristic zero these birational consequences are known. By Jordan-type theorems for birational groups [34, 3, 23] (see also the references in [13]), no birational group contains every finite group, or even every finite alternating group. But H_3 contains the finitary symmetric group, hence every finite group, and the finitary alternating subgroup of H_m is simple and contains every finite alternating group, so its image in a birational group is trivial. Our argument is independent of these results and applies over every field.

The proof is in Section 2.1. As is well known, every nontrivial normal subgroup of H_m contains the finitary alternating subgroup [9, proof of Proposition 5.5]; Lemma 2.4 gives a proof. Thus every homomorphic image is either a faithful copy of H_m, with superpolynomial sofic profile, or a residually finite, virtually abelian quotient with bounded profile on each finite chunk (Proposition 2.5).

We locate the error in Section 7: the extension theorem asserts a finite piece of the kernel that does not depend on the error parameter, but the construction it relies on, that of Elek and Szabó [14], needs kernel elements conjugated along a Følner set whose size grows with the inverse error. Tracking that dependence gives a correct, scale-dependent extension bound (Proposition 7.1), the chunk-indexed analogue of [2, Theorem 5.5]. Since H_3 is amenable, and hence sofic, Theorem 1.1 also answers the last question of Cornulier's Problem 3.18: a sofic group can have sofic profile that is unbounded and not linear in r. Slofstra constructed a finitely presented group whose hyperlinear profile, at relator defect ε and separation δ, is at least $C' \exp(C(\delta/\varepsilon)^{\alpha})$ for any fixed $\alpha < 1/2$ [36, Theorem 1.2], and the proof of [36, Corollary 1.3] gives a chunk of it with sofic profile at least $C' \exp(Cr^{\alpha/2})$, with exponent $\alpha/2$ arbitrarily close to $1/4$, larger than ours; this answers the second of these questions. That group is not known to be sofic, or even hyperlinear (as Slofstra notes after his Theorem 1.2), hence not known to be amenable, and its sofic profile may be infinite. To our knowledge H_3 is the first amenable group shown to have superpolynomial sofic profile.

The lower bound holds for unitary approximations as well. Write $P = \langle a, b, \alpha \mid r_1, \ldots, r_5 \rangle$ for Johnson's presentation (Section 2) and call unitaries $\phi(a), \phi(b), \phi(\alpha) \in U(D)$ an *e-representation* if every relator satisfies $\|\phi(r_i) - I\|_2 \leq e$ in the normalized Hilbert–Schmidt norm.

Theorem 1.4. *Let $\beta_H = 1/(2\log 107+1) \approx 1/14.48$, $K' = 750 \cdot 135 \cdot 3^{\log 107}$ and $e > 0$. Every e-representation of P on $\mathbb{C}^D$ in which the three-cycle $\gamma = \alpha\, a\alpha a^{-1}$ satisfies $\|\phi(\gamma) - I\|_2 \geq 1/2$ has $\log D \geq (K'e)^{-2\beta_H}/16 - 1/16$.*

Conversely, spreading the single defect of our permutation models over many cells gives such representations with $\log D = O(e^{-2/3})$, while our permutation models give only $O(e^{-1})$ (Corollary 6.7). The displayed unitary lower bound is positive only for $e < 1/K' \approx 6.00 \cdot 10^{-9}$.

The same exponent holds when $1/2$ is replaced by any separation margin $\Delta \in (0, 1]$, with explicit constants (Proposition 4.6).

A general criterion. The lower bound needs only finitely many relations and many cheap copies of S_3. For a finite set S and a finite set R of words in $S^{\pm 1}$, let $\mathrm{Area}_R(w)$ be the least number of conjugates of words of $R^{\pm 1}$ whose product is w in the free group on S.

Theorem 1.5 (Counting criterion)**.** *Let G be a group, $S \subset G$ a finite set containing u, v, and R a finite set of words in $S^{\pm 1}$, each representing 1 in G, with $u^2, v^2, (uv)^3 \in R$ and $uv \neq 1$ in G; let l be the largest length of a word in R. There is a chunk E of G, determined by S and R, with the following property. If $w_1, \dots, w_N$ are words in $S^{\pm 1}$ and $\Lambda \geq 1$ is a number with $\mathrm{Area}_R([w_i x w_i^{-1}, w_j y w_j^{-1}]) \leq \Lambda$ for all $i \neq j$ and $x, y \in \{u, v\}$, then*

$$\log \sigma_E(r) \;\geq\; N/16 \qquad \text{for every } r \geq 10^8 N\Lambda(2l-1).$$

Theorem 8.1 proves this together with a unitary version. The chunk does not depend on the words w_i, which may be long; only the areas matter. The scheme, many cheaply transported copies of a finite group, relations controlled by area, and a dimension bound from finite-group stability, is that of Slofstra [36, Sections 2 and 3]. Not in his argument are the entropy bound for approximately commuting copies of S_3, which needs only pairwise commutator data and replaces the stability theorems, a direct treatment of permutation models, in which relations are transported in measure by the standard additivity of relator defects in normalized Hamming distance, and the use of the scheme inside an amenable group. For H_3, Sections 4 and 5 count the copies $u = \alpha$, $v = a\alpha a^{-1}$ conjugated by $w_i = a^{3(i-1)}$. The criterion can never give more than $2^{O(r)}$. For the copies transported by the powers of a along a ray of H_3 the area grows at least quadratically (Proposition 3.4), so the pairwise estimate cannot give more than exponent $1/3$. The sweep strengthening of Theorem 8.3 supersedes that estimate: a quadratic Dehn function or quadratic block coherence would give $\exp(c\sqrt{r}) \leq \sigma_{E_H}(r) \leq \exp(C\sqrt{r}\log r)$ (Proposition 8.4). In a companion paper [11] we apply the criterion to groups that come close to this ceiling, among them an elementary amenable group, built from Houghton groups acting on finite binary configurations, with a chunk whose profile is at least $\exp(r^{1-o(1)})$. Houghton groups embed in Thompson's group V (Röver [35, Proposition 2.6]; Genevois extends this to generalized Houghton groups [17, Corollary 5.11]), so V has a chunk whose profile is finite and superpolynomial.

Cornulier's Proposition 3.17 asserts that the sofic profile of a Baumslag–Solitar group is at most linear: linear when the group is not residually finite, and bounded otherwise [8]. The linear lower bound and the bounded cases are his, from his Fact 3.8 and Meskin's theorem, but his proof of the linear upper bound used Theorem 3.15. We prove the upper bound directly, and raise the lower bound from r to $2r$.

Theorem 1.6 (Baumslag–Solitar groups; Cornulier's Proposition 3.17 with a direct proof)**.** *Let $m, n \neq 0$. Every chunk E of $BS(m, n) = \langle a, t \mid ta^m t^{-1} = a^n \rangle$ has $\sigma_E(r) = O_E(r)$. If $BS(m, n)$ is not residually finite, then some chunk has $\sigma_E(r) \geq 2r$ for every $r > 1$, which sharpens Cornulier's lower bound r.*

By Lemma 2.1, which sharpens Cornulier's Fact 3.5 from r to $2r$, no chunk has unbounded profile below $2r$, so these finitely presented groups, which are not locally embeddable into finite groups, have the smallest possible unbounded profile (Section 8.4).

How the proofs go.

- *Far commutators have polynomial area* (Section 3). The commutator of two conjugates of the basic transposition at distance R has area at most $135R^{\log 107}$ in Johnson's presentation, by a doubling self-similarity of H_3 whose finite relations are certified by machine. Lee reduced the Dehn function of H_3 to such areas and bounded them exponentially [27, Section 4.2]; fed into a buffered form of his collection scheme, our estimates give a polynomial bound on the full Dehn function (Theorem 3.3), improving his exponential upper bound. (Thompson's group V, which contains H_3, has polynomial Dehn function [20].)
- *Displaced copies of S_3 need dimension* (Section 4). An approximate representation gives, through the far commutators, many approximately commuting copies of S_3 whose three-cycles are displaced from the identity. Rounding each copy to an exact one and counting the entropy needed to make a state invariant forces dimension exponential in their number.
- *Transport* (Section 5). For permutations, relator errors add up in measure over a van Kampen diagram; for unitaries, in norm. This converts the area bound into the two exponents.
- *Upper bounds* (Section 6). Permutation models on a torus of clock positions that fail at a single corner, and unitary models that spread that failure over many cells by small rotations.
- *Baumslag–Solitar groups* (Section 8.4). The cycles of a^m and of a^n are cut into tiles of equal length, and t carries one family of tiles onto the other while recording the tile in a fixed finite group, so that reduced words move every point.

Related work. Arzhantseva and Cherix [2] introduce and study profiles of metric approximations of finitely generated groups. Their sofic profile uses one parameter both for the radius of the ball to be approximated and for the inverse accuracy, whereas Cornulier's fixes a chunk and varies only the accuracy [2, Section 4.5]; here the name always means Cornulier's profile of a chunk. The hyperlinear profile and its relation to nonlocal games are studied in [37, 36], and hyperlinear approximations of amenable groups come from sofic ones [7]. Our method follows Slofstra's in combining stability of approximate representations of finite groups, here an entropy bound in place of the stability theorems of Gowers–Hatami [19] and De Chiffre–Ozawa–Thom [10], with an isoperimetric input. Lee proved an exponential isoperimetric inequality for Houghton's groups [27], and our polynomial Dehn bound follows his scheme (Section 3). Houghton's groups are not co-Hopfian [6]; we use the doubling endomorphism behind this. The extension theorem is attributed in [8] to the construction of Elek and Szabó [14]. A companion paper [12] applies the lower bounds to the memory needed for the repeated use of a quantum channel built from H_3, and a third paper [11] pushes the criterion close to its ceiling.

Glebe [18, Corollary 4.5] proves that H_m, $m \geq 3$, is not stable in the unnormalized Schatten p-norms, $1 < p \leq \infty$, by a cup-product obstruction. That qualitative nonstability result differs from the normalized-error dimension profiles considered here: we require separation of a fixed finite piece and bound the dimension of every approximate model. Glebe's obstruction uses unnormalized norms, in which, for $p < \infty$, a defect of fixed normalized size grows with the dimension; our unitary bounds use the normalized Hilbert–Schmidt norm.

Organization. Section 2 defines H_3, the chunk and the profiles. Section 3 proves the area bounds, Section 4 the dimension bounds for approximate representations, and Section 5 the lower bounds of Theorems 1.1 and 1.4. Section 6 gives the upper bounds, Section 7 the correction, Section 8 the general criterion and Baumslag–Solitar groups, and Section 9 open problems.

Changes from version 1.0. This version adds the counting criterion (Theorems 1.5 and 8.1), the Baumslag–Solitar theorem (Section 8.4), the linear floor (Lemma 2.1), the bound for every separation margin (Proposition 4.6), the quadratic lower bound for the Dehn function and the example G_0 (Section 7). Corollary 1.5 of version 1.0 is now Remark 2.2 together with the unnumbered remark after Proposition 2.3, which records the standard dimensions of the irreducible representations, and several statements are renumbered. Appendix A gives the fillings and Appendix B the Dehn-function bound.

2 Houghton's group, chunks and profiles

Let $Y = \{(i, j) : i \in \{1, 2, 3\},\ j \geq 1\}$ be three rays. Permutations act on the right and gh means g followed by h. Let a send $(1, 1) \mapsto (2, 1)$, $(1, j) \mapsto (1, j - 1)$ for $j \geq 2$ and $(2, j) \mapsto (2, j + 1)$, fixing ray 3; let b be defined in the same way with rays 2 and 3 exchanged; let α be the transposition of $(1, 1)$ and $(1, 2)$. These generate H_3, the group of permutations of Y that are eventually translations on each ray. They satisfy

$$r_1 = \alpha^2, \quad r_2 = (\alpha a \alpha a^{-1})^3, \quad r_3 = [\alpha, a^2 \alpha a^{-2}], \quad r_4 = aba^{-1}b^{-1}\alpha^{-1}, \quad r_5 = a\alpha a^{-1} b \alpha^{-1} b^{-1}, \tag{1}$$

with $[x, y] = xyx^{-1}y^{-1}$; each identity can be checked directly on the finitely many points where the word moves anything. By a theorem of Johnson [25] (see also [27, Theorem 2.8]) these relators present H_3. The lower bound below uses only that $r_1, \dots, r_5$ hold in H_3, the far-commutator areas of Section 3, and that the three-cycle $\gamma = \alpha\,(a\alpha a^{-1})$ on $(1, 1), (1, 2), (1, 3)$ is nontrivial. The short exact sequence $1 \to \mathrm{FSym}(\mathbb{N}) \to H_3 \to \mathbb{Z}^2 \to 1$ shows that H_3 is elementary amenable.

We recall Cornulier's definitions [8, Definitions 3.2 and 3.4]. A *chunk* of a group G is a finite subset $E \ni 1$ with the partial multiplication $xy = z$ whenever $x, y, z \in E$ and $xy = z$ in G. For a map $f : E \to \mathrm{Sym}_n$ with $f(1) = \mathrm{id}$, let d_n be the normalized Hamming distance on Sym_n. The map is an *ε-morphism* if $d_n(f(xy), f(x)f(y)) \leq \varepsilon$ whenever xy is defined in E, and it is *$(1 - \varepsilon)$-expansive* if $d_n(f(x), f(y)) \geq 1 - \varepsilon$ for distinct $x, y \in E$. The sofic profile is

$$\sigma_E(r) = \min\{n :\ \text{there is a } (1 - r^{-1})\text{-expansive } r^{-1}\text{-morphism } E \to \mathrm{Sym}_n\}, \qquad r > 1.$$

A group has polynomial sofic profile if every chunk has $\sigma_E \preceq$ some polynomial [8, Definition 3.10].

Let $E_H \subset H_3$ be the symmetric chunk consisting of 1, $a^{\pm 1}, b^{\pm 1}, \alpha^{\pm 1}$, all prefixes of the relators $r_1, \dots, r_5$ and of the word $\alpha a \alpha a^{-1}$ for γ, and the inverses of these elements.

For a matrix $T \in M_D$ write $\tau_D(T) = D^{-1} \operatorname{Tr} T$, $\|T\|_2^2 = \tau_D(T^*T)$ and $\|T\|_{HS}^2 = \operatorname{Tr}(T^*T)$; the dimension subscript on τ is omitted when clear. For a density matrix ρ, $S(\rho) = -\operatorname{Tr}(\rho \log \rho)$, and $h_2(p) = -p \log p - (1 - p) \log(1 - p)$, with $0 \log 0 = 0$.

For permutations, relations can be transported in measure: the normalized Hamming distance d_n is bi-invariant, so a null word that is a product of A conjugates of relators has $d_n(\phi(w), \mathrm{id}) \leq A\delta$ when every relator has $d_n(\phi(r_i), \mathrm{id}) \leq \delta$. For unitaries only $\|\phi(w) - I\|_2 \leq A \max_i \|\phi(r_i) - I\|_2$ is available in general, since $\|(e^{i\vartheta} I)^A - I\|_2$ grows like $A\vartheta$.

The two scales are related as follows. For permutation matrices, $\|P_\sigma - P_\pi\|_2^2 = 2\, d_n(\sigma, \pi)$. So an r^{-1}-morphism, read through permutation matrices, has defect at most $(2/r)^{1/2}$ in $\|\cdot\|_2$, and $(1 - r^{-1})$-expansiveness becomes $\|P_{f(x)} - P_{f(y)}\|_2^2 \geq 2(1 - r^{-1})$. Conversely, a unitary statement at squared defect $\delta = e^2$ corresponds to Hamming accuracy $\delta/2$ when the unitaries are permutation matrices. The permutation bounds (Theorem 1.1, Section 5) are stated in d_n; the unitary bounds (Theorem 1.4, Section 6.2) are stated in $\|\cdot\|_2$.

A chunk E *embeds in a finite group* Q if there is an injective map $E \to Q$ that sends 1 to 1 and preserves every product defined in E. A group is locally embeddable into finite groups

(LEF) if all its chunks do [38]. The next lemma sharpens [8, Fact 3.5] from r to $2r$: a profile cannot be unbounded and small.

Lemma 2.1 (A linear lower bound)**.** *If $\sigma_E(r) < 2r$ for some $r > 1$, then E embeds in a finite group. Hence σ_E is either bounded or at least $2r$ for every $r > 1$.*

Proof. Distinct permutations of n points differ on at least two points, so their normalized Hamming distance is at least $2/n$. If $n < 2r$, an r^{-1}-morphism $E \to \mathrm{Sym}_n$ therefore preserves every product defined in E, and $(1 - r^{-1})$-expansiveness makes it injective. Conversely, the regular representation of a finite group containing an embedded copy of E is an exact, fully separating model at every accuracy, so the profile is bounded. □

2.1 Images and finite-dimensional representations

Proof of Corollary 1.3. Let $f : H_3 \to \Gamma$ and suppose $f(\gamma) \neq 1$, where $\gamma = \alpha a \alpha a^{-1}$ is the three-cycle used in the lower bound. The image chunk $f(E_H)$ retains every multiplication relation among the relator prefixes, even if other labels become identified. An expansive r^{-1}-morphism of this chunk to Sym_D therefore gives generator permutations with relator errors at most $23/r$ and displacement of γ at least $1 - 5/r$. For $r \geq 10$, the proof of Theorem 1.1, using Proposition 5.1, applies unchanged: it only needs these relations and the displacement of this one three-cycle. Thus $\log D \geq r^{1/(\log 107+1)}/1280 - 1/16$, contradicting the polynomial profile of $f(E_H)$. Hence f kills γ, and then all finitary three-cycles by conjugation. These generate the finitary alternating subgroup.

For H_m, restrict to the copy of H_3 on the first three rays and conjugate by finitary permutations of all m rays. The same conclusion follows. The quotient by the finitary alternating subgroup is a central extension of $\mathbb{Z}^{m-1}$ by C_2: its kernel is the parity of a finitary permutation, which conjugation preserves. Its commutators are therefore central of order at most two. Put $d = m - 1$. Commutation defines an alternating bilinear form on $\mathbb{F}_2^d$ with values in C_2: it is independent of the lifts, and $[x^2, y] = [x, y]^2 = 1$. Such a form has an isotropic subspace of dimension at least $\lceil d/2 \rceil$. Indeed, if the form is nonzero, split off a nondegenerate plane spanned by e, f with pairing one, apply induction to its orthogonal complement, and adjoin e; if it vanishes, take the whole space. The inverse image of this isotropic subspace in the central extension is abelian and has index at most $2^{\lfloor d/2 \rfloor}$. Its image under any quotient map has no larger index. The other stated properties also pass to every image.

Finally, the independent Corollary 4.5 of [8] gives polynomial profile for every chunk of $\mathrm{Bir}_K(X)$; replace X by a nonempty affine open if necessary. This proves the birational assertions made after Corollary 1.3; for the last of them, note that an injective homomorphism from H_3 does not kill the finitary alternating subgroup. □

Remark 2.2 (Finite-dimensional representations). Every homomorphism $H_m \to U(D)$, $m \geq 3$, kills the finitary alternating subgroup A_{fin}. This is the standard fact that the infinite alternating group has no nontrivial finite-dimensional unitary representations: A_{fin} is simple (Lemma 2.4) and contains elementary abelian 2-subgroups of every rank, while those of $U(D)$ have order at most 2^D. So every finite-dimensional representation of H_m factors through the parity quotient of Proposition 2.3.

Proposition 2.3 (Parity quotient and sharp image bounds)**.** *Put $d = m - 1$ and $k = \lfloor d/2 \rfloor$. The parity quotient has presentation*

$$H_m/A_{\mathrm{fin}} = \langle t_1, \dots, t_d, z \mid z^2 = 1,\ z \text{ central},\ [t_i, t_j] = z\ (i < j) \rangle.$$

The abelian-index bound in Corollary 1.3 is attained by a finite image.

Proof. Indeed, take t_i to translate from ray 1 to ray $i+1$. Directly in the ray action, every $[t_i, t_j]$ is the same basic transposition α. These translations generate H_m: they generate the translation lattice, and conjugates of α give adjacent transpositions along the connected union of rays, hence every finitary permutation. Their classes and $z = \alpha A_{\mathrm{fin}}$ satisfy the displayed relations. Every abstract word collects to $z^{\epsilon} t_1^{n_1} \cdots t_d^{n_d}$. If its image is trivial, translation forces all $n_i = 0$, and parity forces $\epsilon = 0$. Thus there are no further relations.

For attainment, use $X = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $Y = \left(\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right)$ and $Z = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. On $(\mathbb{C}^2)^{\otimes k}$ take

$$\Gamma_{2j-1} = Z^{\otimes(j-1)} \otimes X \otimes I^{\otimes(k-j)}, \qquad \Gamma_{2j} = Z^{\otimes(j-1)} \otimes Y \otimes I^{\otimes(k-j)}.$$

If $d = 2k+1$, append $\Gamma_{2k+1} = Z^{\otimes k}$. These pairwise anticommute and square to I, so $t_i \mapsto \Gamma_i$, $z \mapsto -I$ defines a representation. The first pair generates all matrices on the first qubit; removing its Z-strings from the next pair and iterating proves irreducibility. Its image F lies in the finite Pauli group. Any abelian subgroup $B \leq F$ has a common eigenvector, whose coset orbit and irreducibility give $[F : B] \geq 2^k$. Corollary 1.3 gives the reverse bound, because a finite group has bounded chunk profiles. Thus its abelian-index bound is sharp already for a finite image. □

Remark (Irreducible representations; standard). The finite-dimensional irreducible unitary representations of H_m, which factor through the parity quotient by Remark 2.2, have dimensions 1 and 2^k, and both occur. This is the standard representation theory of complex Clifford algebras; we include the short argument. Let π be an irreducible unitary representation of the parity quotient. The proof of Corollary 1.3 gives an abelian subgroup of index $q \leq 2^k$; the commuting unitaries representing it have a common eigenvector, and its images under q coset representatives span an invariant subspace, so $\dim \pi \leq 2^k$. Schur's lemma gives $z = \pm I$. For $z = I$ the generators commute, so the dimension is one. For $z = -I$, their matrices T_i pairwise anticommute, and T_i^2 is central. Rescale each T_i by a scalar phase so that $T_i^2 = I$. The operators $H_j = iT_{2j-1}T_{2j}$, $1 \leq j \leq k$, are commuting Hermitian involutions. From a common eigenvector, products of the T_{2j-1} reach all 2^k joint eigenvalue patterns, since each flips exactly the jth sign. These vectors are nonzero and orthogonal. The dimension is at least 2^k, hence exactly 2^k. The trivial representation and the representation constructed in the proof of Proposition 2.3 show that both dimensions occur.

Lemma 2.4 (Normal subgroups; [9, proof of Proposition 5.5])**.** *For $m \geq 3$, every nontrivial normal subgroup of H_m contains A_{fin}.*

Proof. First, A_{fin} is simple. If $N \lhd A_{\mathrm{fin}}$ is nontrivial, choose $1 \neq x \in N$, a moved point p, and distinct points r, s outside the support of x. Put $q = px^{-1} \neq p$ and $c = (p\ r\ s)$. With our right-action convention,

$$[x, c] = (q\ r\ s)(p\ s\ r) = (q\ p\ s).$$

Thus N contains a three-cycle. All three-cycles are conjugate in A_{fin}: any finitary conjugator can be made even by composing with a transposition on two unused points. Since three-cycles generate A_{fin}, we have $N = A_{\mathrm{fin}}$.

Now let $K \lhd H_m$. Both K and A_{fin} are normal, so $[K, A_{\mathrm{fin}}] \subseteq K \cap A_{\mathrm{fin}}$. If the intersection is nontrivial, simplicity gives the conclusion. Otherwise K centralizes A_{fin}. Its centralizer in the full symmetric group on the rays is trivial: if a permutation x moves p, set $q = px^{-1} \neq p$ and choose distinct $r, s \notin \{p, q\}$. The conjugate $x(p\ r\ s)x^{-1}$ has q in its support, whereas $(p\ r\ s)$ does not, so x cannot centralize that three-cycle. Hence $K = 1$ in this case. □

Proposition 2.5 (Classification of polynomial-profile images)**.** *Let $f : H_m \to \Gamma$, $m \geq 3$, and put $F = f(H_m)$. The following are equivalent: every finite chunk of F has polynomial sofic profile; every finite chunk of F has bounded sofic profile; F is residually finite; $f(A_{\mathrm{fin}}) = 1$;*

and f is not injective. Every such image is virtually abelian, with the sharp index bound of Proposition 2.3. Moreover, A_{fin} is the intersection of the kernels of all finite quotients of H_m, and also of all finite-dimensional unitary representations (a standard fact, since A_{fin} is infinite and simple).

Proof. Lemma 2.4 gives the equivalence of noninjectivity and $f(A_{\mathrm{fin}}) = 1$. The polynomial-profile condition implies $f(A_{\mathrm{fin}}) = 1$ by Corollary 1.3. If f kills A_{fin}, write its image using the generators of Proposition 2.3, and let $C = \langle f(z), f(t_1)^2, \dots, f(t_{m-1})^2 \rangle$. This is a finitely generated central abelian subgroup of F of index at most 2^{m-1}, since F/C is generated by $m-1$ commuting involutions; in the parity quotient itself the index is exactly 2^{m-1}, four for H_3. An element outside C is detected by F/C. A nonidentity element $g \in C$ is outside kC for some positive integer k, by the structure theorem for finitely generated abelian groups. The subgroup kC is central and F/kC is finite, so it detects g. Thus F is residually finite. For a fixed finite chunk, residual finiteness separates its finitely many pairwise differences in a finite product of quotients. The regular action of that quotient is an exact, fully separating permutation model at every accuracy, hence the profile is bounded. Bounded implies polynomial. Finally every finite quotient kills A_{fin}, while the parity quotient is residually finite by the argument just given. Finite regular representations separate the same quotient; Remark 2.2 supplies the reverse inclusion for unitary kernels. □

3 Polynomial area of far commutators

Let $P = \langle a, b, \alpha \mid r_1, \dots, r_5 \rangle$, put $\sigma_j = a^j \alpha a^{-j}$, and let $A^*(M)$ be the largest area in P of a commutator $[\sigma_p^{\pm 1}, \sigma_q^{\pm 1}]$ with $0 \le p, q \le M$ and $|p - q| \ge 2$. These commutators are trivial in H_3 because the transpositions have disjoint supports.

Theorem 3.1 (Polynomial far-commutator area)**.** *$A^*(M) \le M^{11}$ for every $M \ge 2$.*

Outline; the full proof is in Appendix A. Let ψ be the endomorphism of the free group on a, b, α with $\psi(a) = a^2$, $\psi(b) = b^2$ and $\psi(\alpha) = [a^2, b^2]$. On H_3 it is induced by the diagonal action on two copies of Y, identified with Y by the bijection $(1, j, e) \mapsto (1, 2j - 1 + e)$ and $(i, j, e) \mapsto (i, 2j - e)$ for $i = 2, 3$ and $e \in \{0, 1\}$. The resulting group endomorphism, not this bijection of sets, is injective and not surjective. It differs from the doubling map of [6] by conjugation by the involution that exchanges consecutive pairs of points on rays 2 and 3; all relations used here are certified for ψ itself. The proof has four steps.

(1) *ψ lifts to P with bounded area multiplier.* Each $\psi(r_i)$ is a null word in P, of area at most $0, 17, 182, 864, 359$ for $i = 4, 1, 5, 3, 2$. Hence $\mathrm{Area}_P(\psi(w)) \le 864\, \mathrm{Area}_P(w)$.

(2) *Two cheap letters.* With $t = b^{-1}a$ and $\rho = a^{-1}b^{-1}ab$, the letter α commutes with t at cost 1 and with ρ at cost at most 5, and $\{t, \rho\}$ is invariant under ψ up to one relation, $b^{-1}\rho b = t\rho t^{-1} \cdot \rho \cdot t\rho t^{-1}$, of area at most 14. The identities $\psi(t) = t\rho t$, $t\rho = b^{-1}tb$ and $\psi(\rho) = t^{-1}\rho^b \rho^{b^2} t\, \rho\, \rho^b$ hold in the free group.

(3) *A doubling recursion.* The words $X_n = b^{-n}a^n$ satisfy $X_{2n} = \psi(X_n)$ and $X_{2n+1} = b^{-1}X_{2n}bt$ freely, and in H_3 act as a reversal of $[1, n]$ followed by translation on the line formed by rays 2 and 3. Writing X_n as a word W_n in t, ρ costs $R(n)$ cells, with $R(2n) \le 864\, R(n) + 84\, \#_\rho(W_n)$ and a comparable odd step. Unrolling gives $R(n) \le 1.35\, n^{\log_2 864}$.

(4) *From X_n to far commutators.* $[\alpha, X_n]$ expands into one conjugate of $[\alpha, t]^{\pm 1}$ or $[\alpha, \rho]^{\pm 1}$ per letter of W_n, so its area is at most $2R(n) + 5|W_n|$; a far commutator $[\sigma_0, \sigma_m]$ reduces to it at linear extra cost. This gives $A^*(m) \le 5.4\, m^{\log_2 864} + 10\, m^{\log_2 323} + 18m \le m^{11}$ for $m \ge 5$, and $A^*(2), A^*(3), A^*(4) \le 1, 21, 47$.

□

Certificates. Every finite relation used in the proof is supplied in machine-checkable form: as an explicit list of cells (z_k, r_{i_k}, s_k) such that the relation equals $\prod_k z_k r_{i_k}^{s_k} z_k^{-1}$ in the free group on a, b, α, which a short independent program verifies by free reduction after recomputing the relation from its definition. The certified relations and their cell counts are: $A(2) = 1$, $A(3) \le 21$, $A(4) \le 47$, $A(5) \le 79$, $A(6) \le 157$; $E(0..3) \le 1, 4, 27, 76$; $U(1..3) \le 1, 5, 32$; $B(2), B(3) \le 9, 12$; the relation of step (2) with 14 cells; $[\alpha, \rho]$ with 5; $\psi(\alpha) = \sigma_1\sigma_2\sigma_0\sigma_1$ with 6; the images $\psi(r_1), \psi(r_2), \psi(r_3), \psi(r_5)$ with $17, 359, 864, 182$; and the rewriting of $\psi(\rho)$ as a word in t and ρ, which uses the relation of step (2) six times, with 84. The recursion of step (3) has also been executed mechanically, transporting certificates through ψ cell by cell: for $n \le 9$ the certificates of $X_n = W_n$ have $0, 0, 14, 84, 392, 995, 2493, 5298, 11878$ cells, each at most the recursion bound, and the resulting far commutators satisfy $A(m) \le 73\,121$ for $m \le 11$, far below m^{11}. Appendix A gives the fillings in words; the ancillary files contain the generator, the verifier and the exported certificates. The exported file holds every finite relation listed above and the recursion for $n \le 5$, which the independent verifier checks; the instances $6 \le n \le 9$ were executed by the generator but are not in the file, and the generator exports them for the verifier on request. The proof uses only the finite relations and the hand induction of Appendix A.

Keeping a derived nine-cell relation atomic during substitution gives a sharper bound.

Theorem 3.2 (Reweighted far-commutator area)**.** *$A^*(M) \le 135\, M^{\log 107}$ for every $M \ge 2$.*

The proof is in Appendix A. The same area estimates also control arbitrary null words.

Theorem 3.3 (Polynomial Dehn function)**.** *Let $\delta_P(n)$ be the maximum area of a null word of length at most n in Johnson's presentation. Then*

$$\delta_P(n) \le 400n^4(1 + A^*(4n + 2)) = O(n^{\log 107+4}) \qquad (n \ge 1).$$

The unweighted estimate of Theorem 3.1 alone gives $\delta_P(n) = O(n^{15})$.

The scheme is Lee's [27, Section 4.2]. He reduced the Dehn function of H_3 to the areas of far commutators and of the words comparing the conjugates of α transported along two rays, by powers of a and of b, and bounded these areas exponentially [27, Lemma 4.7 and Remark 4.8]. Our polynomial bounds for these words (Theorems 3.1 and 3.2 and Lemma A.1), fed into a buffered form of the same collection scheme, give Theorem 3.3. The proof in Appendix B collects a buffered null word into adjacent transpositions, then fills it using a uniform Coxeter insertion algorithm. This full Dehn bound does not improve the profile exponents: the direct far-commutator estimate is sharper for the words used there. The explicit integer weights give exponent $\log 107 \approx 6.7415$. The auxiliary substitution matrix has spectral radius $\rho \approx 106.3385$ and also gives the asymptotic bound $O(M^{\log \rho})$ (Appendix A). Neither is an optimal-area claim.

The Dehn function is at least quadratic, as for every finitely presented group that is not hyperbolic [32] (H_3 contains the infinite torsion group $\mathrm{FSym}(\mathbb{N})$). Explicitly, the word $w_m = [a^m, b^m]^2$ is null in H_3, because $[a^m, b^m]$ is a product of m disjoint transpositions, and it has length $8m$. Send a, b and α to the lattice steps $(1, 0)$, $(0, 1)$ and $(0, 0)$. The signed area enclosed by the path of a null word is additive over products of conjugates and unchanged by free reduction. Among the relators only r_4 encloses nonzero area, namely one unit, while w_m encloses $2m^2$. Hence every filling of w_m uses at least $2m^2$ cells of type r_4, and $\delta_P(n) \ge 2\lfloor n/8 \rfloor^2$. The far commutators themselves also have at least quadratic area, and the cost again sits in r_4.

Proposition 3.4 (Quadratic lower bound). *For odd $k \geq 9$, every filling of $C_k = [\alpha, a^k \alpha a^{-k}]$ uses at least $(k-7)(k-5)/8$ cells of type r_4, whatever the number of cells of the other types. Hence $A^*(R) \geq R^2/16$ for every integer $R \geq 23$.*

Proof. In a permutation model in which r_1, r_2, r_3, r_5 act trivially and r_4 moves at most a fraction δ of the points, a null word filled with A_4 cells of type r_4 moves at most a fraction $A_4\delta$ (transport in measure, Section 2). The corner models of Proposition 6.1 are of this kind with $\delta \leq M^{-2}$, while C_{2M-1} moves a fraction at least $(1-s^{-2})(M-4)(M-3)/(2M^2)$. Hence $A_4 \geq (1-s^{-2})(M-4)(M-3)/2$ for every s, which gives the claim with $k = 2M-1$. For integer $R \geq 23$ apply it to the largest odd $k \leq R$. □

Thus the pairwise estimate for transport by the powers of a along ray 1 cannot give more than the exponent $1/3$ in r for permutation approximations, and $2/5$ in $1/e$ for unitary approximations with relator defect e (Section 9). This is a limit of the pairwise area estimate; the sweep strengthening of Theorem 8.3 can improve it: Theorem 8.1 states the count for any family of copies, and [11] applies it to groups that move many copies pairwise at small cost.

4 Stability and counting

The full S_3^N sector bound below feeds the permutation-profile estimate; the phase-robust version later in the section feeds the unitary estimate and the weighted extraction of [12]. The entropy arguments use monotonicity of the Holevo quantity under partial trace, a consequence of strong subadditivity [28].

4.1 An entropic sector bound

Lemma 4.1 (Projection products). *Let $P_1, \dots, P_m$ be orthogonal projections on a Hilbert space and $\|x\| = 1$. Then $1 - \mathrm{Re}\langle x, P_m \cdots P_1 x\rangle \leq 2\sum_j \|(I-P_j)x\|^2$.*

Proof. Put $x_0 = x$, $x_j = P_j x_{j-1}$ and $r_j = x_{j-1} - x_j \in \mathrm{ran}(I-P_j)$. Then $\sum_j \|r_j\|^2 = 1 - \|x_m\|^2 =: b$, and $a = 1 - \mathrm{Re}\langle x, x_m\rangle = \mathrm{Re}\sum_j \langle (I-P_j)x, r_j\rangle \leq \sqrt{Eb}$ with $E = \sum_j \|(I-P_j)x\|^2$. Since $0 \leq \|x - x_m\|^2 = 2a - b$, we get $a \leq \sqrt{2Ea}$, that is $a \leq 2E$. □

Lemma 4.2 (Entropic sector bound). *Let $u_1, \dots, u_N : S_3 \to U(D)$ be unitary representations, not assumed to commute, such that for a three-cycle c*

$$\|u_i(c) - I\|_2 \geq \Delta \quad (1 \leq i \leq N), \qquad \|u_i(g)u_j(h) - u_j(h)u_i(g)\|_2 \leq \eta \quad (i \neq j,\ g, h \in S_3).$$

Put $t = \sqrt{2N}\,\eta$. If $t \leq \Delta$ and $t^2 \leq 5/6$, then

$$\log D \ \geq\ N\Big[\frac{(\Delta - t)^2}{3} - h_2(t^2) - t^2 \log 5\Big].$$

Proof. Let $G = S_3^N$ and $f(g) = u_1(g_1) \cdots u_N(g_N)$. For h in factor i, $\|f(gh) - f(g)u_i(h)\|_2 = \|S^* X S - X\|_2$ with $X = u_i(h)$ and $S = u_{i+1}(g_{i+1}) \cdots u_N(g_N)$, and $\|S^*XS - X\|_2^2 = 2 - 2\,\mathrm{Re}\,\tau(X^*S^*XS)$. The averages $T_j(Y) = \frac{1}{6}\sum_a u_j(a)^* Y u_j(a)$ are orthogonal projections on M_D with the normalized Hilbert–Schmidt inner product, the average of S^*XS over g is $T_N \cdots T_{i+1}(X)$, and $\|(I-T_j)X\|_2^2 = \frac{1}{12}\sum_a \|[X, u_j(a)]\|_2^2 \leq \eta^2/2$. Lemma 4.1 therefore gives

$$\mathbb{E}_g \|f(gh) - f(g)u_i(h)\|_2^2 \leq 2(N-i)\eta^2 \leq t^2.$$

Let R be the right regular representation of G and $W : \mathbb{C}^D \to \ell^2(G) \otimes \mathbb{C}^D$ the isometry $Wv = |G|^{-1/2}\sum_g |g\rangle \otimes f(g)v$. Then $\|R(h)W - W u_i(h)\|_{HS} \leq \sqrt{D}\,t$ for h in factor i. Put

$P = WW^*$ and $\rho = P/D$, a state of entropy $\log D$. For the three-cycle c_i of factor i and $Z_i = 2I - R(c_i) - R(c_i)^*$,

$$\mathrm{Tr}(\rho Z_i) = D^{-1}\|(R(c_i) - I)W\|_{HS}^2 \geq (\Delta - t)^2.$$

Let $\mathcal{E}_i(X) = \frac{1}{6}\sum_h R(h_i)XR(h_i)^*$, with h_i the element h placed in factor i; these conditional expectations commute. Since $(I - P)W = 0$, $\mathrm{Tr}[(I - P)\mathcal{E}_i(\rho)] \leq t^2$. The state $\mathcal{E}_i(\rho)$ is an average of six states of rank D, one of them ρ, so after pinching by P and $I - P$ its two blocks have rank at most D and $5D$. Pinching does not decrease entropy and $x \mapsto h_2(x) + x\log 5$ increases on $[0, 5/6]$, so $S(\mathcal{E}_i(\rho)) \leq \log D + h_2(t^2) + t^2\log 5$. For a conditional expectation $S(\mathcal{E}(\sigma)) - S(\sigma) = D(\sigma\|\mathcal{E}(\sigma))$; since the $\mathcal{E}_j$ commute, monotonicity of relative entropy under $\mathcal{E}_{i-1}\cdots\mathcal{E}_1$ bounds the entropy increment of $\mathcal{E}_i$ applied after $\mathcal{E}_{i-1}\cdots\mathcal{E}_1$ by $D(\rho\|\mathcal{E}_i\rho)$. Hence $\bar\rho = \mathcal{E}_N\cdots\mathcal{E}_1(\rho)$ satisfies $S(\bar\rho) \leq \log D + N[h_2(t^2) + t^2\log 5]$.

The state $\bar\rho$ commutes with $R(G)$, so $\bar\rho = \bigoplus_\lambda p_\lambda(I_{d_\lambda}/d_\lambda \otimes \sigma_\lambda)$ over the irreducible representations λ of S_3^N, and $S(\bar\rho) \geq \sum_\lambda p_\lambda \log d_\lambda = \sum_\lambda p_\lambda k_\lambda$, where k_λ is the number of factors of λ carrying the two-dimensional representation of S_3. The operator Z_i is central in the group algebra, so $\mathrm{Tr}(\bar\rho Z_i) = \mathrm{Tr}(\rho Z_i)$, and it acts as $3I$ on the sectors whose factor i is two-dimensional and as 0 on the others. Hence $N(\Delta - t)^2 \leq \sum_i \mathrm{Tr}(\bar\rho Z_i) = 3\sum_\lambda p_\lambda k_\lambda \leq 3S(\bar\rho)$. □

Lemma 4.3 (Exact S_3 rounding)**.** *If $A, B \in U(D)$ satisfy $\|A^2 - I\|_2, \|B^2 - I\|_2, \|(AB)^3 - I\|_2 \leq e$, there is a unitary representation u of $S_3 = \langle s, t \mid s^2, t^2, (st)^3\rangle$ on $\mathbb{C}^D$ with $\|u(s) - A\|_2 \leq e$, $\|u(t) - B\|_2 \leq 8e$ and $\|u(st) - AB\|_2 \leq 9e$, and each normal form $1, s, t, st, ts, sts$ evaluated in u is within $10e$ of the same word in A, B.*

Proof. Round the spectra of A and B to the nearer of ± 1. For $|z| = 1$ this moves z by at most $|z^2 - 1|$, so the resulting Hermitian unitaries S, T satisfy $\|S - A\|_2, \|T - B\|_2 \leq e$. Then $C = ST$ satisfies $SCS = C^*$ and $\|C^3 - I\|_2 \leq 7e$. Round the spectrum of C to the nearest cube root of unity, choosing 1 at ties except for the following explicit exception: send -1 to 1, although it is not the nearest root there. At this exceptional point the displacement is $2 = |(-1)^3 - 1|$. This rule commutes with complex conjugation and moves z by at most $|z^3 - 1|$, so C_0 satisfies $C_0^3 = I$, $SC_0S = C_0^*$ and $\|C_0 - C\|_2 \leq 7e$. Then $u(s) = S$ and $u(t) = SC_0$ are involutions with $u(s)u(t) = C_0$, and the bounds follow from the triangle inequality. □

4.2 A phase-robust form

The count of Lemma 4.2 can be run on products of three-cycles alone. It then needs no relation between the transpositions of different copies, so a coherent phase in their commutators does no harm. Write $\kappa(u, v) = \|uv - vu\|_2$ and $G(t) = h_2(\min\{\frac{2}{3}t, \frac{2}{3}\}) + \min\{\frac{2}{3}t, \frac{2}{3}\}$.

Theorem 4.4 (Phase-robust sector bound)**.** *Let $c_i, x_i \in U(D)$, $1 \leq i \leq N$, satisfy $c_i^3 = x_i^2 = I$ and $x_ic_ix_i = c_i^{-1}$, with no condition on the commutators of x_i and x_j. Put $\Delta_i = \|c_i - I\|_2$ and $f(g) = c_1^{g_1}\cdots c_N^{g_N}$ for $g \in \mathbb{Z}_3^N$, let $g^{(i)}$ be g with its i-th coordinate negated, and, for uniform g,*

$$\hat a_i^2 = \mathbb{E}_g\|f(g + e_i) - f(g)c_i\|_2^2, \qquad \hat b_i^2 = \mathbb{E}_g\|x_if(g^{(i)}) - f(g)x_i\|_2^2.$$

Then $\hat a_i^2 \leq \frac{4}{3}\sum_{j>i}\kappa(c_i, c_j)^2$, $\hat b_i^2 \leq \frac{4}{3}\sum_{j\neq i}\kappa(x_i, c_j)^2$, and

$$\log D \;\geq\; \sum_{i=1}^N \Big[\frac{(\Delta_i^2 - \hat a_i^2)_+}{3} - \frac{\hat b_i^2}{2\ln 2} - G(\hat a_i^2)\Big].$$

In particular $\Delta_i \geq 49/100$ and $\hat a_i, \hat b_i \leq 1/100$ for all i give $\log D \geq N/16$.

Proof. The ordered-product errors. For h unitary, $\|(I-E_j)h\|_2^2 = \frac{1}{3}\kappa(h,c_j)^2$ for the average E_j of conjugation by the powers of c_j. Moving c_i through the suffix $c_{i+1}^{g_{i+1}}\cdots c_N^{g_N}$ and averaging over its exponents gives $\hat{a}_i^2 = 2 - 2\,\mathrm{Re}\langle c_i, E_N\cdots E_{i+1}c_i\rangle$, and Lemma 4.1 gives the first bound. For the second, $T_j(Z) = \frac{1}{3}\sum_k c_j^{-k} Z c_j^{s_j k}$ with $s_i = -1$ and $s_j = 1$ otherwise are orthogonal projections, $T_i(x_i) = x_i$ by the local relation, $\|(I-T_j)x_i\|_2^2 = \frac{1}{3}\kappa(x_i,c_j)^2$ for $j \neq i$, and $\hat{b}_i^2 = 2 - 2\,\mathrm{Re}\langle x_i, T_N\cdots T_1 x_i\rangle$.

Embedding. Let $V\xi = 3^{-N/2}\sum_g f(g)\xi\otimes|g\rangle$, let C_i shift the i-th coordinate of $\ell^2(\mathbb{Z}_3^N)$ by -1, let F_i negate it, and put $X_i = x_i\otimes F_i$. The C_i commute, $X_iC_iX_i = C_i^{-1}$, $X_iC_jX_i = C_j$ for $j\neq i$, and $D^{-1/2}\|C_iV - Vc_i\|_{HS} = \hat{a}_i$, $D^{-1/2}\|X_iV - Vx_i\|_{HS} = \hat{b}_i$.

Entropy. Let $\rho = VV^*/D$ and let $\mathcal{T}_i$ average conjugation by the powers of C_i. The twirl moves weight at most $\min\{\frac{2}{3}\hat{a}_i^2, \frac{2}{3}\}$ outside the range of VV^*, into a block of rank at most $2D$, so $S(\mathcal{T}_i\rho) \le \log D + G(\hat{a}_i^2)$; as in Lemma 4.2 the twirls commute, and the increment of each twirl, the Holevo quantity of the conjugates of the current state, can only decrease under the earlier twirls by data processing for the Holevo quantity, so $\bar{\rho} = \mathcal{T}_1\cdots\mathcal{T}_N\rho$ has $S(\bar{\rho}) \le \log D + \sum_i G(\hat{a}_i^2)$.

Sectors. Let p be the joint distribution of the eigenvalue labels $Z_i \in \{1,\omega,\omega^2\}$ of the C_i, the same for ρ and $\bar{\rho}$. Then $S(\bar{\rho}) \ge H(p) \ge \sum_i H(Z_i \mid Z_{-i})$. Let Q_a and P_b be the spectral projections of c_i and C_i. The numbers $r_{ab} = D^{-1}\|P_bVQ_a\|_{HS}^2$ form a joint distribution whose marginals are the spectral distribution of c_i in the normalized trace and that of C_i in ρ. Distinct cube roots of unity are at squared distance 3, so $\hat{a}_i^2 = 3\Pr(a\neq b)$ and $\Delta_i^2 = 3\Pr(a\neq 1)$, whence $s_i := \Pr(Z_i\neq 1) \ge (\Delta_i^2 - \hat{a}_i^2)_+/3$. Next, the normalized vectorizations of V and of $X_iVx_i^*$ are at squared distance $\hat{b}_i^2$, so their real overlap is $1-\hat{b}_i^2/2$. Measuring the joint spectral projections of the C_j gives p from the first and its flip $p^{(i)}$ in coordinate i from the second, because X_i exchanges ω and ω^2 in coordinate i and fixes the other labels; by the Cauchy–Schwarz inequality outcome by outcome, $\sum_z \sqrt{p(z)p(z^{(i)})} \ge 1-\hat{b}_i^2/2$. For $0\le t\le 1$ one has $1-h_2(t) \le (1-2\sqrt{t(1-t)})/\ln 2$: with $y = 2t-1$ the difference of $1-\sqrt{1-y^2}$ and $\ln 2\,(1-h_2(t))$ vanishes with its derivative at $y=0$ and has second derivative $(1-y^2)^{-3/2} - (1-y^2)^{-1} \ge 0$. So for fixed other labels with masses u, w on the two nontrivial values of Z_i, the conditional entropy receives at least $u+w-(u+w-2\sqrt{uw})/\ln 2$, and summing gives $H(Z_i\mid Z_{-i}) \ge s_i - \hat{b}_i^2/(2\ln 2)$. The numbers in the last claim give $\frac{0.49^2-10^{-4}}{3} - \frac{10^{-4}}{2\ln 2} - G(10^{-4}) > \frac{1}{16}$. □

The theorem isolates the row quantities $\hat{a}_i, \hat{b}_i$ that control the dimension estimate used below.

4.3 From an approximate presentation to dimension

Call $\phi : \{a,b,\alpha\} \to U(D)$ an e-representation if $\|\phi(r_i) - I\|_2 \le e$ for $i = 1,\dots,5$, where ϕ is extended to the free group. If w is null in P with area A, then $\|\phi(w) - I\|_2 \le Ae$, because w is freely a product of A conjugates of relators and the norm is unitarily invariant.

Proposition 4.5. *Let ϕ be an e-representation with $\|\phi(\gamma) - I\|_2 \ge 1/2$, and let $N \ge 1$ satisfy $750\sqrt{N}\,A_Ne \le 1$, where $A_N = \max\{1, A^*(3N)\}$. Then $\log D \ge N/16$.*

Proof. Put $A_i = \phi(\sigma_{3(i-1)})$ and $B_i = \phi(\sigma_{3(i-1)+1})$ for $1\le i\le N$; all indices are at most $3N-2$. In H_3 each pair generates a copy of S_3 on $\{(1,3i-2),(1,3i-1),(1,3i)\}$; its involution and braid relations are conjugates of r_1 and r_2, with defect at most e. Lemma 4.3 gives exact local pairs $x_i = u_i(s)$ and $c_i = u_i(st)$ with $\|x_i - A_i\|_2 \le e$ and $\|c_i - A_iB_i\|_2 \le 9e$, and $A_iB_i = \phi(a)^{3(i-1)}\phi(\gamma)\phi(a)^{-3(i-1)}$, so $\Delta_i \ge \frac{1}{2} - 9e$. Pairs with different i have indices differing by at least two, so four and two far-commutator interchanges give $\kappa(c_i,c_j) \le 4A_Ne + 36e \le 40A_Ne$ and $\kappa(x_i,c_j) \le 2A_Ne + 20e \le 22A_Ne$. By Theorem 4.4 and the

concavity of G, $\log D \ge N[(\overline{\Delta^2} - \bar{a})/3 - \bar{b}/(2\ln 2) - G(\bar{a})]$ with $\bar{a} = N^{-1}\sum_i \hat{a}_i^2 \le \frac{2\cdot 40^2}{3\cdot 750^2}$, $\bar{b} = N^{-1}\sum_i \hat{b}_i^2 \le \frac{4\cdot 22^2}{3\cdot 750^2}$ and $\overline{\Delta^2} \ge (\frac{1}{2} - \frac{9}{750})^2$. Exact rational arithmetic, using $\ln 2 > 69/100$ and $h_2(q) \le q\log(1/q) + q/\ln 2$, gives a margin of at least $141043/1397250000$ above $1/16$. Explicitly, put $a = 2\cdot 40^2/(3\cdot 750^2)$, $b = 4\cdot 22^2/(3\cdot 750^2)$ and $q = 2a/3 = 64/50625$. The integer inequality $64^3 2^{29} \ge 50625^3$ gives $\log(1/q) \le 29/3$, and the resulting rational lower margin is

$$\frac{(488/1000)^2 - a}{3} - \frac{b}{2(69/100)} - \left(\frac{32}{3} + \frac{100}{69}\right) q - \frac{1}{16} = \frac{141043}{1397250000} > 0.$$

□

With $A^*(M) \le CM^p$ the condition reads $N^{p+1/2}e = O(1)$, which gives $\log D \ge c\, e^{-2/(2p+1)}$: the exponent is $2/(2p+1)$ in $1/e$, or $1/(2p+1)$ in the inverse squared defect $1/e^2$. For $p = \log 107$ this is the exponent $2\beta_H$ of Theorem 1.4. The margin $1/2$ enters only through $\Delta_i \ge \frac{1}{2} - 9e$, so every margin gives the same exponent.

Proposition 4.6 (Every separation margin)**.** *Let $0 < \Delta \le 1$, and put $K_\Delta = \sqrt{80000\,(1 + \log(1/\Delta))}/\Delta$ and $K'_\Delta = 135\cdot 3^{\log 107} K_\Delta$. Every e-representation of P on $\mathbb{C}^D$ with $\|\phi(\gamma) - I\|_2 \ge \Delta$ satisfies*

$$\log D \;\ge\; \frac{\Delta^2}{12}\Big((K'_\Delta e)^{-2\beta_H} - 1\Big).$$

Proof. Write $K = K_\Delta$ and run the proof of Proposition 4.5 with 750 replaced by K: if $K\sqrt{N}\, A_N e \le 1$, then $e \le 1/K$, $\Delta_i \ge \Delta - 9/K$, and

$$\log D \;\ge\; N\Big[\frac{(\Delta - 9/K)^2 - a}{3} - \frac{b}{2\ln 2} - G(a)\Big], \qquad a = \frac{3200}{3K^2}, \quad b = \frac{1936}{3K^2}.$$

We show that the bracket is at least $\Delta^2/12$. Since $K \ge 36/\Delta$, $(\Delta - 9/K)^2 \ge 9\Delta^2/16$, so it suffices that $a/3 + b/(2\ln 2) + G(a) \le 5\Delta^2/48$. Here $a/3 + b/(2\ln 2) < 821.2/K^2$. With $q = 2a/3 < 711.2/K^2$, the bound $(1-q)\log\frac{1}{1-q} \le q\log e$ gives $G(a) = h_2(q) + q \le q\log(2e/q)$, and $x\log(2e/x)$ increases for $x < 2$. Put $L = \log(1/\Delta)$, so that $K^2 = 80000(1+L)/\Delta^2$. Multiplying by K^2 and using $\log(1+L) \le L + \frac{1}{10}$ for $L \ge 0$ (the maximum of $\log(1+L) - L$ is below 0.087), the left side becomes at most

$$821.2 + 711.2\Big[\log\frac{2e\cdot 80000}{711.2} + \log(1+L) + 2L\Big] \;\le\; 7476 + 2134L \;\le\; \frac{5}{48}\,80000\,(1+L),$$

which is the required inequality. Finally $A_N \le 135(3N)^{\log 107}$ by Theorem 3.2, so the condition holds for $N = \lfloor (K'_\Delta e)^{-2\beta_H} \rfloor$, and $\log D \ge N\Delta^2/12$. □

At $\Delta = 1/2$ Proposition 4.5 gives better constants.

5 Lower bounds

Proposition 5.1. *Let ϕ assign permutations of a finite set Ω to a, b, α, with $d_n(\phi(r_i), \mathrm{id}) \le \delta$ for all i and $d_n(\phi(\gamma), \mathrm{id}) \ge 1/2$, and let $N \ge 1$ satisfy $10^8 N\, A_N \delta \le 1$, where $A_N = \max\{1, A^*(3N)\}$. Then $\log|\Omega| \ge N/16$.*

Proof. Use the pairs $A_i = \phi(\sigma_{3(i-1)})$, $B_i = \phi(\sigma_{3(i-1)+1})$, $1 \le i \le N$, as permutation matrices, with $\|P - I\|_2^2 = 2d_n(P, \mathrm{id})$. Their S_3 relators have $d_n \le \delta$, hence $\|\cdot\|_2$-defect at most $e = \sqrt{2\delta}$, and Lemma 4.3 gives exact representations u_i within $10e$ of the normal forms. Commuting

two normal forms from different pairs takes at most nine far-commutator interchanges, so by transport in measure their commutator has $d_n \le 9A_N\delta$ and $\|\cdot\|_2 \le 3\sqrt{2A_N\delta}$; rounding adds at most $40e$, so $\eta \le 43\sqrt{2A_N\delta}$. The element $u_i(st)$ lies within $9e$ of a conjugate of $\phi(\gamma)$, whose $\|\cdot\|_2$-displacement is at least 1, so $\Delta \ge 1 - 9\sqrt{2\delta} \ge \frac{1}{2}$. Then $t^2 = 2N\eta^2 \le 7396NA_N\delta < 10^{-4}$, and Lemma 4.2 gives $\log|\Omega| \ge N[(49/100)^2/3 - h_2(10^{-4}) - 10^{-4}\log 5] \ge N/16$. □

With $A^*(M) \le CM^p$ the condition reads $N^{p+1}\delta = O(1)$, which gives the sofic exponent $1/(p+1)$.

Proof of Theorem 1.1. Let $f : E_H \to \mathrm{Sym}_n$ be a $(1-r^{-1})$-expansive r^{-1}-morphism with $r \ge 10$, and regard permutations as unitary matrices. For permutation matrices, $\|P-P'\|_2^2 = 2\,d_n(P,P')$, and d_n is bi-invariant. Put $\phi(s) = f(s)$ for $s \in \{a,b,\alpha\}$ and extend ϕ to the free group. Since $s\cdot s^{-1} = 1$ in the chunk, $d_n(f(s)^{-1}, f(s^{-1})) \le r^{-1}$. For a relator $s_1\cdots s_l$ with prefixes p_k, the chunk contains $p_k s_{k+1} = p_{k+1}$ and $p_l = 1$, so $l-1$ merges and at most l inverse replacements give $d_n(\phi(r_i), \mathrm{id}) \le (2l-1)/r \le 23/r$, that is $\|\phi(r_i) - I\|_2 \le \sqrt{46/r}$. In the same way $d_n(\phi(\gamma), f(\gamma)) \le 4/r$, while expansiveness gives $d_n(f(\gamma), \mathrm{id}) \ge 1 - 1/r$ because $\gamma \ne 1$; so $d_n(\phi(\gamma), \mathrm{id}) \ge 1 - 5/r \ge 1/2$. Apply Proposition 5.1 with $\delta = 23/r$ and $A^*(3N) \le 135(3N)^{\log 107}$ (Theorem 3.2): the condition $10^8 N \cdot 135(3N)^{\log 107}\delta \le 1$ holds for $N = \lfloor (r/(23\cdot 10^8\cdot 135\cdot 3^{\log 107}))^{1/(\log 107+1)} \rfloor$. Therefore, whenever $N \ge 1$ (which forces $r \ge 10$),

$$\log \sigma_{E_H}(r) \;\ge\; \frac{N}{16} \;\ge\; c_S r^{1/(\log 107+1)} - \frac{1}{16},$$

with $c_S = (23\cdot 10^8\cdot 135\cdot 3^{\log 107})^{-1/(\log 107+1)}/16 > 1/1280$, and for smaller r the right side is negative. For unitary approximations Proposition 4.5 gives the corresponding bound: by Theorem 3.2, with $K' = 750\cdot 135\cdot 3^{\log 107}$ and $N = \lfloor (K'e)^{-2\beta_H} \rfloor$, every e-representation of P on $\mathbb{C}^D$ with $\|\phi(\gamma) - I\|_2 \ge 1/2$ has $\log D \ge (K'e)^{-2\beta_H}/16 - 1/16$, where $\beta_H = 1/(2\log 107+1)$; this is Theorem 1.4. The upper bound $\log\sigma_{E_H}(r) = O(\sqrt{r}\log r)$ is Proposition 6.2. □

6 Upper bounds

6.1 Permutation models

The upper bounds use permutation models that follow the quotient $\mathbb{Z}^2$ on a torus of clock positions and fail at a single corner. Fix integers $M \ge 2$ and $s \ge 2$, put $\ell = 2M+2$ and $L = 3\ell$, and let $\Omega_{M,s} = (\mathbb{Z}/M\mathbb{Z})^2 \times [s]^L$. For $2 \le r \le L$ let $c_r = (1\ r\ r{-}1\ \cdots\ 2) \in \mathrm{Sym}(L)$ be the reverse cycle and $\tau = (1\ 2)$; products of permutations are read from left to right, and a permutation π of the coordinates acts on colorings by $(x\cdot\pi)(t) = x(t\pi^{-1})$. For $0 \le i,j < M$ put $A_{i,j} = c_{\ell-i-j}$ and, for $j < M-1$, $B_{i,j} = c_{2\ell-j}$. On the top row put $B_{i,M-1} = D_i$, where

$$D_0 = c_{2\ell-M+1}, \qquad D_{i+1} = c_{\ell-i-(M-1)}^{-1}\,\tau\,D_i\,c_{\ell-i} \qquad (0 \le i < M-1). \tag{2}$$

Define permutations of $\Omega_{M,s}$, with clock coordinates modulo M, by

$$(i,j,x)\cdot a = (i+1, j, x\cdot A_{i,j}), \qquad (i,j,x)\cdot b = (i, j+1, x\cdot B_{i,j}), \qquad (i,j,x)\cdot\alpha = (i,j,x\cdot\tau).$$

Proposition 6.1 (Corner models)**.** *In these permutations r_1, r_2, r_3, r_5 hold exactly and r_4 fails only at the clock $(M-1, M-1)$, where it acts as the product $R_M = (1\ M{+}4)\prod_{t=5}^{M+3}(t\ t{+}M)$ of M disjoint transpositions of the colours; so every relator moves at most a fraction M^{-2} of $\Omega_{M,s}$, and γ moves a fraction $1 - s^{-2}$. For $M \ge 5$ the far commutator $C_{2M-1} = [\alpha, a^{2M-1}\alpha a^{1-2M}]$ moves a fraction at least $(1-s^{-2})(M-4)(M-3)/(2M^2)$. A word of length at most t that represents a nonidentity element of H_3 fixes at most a fraction $s^{-1} + 4t/M$ of $\Omega_{M,s}$ when $M > 2t$.*

Proof. Cycle identities. For $r \geq 4$, $c_r \tau c_r^{-1} = (2\ 3)$ and $c_r(2\ 3)c_r^{-1} = (3\ 4)$, and for $3 \leq r < t$

$$c_r c_t c_{r-1}^{-1} c_t^{-1} = \tau; \tag{3}$$

both sides exchange 1 and 2 and fix every other point. By induction on i, the recursion (2) gives $D_i(t) = t - 1$ for $2 \leq t \leq M + 3 - i$: this holds for the reverse cycle D_0, and a point $t \leq M + 2 - i$ passes the four factors of D_{i+1} as $t \mapsto t + 1 \mapsto t + 1 \mapsto t \mapsto t - 1$. Since $M + 3 - i \geq 4$, $D_i^{-1}(1) = 2$ and $D_i^{-1}(2) = 3$, so $D_i \tau D_i^{-1} = (2\ 3)$.

Relators. At every clock the colour permutation of $a\alpha a^{-1}$ is $(2\ 3)$, that of $a^2\alpha a^{-2}$ is $(3\ 4)$, and that of $b\alpha b^{-1}$ is $(2\ 3)$, on the top row by the property of D_i. Hence r_1, r_2, r_3, r_5 act trivially everywhere, and γ acts at every clock as a three-cycle of coordinates, which fixes a fraction s^{-2} of the colorings. At the clock (i, j) the colour permutation of r_4 is $A_{i,j} B_{i+1,j} A_{i,j+1}^{-1} B_{i,j}^{-1} \tau$. For $j < M - 1$ it is trivial by (3) with $r = \ell - i - j$ and $t = 2\ell - j$, also for $i = M - 1$ because $B_{i,j}$ does not depend on i off the top row. For $j = M - 1$ and $i < M - 1$ it is trivial by (2). Only the clock $(M - 1, M - 1)$ remains. There the colour permutation is $c_4 D_0 c_{M+3}^{-1} D_{M-1}^{-1} \tau$. With $E_r = c_r^{-1}\tau = (2\ 3\ \cdots\ r)$ the recursion unrolls to $D_{M-1} = E_5 E_6 \cdots E_{M+3}\, D_0\, c_{2M+2} c_{2M+1} \cdots c_{M+4}$, and $D_0 c_r^{-1} D_0^{-1} = E_{r+1}$ for $r \leq 3M + 4$, so the corner permutation is $c_4\, H\, G^{-1} \tau$ with $H = E_{M+4} \cdots E_{2M+3}$ and $G = E_5 \cdots E_{M+3}$. Following the points, H sends $t \mapsto t + M$ for $2 \leq t \leq M + 3$ and $t \mapsto 2M + 5 - t$ for $M + 4 \leq t \leq 2M + 3$, and G sends $t \mapsto t + M - 1$ for $2 \leq t \leq 4$ and $t \mapsto M + 5 - t$ for $5 \leq t \leq M + 3$, both fixing all other points. Hence $HG^{-1} = (2\ 3\ 4\ \ M{+}4) \prod_{t=5}^{M+3} (t\ \ t{+}M)$, and $c_4 H G^{-1} \tau = R_M$.

Far commutators. Let $k = 2M - 1$, fix a clock (i, j) with $i + j \geq M + 3$ (so $i \geq 4$), and put $F = \ell - j$. The colour permutation of a^k read from this clock is $P = \prod_{t=0}^{k-1} c_{F-((i+t) \bmod M)}$, and that of σ_k is $P\tau P^{-1}$, which moves $P^{-1}(1)$ and $P^{-1}(2)$. With $d_r = c_r^{-1} = (1\ 2\ \cdots\ r)$,

$$x \cdot (d_r d_{r+1} \cdots d_t) = \begin{cases} x + t - r + 1, & x < r, \\ t - x + 1, & r \leq x \leq t, \\ x, & x > t. \end{cases}$$

The factors of P^{-1} come in three blocks, with cycle lengths $F - i + 2, \ldots, F$, then $F - M + 1, \ldots, F$, then $F - M + 1, \ldots, F - i$. The first block sends $(1, 2)$ to $(i, i + 1)$; since $i \geq F - M + 1$, the second sends $(i, i + 1)$ to $(F - i + 1, F - i)$; the third fixes $F - i + 1$ and sends $F - i$ to 1. So σ_k acts as the transposition $(1\ \ 2M + 3 - i - j)$, whose second point is at least 5, and C_k acts as a three-cycle. There are $(M - 4)(M - 3)/2$ such clocks.

Short words. Let $\mathcal{S}_{i,j}$ consist of the first $\ell - i - j$, $\ell + i$ and $\ell + j$ points of rays $1, 2, 3$, ordered ray by ray; it has L points. When the clock step does not wrap, the generator a of H_3 maps $\mathcal{S}_{i,j}$ onto $\mathcal{S}_{i+1,j}$ and induces $c_{\ell-i-j}$ on the ordered coordinates; likewise b induces $c_{2\ell-j}$ and α induces τ. So on the clocks at distance at least t from the lines where the clock wraps, a fraction at least $1 - 4t/M$ when $M > 2t$, a word of length at most t acts through the element of H_3 it represents. If that element has nonzero translation it fixes no state. If it is nontrivial with zero translation, it moves only points among the first $t + 2$ of each ray, which lie in $\mathcal{S}_{i,j}$, so it acts as a nonidentity permutation of the coordinates and fixes at most a fraction s^{-1} of the colorings. Its fixed-point fraction is therefore at most $s^{-1} + 4t/M$. □

Proposition 6.2 (Upper bound for every fixed chunk)**.** *For every finite chunk $E \subset H_3$, $\log \sigma_E(r) = O_E(\sqrt{r} \log r)$. In particular this holds for E_H.*

Proof. Fix representative words for the elements of E, with the empty word representing 1, and let t bound the lengths of their pairwise differences. The finitely many chunk relations have some maximum area B_E in Johnson's presentation. Take $s = M$ and the diagonal action

on two independent copies of $\Omega_{M,M}$. For $M > 4t$, the short-word estimate in Proposition 6.1 bounds the fixed-point fraction of every nonidentity difference by $((1+8t)/M)^2$ in the product: the fractions square. Each relator has defect at most $2/M^2$ in the product, so transport in measure bounds every chunk-relation defect by $2B_E/M^2$. Choosing M at least a sufficiently large E-dependent multiple of $\sqrt{r}$ gives an r^{-1}-morphism that is $(1 - r^{-1})$-expansive. Its degree is $|\Omega_{M,M}|^2 = M^{4+2(6M+6)}$, whose logarithm is $O(M \log M)$, proving the claim. □

6.2 Unitary models

In a permutation model a relator that fails costs the fraction of points it moves. In a unitary model it costs a sum of squared angles, so a single defect can be spread thin. Doing this for the corner models gives unitary models whose relator defect has squared norm of order M^{-3}, not M^{-2}, at the same size. Throughout, products are read from left to right, the left factor acting first, as in Section 6, and $e_1 = (1, 0)$, $e_2 = (0, 1)$.

One-particle models. No generator moves the last $2M+2$ of the $6M+6$ colour coordinates. Use the reduced colour coordinates $1, \dots, L$ with $L = 4M+4$, and read each link permutation A_c, B_c and τ of Section 6 as a permutation matrix on $\mathbb{C}^L$. A one-particle model replaces some of these link matrices by unitaries. For a word w and a clock c the holonomy $W_c(w) \in U(L)$ is the product of the link matrices along the path of w from c, with inverse matrices for inverse letters. For a two-element set $S = \{x, y\}$ let Q_S be the projector onto $(e_x - e_y)/\sqrt{2}$; then $\exp(i\pi Q_S)$ is the transposition $(x\ y)$, and $\exp(i\vartheta Q_S)$ is a phase rotation in the plane of S.

Fermionic quantization. For $W \in U(L)$ let $\Lambda(W)$ act on the exterior algebra $\Lambda(\mathbb{C}^L)$, of dimension 2^L, by $v_1 \wedge \cdots \wedge v_m \mapsto Wv_1 \wedge \cdots \wedge Wv_m$. It is multiplicative on all of $U(L)$, and $\operatorname{Tr}\Lambda(W) = \det(I + W)$, since in degree m its eigenvalues are the products of m eigenvalues of W. For $k \geq 1$ put $\Lambda_k = \Lambda^{\otimes k}$, so that $2^{-kL} \operatorname{Tr}\Lambda_k(W) = \det\big((I + W)/2\big)^k$. The quantized model acts on $\mathcal{H} = \mathbb{C}^{M^2} \otimes \Lambda(\mathbb{C}^L)^{\otimes k}$, with normalized trace $\tau_{\mathcal{H}}$, distinct from the transposition $\tau = (1\ 2)$. A letter moves the clock c to $c + \pi(s)$ and applies Λ_k of its link matrix at c. These right-action operators compose in reverse order; taking their adjoints defines the usual free-group representation ρ on a, b, α. A word whose translation is nonzero modulo M has $\tau_{\mathcal{H}}(\rho(w)) = 0$, and a word that closes has

$$\|\rho(w) - I\|_2^2 = \mathbb{E}_c\Big[2 - 2\operatorname{Re}\det\big((I + W_c(w))/2\big)^k\Big],$$

with $\mathbb{E}_c$ the average over clocks. For a three-cycle σ of coordinates $\det((I + \sigma)/2) = \frac{1}{4}$.

Lemma 6.3 (Rotating the links)**.** *In a corner model, replace every link A_c by U_cA_c and every link B_c by V_cB_c, where U_c commutes with $(2\ 3)$ and $(3\ 4)$, and V_c commutes with $(2\ 3)$. Then at every clock the holonomies of α, $a\alpha a^{-1}$, $a^2\alpha a^{-2}$ and $b\alpha b^{-1}$ are $(1\ 2)$, $(2\ 3)$, $(3\ 4)$ and $(2\ 3)$, so r_1, r_2, r_3, r_5 have identity holonomy everywhere and γ acts as a three-cycle.*

Proof. By the proof of Proposition 6.1, $A_c\tau A_c^{-1} = (2\ 3)$, $A_c(2\ 3)A_c^{-1} = (3\ 4)$ and $B_c\tau B_c^{-1} = (2\ 3)$ at every clock. So the holonomy of $a\alpha a^{-1}$ at c is $U_c(2\ 3)U_c^{-1} = (2\ 3)$, that of $a^2\alpha a^{-2}$ is $U_cA_c\,U_{c+e_1}(2\ 3)U_{c+e_1}^{-1}A_c^{-1}U_c^{-1} = U_c(3\ 4)U_c^{-1} = (3\ 4)$, and that of $b\alpha b^{-1}$ is $V_c(2\ 3)V_c^{-1} = (2\ 3)$. The relators r_1, r_2, r_3, r_5 are products of these conjugates, as in the proof of Proposition 6.1. □

A plaquette q with base v_0 has the vertices v_0, $v_1 = v_0 + e_1$, $v_2 = v_0 + e_2$, $v_0 + e_1 + e_2$; its bottom and top edges are the a-links at v_0 and v_2, and its left and right edges the b-links at v_0 and v_1. Its r_4 holonomy is $W_q = A_{v_0}B_{v_1}A_{v_2}^{-1}B_{v_0}^{-1}\tau$. A *region* consists of a set R of plaquettes, a source $q_0 \in R$, a sign $s_R = \pm 1$, and a two-element label $\operatorname{lab}(v) \subset \{1, \dots, L\}$ at every vertex of every plaquette of R. An edge is *interior* if it lies on two plaquettes of R.

Lemma 6.4 (Spreading a defect)**.** *Take a permutation model satisfying the local identities $A_c\tau A_c^{-1} = (2\ 3)$, $A_c(2\ 3)A_c^{-1} = (3\ 4)$ and $B_c\tau B_c^{-1} = (2\ 3)$ of Lemma 6.3 at every clock. Suppose W_q is the identity except at finitely many sources, where it is the product of the transpositions given by the base labels of the regions with that source. Let finitely many regions satisfy:*

(R1) for each region R, every edge of a plaquette $q \in R$ other than its own source $q_0(R)$, from v to w, carries $\mathrm{lab}_R(v)$ onto $\mathrm{lab}_R(w)$ under its link permutation, even if q is another region's source;

(R2) every base label avoids $\{1, 2\}$, and the label at the initial vertex of every interior edge avoids $\{2, 3, 4\}$;

(R3) each region is connected through its interior edges, and the only interior edge of its source is the bottom edge;

(R4) labels of different regions are disjoint wherever both are defined; on an interior link carrying a region's rotation, its label is disjoint from any additional modification of that link.

Here the link permutations belong to the base model, including any preliminary exact moves. Potentials vanish on boundary links; no disjointness from an additional modification is required for a region on a link where that region carries no rotation. Let $F_{R,q}$ be real numbers with $\sum_{q\in R} F_{R,q} = s_R\pi$ for each region. Then there are potentials φ_R on the interior edges of R such that the model with $U_c = \prod_R \exp(i\varphi_R(a\text{-edge at } c)\, Q_{\mathrm{lab}_R(c)})$ and $V_c = \prod_R \exp(i\varphi_R(b\text{-edge at } c)\, Q_{\mathrm{lab}_R(c)})$ still has r_1, r_2, r_3, r_5 exact, and its r_4 holonomy at every plaquette q is $\prod_{R\ni q} \exp(iF_{R,q}Q_{\mathrm{lab}_R(v_0)})$, with v_0 the base of q.

Proof. Put $\varphi_R = 0$ off the interior edges and $(\mathrm{curl}\,\varphi)(q) = \varphi(\text{bottom}) + \varphi(\text{right}) - \varphi(\text{top}) - \varphi(\text{left})$. An interior edge enters the curls of its two plaquettes with opposite signs, so the curls over R sum to zero. Choose a spanning tree of the plaquettes of R joined through interior edges, rooted at the source, and fix the potential of each tree edge, from the leaves inward, so that $\mathrm{curl}\,\varphi_R = F_R - s_R\pi\,\delta_{q_0}$ at the plaquette below it; since both sides sum to zero, the equation then also holds at the root. By (R2) and (R4) all rotations on a link commute with the transpositions of Lemma 6.3, so r_1, r_2, r_3, r_5 stay exact.

For the r_4 holonomy, the unmodified identity $A_{v_0}B_{v_1}A_{v_2}^{-1}B_{v_0}^{-1} = W_q\tau$ gives

$$W_q' = U_{v_0}\,(A_{v_0}V_{v_1}A_{v_0}^{-1})\,W_q\,\tau\,(B_{v_0}U_{v_2}B_{v_0}^{-1})^{-1}V_{v_0}^{-1}\tau.$$

Conjugation by a link permutation moves the plane of a rotation. For each region separately, at a plaquette other than its own source, (R1) moves the planes $\mathrm{lab}(v_1)$ and $\mathrm{lab}(v_2)$ to $\mathrm{lab}(v_0)$. At that region's own source only its bottom edge carries its potential; potentials of other regions are still treated by their respective cases. By (R4) their planes commute with the source defects of the other regions. So all rotations in W_q' lie in the planes $\mathrm{lab}_R(v_0)$; they avoid $\{1, 2\}$ by (R2), hence commute with τ, and planes of different regions are disjoint by (R4). For region R the angles add up to $(\mathrm{curl}\,\varphi_R)(q)$, and at its source W_q contributes the rotation by $s_R\pi$ in the plane of the source label. The total is $F_{R,q}$. □

Proposition 6.5 (Unitary models)**.** *For every $M \geq 48$ the model above with $k = 2$ and suitable rotations of the links gives unitaries $\rho(a), \rho(b), \rho(\alpha)$ on a space of dimension $M^2\,2^{8M+8}$ in which r_1, r_2, r_3, r_5 act as the identity, $\|\rho(r_4) - I\|_2^2 \leq 14.26\,M^{-3}$ and $\|\rho(\gamma) - I\|_2^2 = 15/8$.*

Proof. Two exact moves. Let $c^* = (M-1, M-1)$, where the r_4 holonomy is R_M (Proposition 6.1). Replace A_{c^*} by $\eta_1 A_{c^*}$ and $B_{(0,M-2)}$ by $\eta_2 B_{(0,M-2)}$, with $\eta_1 = (1\ \ M{+}4)$ and $\eta_2 = (5\ \ M{+}5)$. These satisfy the hypotheses of Lemma 6.3, so r_1, r_2, r_3, r_5 stay exact. The

modified links lie on three plaquettes. At c^* the holonomy becomes $\eta_1 R_M = \prod_{t=5}^{M+3}(t\ \ t{+}M)$. At $(0, M-2)$, where it was the identity, it becomes $\tau\eta_2\tau = \eta_2$. At $(M-1, M-2)$ it becomes $(A\eta_2 A^{-1})\,\tau(B\eta_1 B^{-1})\tau$ with $A = A_{(M-1,M-2)} = c_5$ and $B = B_{(M-1,M-2)} = c_{3M+6}$. Here $A\eta_2 A^{-1} = (1\ \ M{+}5)$ and $B\eta_1 B^{-1} = (2\ \ M{+}5)$, so $\tau(B\eta_1 B^{-1})\tau = (1\ \ M{+}5)$ and the holonomy stays the identity. Use these modified permutations as the base model in Lemma 6.4.

The regions. For $5 \le t \le M+3$ let R_t have source c^* and sign $(-1)^{t-5}$, and consist of c^*, the plaquettes $(M-1, j)$ with $1 \le j \le M-2$, and the plaquettes (i, j) with $0 \le i \le M-3$, $0 \le j \le M-2$ and $i+j \le t+M-8$; label the vertex (i, j) by $\{t{+}M{-}1{-}i{-}j,\ t{+}2M{-}1{-}j\}$ for $0 \le i \le M-2$ and $(M-1, j)$ by $\{t+M-1-j,\ t+2M-1-j\}$. Let T have source $(0, M-2)$ and sign $(-1)^{M-1}$, and consist of $(0, M-2)$ and the plaquettes (i, j) with $i, j \ge 0$ and $i+j \le M-3$, with labels $\{M+3-i-j,\ 2M+3-j\}$. The source labels are the defect pairs $\{t, t+M\}$ at c^* and $\{5, M+5\}$ at $(0, M-2)$.

These regions satisfy (R1)–(R4). Off the top row $B_{i,j} = c_{4M+4-j}$ lowers every label coordinate by one, which is the change of the labels from j to $j+1$. In the triangles $A_{i,j} = c_{2M+2-i-j}$ lowers the first label coordinate, which is at most $2M+2-i-j$ since $t \le M+3$, and fixes the second, which exceeds $2M+2-i-j$ since $t \ge 5$; this is the change from i to $i+1$, and the same holds in T. On the column $i = M-1$, $A_{M-1,j} = c_{M+3-j}$ fixes both label coordinates, which are at least $t+M-1-j \ge M+4-j$, and the labels at $(M-1, j)$ and $(0, j)$ agree; this is the only crossing of a seam, since the rows stay below $M-1$ and c^* is attached only through its bottom edge. Every base label and every label at the initial vertex of an interior edge has both coordinates at least 5. Two labels of bulk regions at one vertex differ in t by at most $M-2$, so they are disjoint; a bulk label meets a label of T only if $t \in \{4,\ 4-M-i,\ M+4+i\}$, which is impossible; and η_1, η_2 are disjoint from the labels $\{t, t+M\}$ and $\{t+1, t+M+1\}$ on the links where those regions carry rotations. At $(0, M-2)$ the label of T equals the support of η_2, but that b-link is a boundary link of T and carries no T-rotation. Each region is connected through interior edges: the triangle is a staircase, the column is glued to it through the b-edges at $(0, j)$, and each source through its bottom edge. Finally $R_t \subset R_{t+1}$, and every region has at least $m_M = (M-1)(M-2)/2 + 1$ plaquettes.

The defect. Take $F_{R,q} = s_R\pi/|R|$ on each region and $k = 2$. By Lemma 6.4 the r_4 holonomy at q is a product of phase rotations by angles $F_{R,q}$ in disjoint planes, so $\det((I + W_q')/2) = \prod_R (1 + e^{iF_{R,q}})/2 = e^{i\Theta_q} C_q$ with $\Theta_q = \frac{1}{2}\sum_R F_{R,q}$ and $C_q = \prod_R \cos(F_{R,q}/2) \in [0, 1]$. Using $1 - C^k \le k(1-C)$, $1 - \prod_R c_R \le \sum_R (1 - c_R)$ and $1 - \cos x \le x^2/2$,

$$2 - 2\,\mathrm{Re}(e^{ik\Theta_q} C_q^k) = 2(1 - C_q^k) + 2C_q^k(1 - \cos k\Theta_q) \le \frac{k}{4}\sum_R F_{R,q}^2 + k^2\Theta_q^2.$$

Summed over the plaquettes, $\sum_q \sum_R F_{R,q}^2 = \sum_R \pi^2/|R| \le M\pi^2/m_M$. The bulk regions containing q are those with t at least some t_q; their signs alternate and their magnitudes $\pi/|R_t|$ decrease, so they sum to at most π/m_M in absolute value, and T adds at most π/m_M. Hence $|\Theta_q| \le \pi/m_M$, and averaging over the M^2 clocks,

$$M^3 e_4^2 := M^3 \|\rho(r_4) - I\|_2^2 \le \frac{\pi^2 M^2}{2m_M} + \frac{4\pi^2 M^3}{m_M^2} \le \frac{\pi^2 M^2}{(M-1)(M-2)} + \frac{16\pi^2 M^3}{(M-1)^2(M-2)^2},$$

which decreases in M and is below 14.26 at $M = 48$. So $e_4^2 \le 14.26\, M^{-3}$ for $M \ge 48$.

The three-cycle. By Lemma 6.3, γ acts at every clock as a three-cycle σ of coordinates, so $\tau(\rho(\gamma)) = \det((I+\sigma)/2)^2 = 4^{-2}$ and $\|\rho(\gamma) - I\|_2^2 = 2 - 2\cdot 4^{-2} = 15/8$. □

Proposition 6.6 (Sharp defect within the fixed-region construction)**.** *Fix an integer $k \ge 1$, the clock torus, the regions and their signs in the proof above, and the k-fold exterior-algebra*

representation. Allow arbitrary real fields $F_{R,q}$ with $\sum_{q\in R} F_{R,q} = s_R\pi$, realized by Lemma 6.4. If $E_{M,k}$ is the infimum of the squared r_4 defect over these fields, then

$$E_{M,k} = \frac{k\pi^2\sqrt{2}\ln(1+\sqrt{2})}{4}\, M^{-3}(1+o(1)) \qquad (M\to\infty).$$

This optimizes only these fixed regions and commuting phase planes, not arbitrary unitary models of H_3.

Proof. Put $S_M = \sum_R |R|^{-1}$ and $m_M = 1+(M-1)(M-2)/2$. Reduce each field angle modulo 2π to $\theta_{R,q} \in [-\pi,\pi]$. Its sum over a region is an odd multiple of π, so $\sum_{q\in R}\theta_{R,q}^2 \ge \pi^2/|R|$ by Cauchy–Schwarz. The absolute normalized exterior trace at q is $\prod_{R\ni q}|\cos(\theta_{R,q}/2)|^k \le \exp(-k\sum_{R\ni q}\theta_{R,q}^2/8)$. Here $\cos x \le e^{-x^2/2}$ for $|x|\le\pi/2$, as follows by differentiating $\log\cos x + x^2/2$. For nonnegative a_q, $\sum_q(1-e^{-a_q}) \ge 1-e^{-\sum_q a_q}$. Consequently

$$\frac{2}{M^2}(1-e^{-k\pi^2 S_M/8}) \le E_{M,k} \le \frac{k\pi^2 S_M}{4M^2} + \frac{k^2\pi^2}{m_M^2}.$$

The upper bound is the constant field in each region, with the alternating signs and phase estimate in the preceding proof. Directly counting the staircase gives $|R_t| = (M-1)^2 - h(h+1)/2$ for $h = M+3-t \in \{0,\dots,M-2\}$ and $|T| = m_M$. The Riemann sum therefore yields

$$MS_M \longrightarrow \int_0^1 \frac{dx}{1-x^2/2} = \sqrt{2}\ln(1+\sqrt{2}).$$

Since $S_M = O(M^{-1})$ and m_M is of order M^2, the two bounds have the same leading term. □

Corollary 6.7 (Unitary upper bound)**.** *Let $0 < e \le e_0 = 14.26^{1/2}\cdot 48^{-3/2}$ and $M = \lceil (14.26/e^2)^{1/3}\rceil$. Johnson's presentation has an e-representation on $\mathbb{C}^D$ with $\|\phi(\gamma)-I\|_2 \ge 1/2$ and $\log D \le 8M+8+2\log M$. So such representations exist with $\log D = O(e^{-2/3})$.*

Proof. The choice of M gives $M \ge 48$ and $14.26\,M^{-3} \le e^2$; apply Proposition 6.5. □

Our permutation construction gives the weaker bound in this norm. Read as permutation matrices, the corner models with $s=2$ have relator defects at most $\sqrt{2}/M$ in $\|\cdot\|_2$, since a permutation moving a fraction f of the points has $\|P-I\|_2^2 = 2f$, and $\|\phi(\gamma)-I\|_2^2 = 3/2$, at $\log D = 6M+6+2\log M$; this gives only $\log D = O(e^{-1})$. Spreading the defect lowers the exponent of $1/e$ in our upper bounds from 1 to $\frac{2}{3}$; whether optimal permutation models are weaker than optimal unitary ones is open (Section 9). The lower bound of Theorem 1.4 has exponent $2\beta_H \approx 0.138$.

7 The correction

We now compare Corollary 1.2 with [8, Theorem 3.15]. That theorem asserts a symmetric chunk E' of the kernel and a finite symmetric $S \subset \mathbb{Z}^2$ with $\sigma_{E_H}(r) \le \sigma_{E'}(r)\,\alpha_{\mathbb{Z}^2,S}(r)$ for all $r>1$, where $\alpha_{Q,S}(r)$ is the least size of a finite $F\subset Q$ with $|SF\setminus F| < |F|/r$. The kernel $\mathrm{FSym}(\mathbb{N})$ is locally finite, so E' lies in a finite subgroup and $\sigma_{E'}$ is bounded [8, Fact 3.5]; boxes show $\alpha_{\mathbb{Z}^2,S}(r) = O(r^2)$. The resulting bound $\sigma_{E_H}(r) = O(r^2)$ is incompatible with Theorem 1.1. Since H_3 is elementary amenable, Theorem 1.1 also contradicts the conclusion of [8, Example 3.16]. The comparison does not need Theorem 3.2: with the certified bound of Theorem 3.1 alone, the proof of Theorem 1.1 gives $\sigma_{E_H}(r) \ge 2^{r^{1/12}/270-1/16}$.

Where the argument breaks. Theorem 3.15 is stated in [8] as what the explicit construction of Elek and Szabó [14] yields "without any change". That construction, given a finite $F \subset G$ and $\varepsilon > 0$, chooses a Følner set $A \subset Q$ with $|Ag \setminus A| \le \varepsilon |A|$ for $g \in F$, a section $\sigma : Q \to G$, and then an approximation of the kernel on the finite set

$$H_\varepsilon = N \cap (\sigma(A)\, F\, \sigma(A)^{-1}),$$

acting on $B \times A$ with B the kernel's approximation space [14, proof of Theorem 1(3)]. The kernel chunk is therefore H_ε, and it depends on $\varepsilon = 1/r$ through A. For H_3 with the section $(m,n) \mapsto a^m b^n$ and $A = [0,M)^2$, it contains $a^m b^n \alpha b^{-n} a^{-m}$, the transposition of $(1, m+n+1)$ and $(1, m+n+2)$, for all $0 \le m, n < M$: at least $2M-1$ distinct transpositions, which generate the symmetric group of $2M$ points, with M of order r. No chunk independent of r contains them. What the construction does give is the bound of Proposition 7.1 below, in which the kernel chunk grows with r. The upper bound by the isoperimetric profile [8, Proposition 3.14] is not affected, and H_3, having exponential growth, has isoperimetric profile at least exponential. The proof of the linear upper bound in [8, Proposition 3.17] for Baumslag–Solitar groups applies Theorem 3.15 in the same way, so it is incomplete as well; Section 8.4 proves it by a direct construction.

A scale-dependent extension bound. The construction of Elek and Szabó, with the accuracy tracked explicitly, gives the following correct form of the extension bound. Let $1 \to N \to G \xrightarrow{\pi} Q \to 1$ be an extension, $s : Q \to G$ a section with $s(e) = 1$, and put $c(a,g) = s(a)\, g\, s(a\pi(g))^{-1} \in N$ for $a \in Q$ and $g \in G$.

Proposition 7.1. *Let $E \subset G$ be a symmetric chunk, $r > 1$, and $A \subset Q$ a finite nonempty set with $|A\pi(g) \setminus A| \le |A|/(4r)$ for every $g \in E$. Then $K_{A,E} = \{c(a,g) :\ g \in E,\ a \in A,\ a\pi(g) \in A\}$ is a symmetric chunk of N, and*

$$\sigma_E(r) \ \le\ |A|\, \sigma_{K_{A,E}}(2r).$$

Proof. The set $K_{A,E}$ contains $c(a,1) = 1$, and $c(a,g)^{-1} = c(a\pi(g), g^{-1})$, so it is symmetric. The cocycle identity $c(a,g)\, c(a\pi(g), h) = c(a, gh)$ holds in G. Let $f_K : K_{A,E} \to \mathrm{Sym}_m$ be a $(1 - \frac{1}{2r})$-expansive $\frac{1}{2r}$-morphism with $m = \sigma_{K_{A,E}}(2r)$, and on $A \times [m]$ let $f(g)$ send (a,x) to $(a\pi(g), x \cdot f_K(c(a,g)))$ whenever $a\pi(g) \in A$, completed arbitrarily to a permutation; $f(1)$ is the identity. Call a fiber a good for a relation $gh = k$ in E if a, $a\pi(g)$ and $a\pi(k)$ lie in A. Then $c(a,g)c(a\pi(g),h) = c(a,k)$ is a relation of the chunk $K_{A,E}$, so on a good fiber $f(g)f(h)$ and $f(k)$ differ on at most a fraction $\frac{1}{2r}$ of the points, and the bad fibers form a fraction at most $\frac{2}{4r}$. So f is an r^{-1}-morphism. For distinct $g, h \in E$: if $\pi(g) \ne \pi(h)$, then on every fiber a with $a\pi(g), a\pi(h) \in A$ the two images have different quotient coordinates, a fraction at least $1 - \frac{1}{2r}$ of the points; if $\pi(g) = \pi(h)$, then $c(a,g) \ne c(a,h)$ are distinct elements of $K_{A,E}$, so on the fibers with $a\pi(g) \in A$ the images differ on a fraction at least $1 - \frac{1}{2r}$ of each fiber, and in total on at least $(1 - \frac{1}{4r})(1 - \frac{1}{2r}) \ge 1 - \frac{1}{r}$. So f is $(1 - r^{-1})$-expansive, on $|A|\, m$ points. □

In the setting of Arzhantseva and Cherix, where one parameter is both the radius of the ball and the inverse accuracy, a scale-dependent extension bound for finitely generated kernels is [2, Theorem 5.5 and Corollary 5.6] (Theorem 38 and Corollary 39 in arXiv:2008.12954v1); Proposition 7.1 is the chunk-indexed analogue for arbitrary kernels, which is what the extension theorem of [8] needs.

When the cocycle takes finitely many values on $Q \times E$, for instance for direct products, $K_{A,E}$ lies in a fixed finite set and a bound with a fixed kernel chunk follows, as Theorem 3.15 asserts. For H_3 with boxes $A = [0,M)^2$ and M of order r, K_{A,E_H} lies in the finite group

of permutations of the first $2M + O(1)$ points of ray 1, whose regular representation gives $\sigma_{K_{A,E_H}} \le (2M + O(1))!$; so the proposition gives $\log \sigma_{E_H}(r) = O(r \log r)$, consistent with, and weaker than, Proposition 6.2.

Kernel and quotient do not determine the profile. The group $G_0 = \mathrm{FSym}(\mathbb{Z}^2) \rtimes \mathbb{Z}^2$, with $\mathbb{Z}^2$ acting by translation, has kernel and quotient isomorphic to those of H_3, since $\mathrm{FSym}(\mathbb{Z}^2) \cong \mathrm{FSym}(\mathbb{N})$, but every chunk of G_0 has bounded profile. It is the $\mathbb{Z}^2$-analogue of the Vershik–Gordon example $\mathrm{FSym}(\mathbb{Z}) \rtimes \mathbb{Z}$, and it is LEF [38, 4, 1]; we include the short argument. Indeed, let $E \subset G_0$ be a chunk, and let $B \subset \mathbb{Z}^2$ be a finite set containing the translation parts of the elements of E, the supports of their permutation parts, and all translates of these supports by the translation parts and their negatives. For q larger than twice the largest absolute value of a coordinate in B, reduction modulo q is injective on B, so reducing the permutation parts and translation parts gives an injective map of E into the finite group $\mathrm{Sym}((\mathbb{Z}/q)^2) \rtimes (\mathbb{Z}/q)^2$ that preserves every product defined in E, and Lemma 2.1 applies. So an extension bound must use the action or the cocycle, as Proposition 7.1 does through the chunks $K_{A,E}$.

Statement identifiers in [8] refer to the published journal version: Theorem 3.15, Example 3.16 and Proposition 3.17 are on p. 833, and Problem 3.18 and Example 3.19 are on p. 834. Proposition 3.17 is labelled Example 3.17 in arXiv:1305.0993v1.

8 Other groups

8.1 A criterion from finitely many relations

For a finite set S and a finite set R of words in $S^{\pm 1}$, let $\mathrm{Area}_R(w)$ be the least number of conjugates of words of $R^{\pm 1}$ whose product is w in the free group on S, and ∞ if there is none.

Theorem 8.1 (Finitely many relations)**.** *Let G be a group, $S \subset G$ a finite set containing u, v, and R a finite set of words in $S^{\pm 1}$, each representing 1 in G, with $u^2, v^2, (uv)^3 \in R$ and $uv \neq 1$ in G. Let l be the largest length of a word in R, and let E be the symmetric chunk consisting of 1, $S^{\pm 1}$, the prefixes of the words of R and of uv, and the inverses of these elements. Let $w_1, \dots, w_N$ be words in $S^{\pm 1}$ and $\Lambda \ge 1$ a number such that*

$$\mathrm{Area}_R([w_i x w_i^{-1}, w_j y w_j^{-1}]) \le \Lambda \qquad (i \neq j,\ x, y \in \{u, v\}).$$

Then $\log \sigma_E(r) \ge N/16$ for every $r \ge 10^8 N \Lambda (2l - 1)$. Moreover, let $0 < \Delta \le 1$, let K_Δ be as in Proposition 4.6, and let ϕ assign unitaries in $U(D)$ to S so that every word of R is mapped within e of I in the norm $\|\cdot\|_2$ and $\|\phi(uv) - I\|_2 \ge \Delta$. If $K_\Delta \sqrt{N}\, \Lambda e \le 1$, then $\log D \ge N\Delta^2/12$.

The chunk E does not depend on the words w_i, which may be long; only the areas matter. For H_3, Sections 4 and 5 use $u = \alpha$ and $v = a\alpha a^{-1}$, with $uv = \gamma$, conjugated by $w_i = a^{3(i-1)}$.

Proof. Take the copies $A_i = \phi(w_i)\phi(u)\phi(w_i)^{-1}$ and $B_i = \phi(w_i)\phi(v)\phi(w_i)^{-1}$, $1 \le i \le N$, in place of the pairs used in Propositions 4.5 and 5.1. Those proofs use three properties of their pairs, and each holds here. The S_3 relations of each pair are conjugates of $u^2, v^2, (uv)^3$, so they have the defect of a relator. The product $A_i B_i = \phi(w_i)\phi(uv)\phi(w_i)^{-1}$ has the displacement of $\phi(uv)$. A commutator of two letters from copies $i \neq j$, or of their inverses, is a conjugate of $[w_i x w_i^{-1}, w_j y w_j^{-1}]^{\pm 1}$, so its area is at most Λ, which takes the place of A_N. For unitaries the proof of Proposition 4.6 then gives $\log D \ge N\Delta^2/12$. For permutations, let $f : E \to \mathrm{Sym}_n$ be a $(1 - r^{-1})$-expansive r^{-1}-morphism and put $\phi(s) = f(s)$ for $s \in S$. As in the proof of Theorem 1.1, every word of R has defect at most $\delta = (2l - 1)/r$, and

$d_n(\phi(uv), \mathrm{id}) \geq 1 - 2/r \geq \frac{1}{2}$. Proposition 5.1, with γ replaced by uv and A_N by Λ, gives $\log n \geq N/16$ whenever $10^8 N\Lambda\delta \leq 1$. □

Corollary 8.2 (Transport by one element)**.** *Let G, S, R, l and E be as in Theorem 8.1, let $t \in S$, and suppose that for some $C \geq 1$ and $p \geq 0$*

$$\mathrm{Area}_R([x, t^k y t^{-k}]) \leq C k^p \qquad (x, y \in \{u, v\},\ k \geq 1).$$

Then

$$\log \sigma_E(r) \ \geq \ \frac{1}{16} \Big\lfloor \Big(\frac{r}{10^8 C(2l-1)} \Big)^{1/(p+1)} \Big\rfloor \qquad (r > 1),$$

and every unitary assignment ϕ as in Theorem 8.1 has $\log D \geq \frac{\Delta^2}{12}\big((CK_\Delta e)^{-2/(2p+1)} - 1\big)$.

Proof. Take $w_i = t^{i-1}$ for $1 \leq i \leq N$. For $i < j$ the commutator $[t^{i-1} x t^{1-i}, t^{j-1} y t^{1-j}]$ is freely conjugate to $[x, t^{j-i} y t^{i-j}]$, and for $i > j$ it is the inverse of such a commutator with x and y exchanged, so $\Lambda = CN^p$ bounds every area. The condition $r \geq 10^8 N \cdot CN^p(2l-1)$ holds for $N = \lfloor (r/(10^8 C(2l-1)))^{1/(p+1)} \rfloor$, and the condition $K_\Delta \sqrt{N}\, CN^p e \leq 1$ for $N = \lfloor (CK_\Delta e)^{-2/(2p+1)} \rfloor$. When $N = 0$ the bounds are trivial. □

The group G need not be finitely presented; only the finite set R matters. The finite set R cannot be replaced by an infinite family of relations: in $S_3 \wr \mathbb{Z}$ every far commutator is a relation of the group, but every chunk embeds in a finite group $S_3 \wr (\mathbb{Z}/L)$ for large L, so its profile is bounded. The profile depends only on the chunk and its partial multiplication, which injective homomorphisms preserve; so a bound for a chunk of a subgroup holds in every larger group.

The conditions of Theorem 8.1 force $N \leq r/10^8$ and $N \leq (K_\Delta e)^{-2}$, so the bounds it yields never exceed $\log \sigma_E(r) \geq r/(16 \cdot 10^8)$ or $\log D \geq \Delta^2/(12 K_\Delta^2 e^2)$, that is, order r and order e^{-2}. Transport by one element places N copies at distances up to N, and with area of order k^p at distance k it reaches N of order $r^{1/(p+1)}$. If words of length $O(\log N)$ carry a copy of S_3 to N pairwise commuting copies whose commutators have area polylogarithmic in N, the theorem gives $\log \sigma_E(r)$ of order r up to logarithms. The companion paper [11] realizes this in an elementary amenable group, and in a subgroup of Thompson's group V. The recent preprint of Chaudkhari, Juschenko and Muliarchyk, arXiv:2610.25415v1, Theorem 1.1 [40], proves soficity of V; using that result, the subgroup's chunk has finite profile.

8.2 Whole suffix sweeps

Pairwise commutator estimates can overcount the cost of moving one copy past an ordered suffix. The following form charges one filling for the whole suffix. It uses permutation rounding in Hamming distance before passing to the entropy estimate.

Theorem 8.3 (Sweep counting criterion)**.** *Write $a_0 = u$, $b_0 = v$ for the two abstract S_3 letters. Use the group, relators, chunk and transport words of Theorem 8.1, and suppose $N \geq 1$. For each of the six fixed normal forms*

$$\mathcal{N} = \{1, a_0, b_0, a_0 b_0, b_0 a_0, a_0 b_0 a_0\}, \qquad W_i(g) = w_i g w_i^{-1},$$

assume that, for every i, every $h \in \mathcal{N}$ and every choice $g_j \in \mathcal{N}$ for $j > i$,

$$\mathrm{Area}_R\left([W_i(h), W_{i+1}(g_{i+1}) \cdots W_N(g_N)]\right) \leq \Lambda_N.$$

Then

$$(\Lambda_N + 40N)(2l-1) \leq r/20000 \quad \Longrightarrow \quad \log \sigma_E(r) \geq N/16. \tag{4}$$

If only the sweeps with $h = a_0, b_0$ are bounded by Λ_N, the same conclusion holds with $r/60000$ in the hypothesis in place of $r/20000$. In both versions every later factor ranges over all six normal forms.

Proof. Relator errors. Given a chunk model $f : E \to \mathrm{Sym}_n$, evaluate generator words by $\phi(s) = f(s)$ and $\phi(s^{-1}) = f(s)^{-1}$. Merging successive prefixes of a relator of length m costs at most $(m-1)/r$; replacing inverse letters costs at most m/r, since $s, s^{-1}, 1 \in E$. Thus each relator has Hamming defect at most $\epsilon = (2l-1)/r$. Bi-invariance implies that a word of area A has defect at most $A\epsilon$. The prefix of $a_0 b_0$ and expansiveness give $d(\phi(a_0 b_0), 1) \geq 1 - 2/r$. All transported pairs have the same square and braid errors and three-cycle displacement: conjugation is an exact Hamming isometry, regardless of the transport word's length.

Native rounding. Let A, B be any transported pair. Replace cycles of length at least three in each permutation by fixed points, obtaining involutions P, Q with $d(P, A), d(Q, B) \leq \epsilon$. Put $R = PQ$. Then $d(R^3, 1) \leq 7\epsilon$. The identity $QRQ = R^{-1}$ implies that $\mathrm{Fix}(R^3)$ is invariant under Q, R, and hence under $P = RQ$. Keep P, Q on this set and make them both trivial on its invariant complement. They now give an exact permutation representation $u_i : S_3 \to \mathrm{Sym}_n$. Every normal form has length at most three: its first replacement costs at most 3ϵ, and deleting bad orbits costs at most 7ϵ. Consequently

$$d(u_i(g), \phi(W_i(g))) \leq 10\epsilon \qquad (g \in \mathcal{N}).$$

This construction allows trivial and sign orbits. In particular, the three-cycle loses at most 10ϵ of its Hamming displacement.

One ordered error. Put $U_{>i} = u_{i+1}(g_{i+1}) \cdots u_N(g_N)$ and let $V_{>i}$ be its raw-word counterpart. Replacing both occurrences of the first factor and of the suffix in the commutator gives

$$d([u_i(h), U_{>i}], 1) \leq (\Lambda_N + 20(N - i + 1))\epsilon \leq (\Lambda_N + 40N)\epsilon.$$

Let $F(g) = u_1(g_1) \cdots u_N(g_N)$, for uniform $g \in S_3^N$. Exact same-copy multiplication shows that the error $F(gh) - F(g)u_i(h)$ is exactly the error in commuting $u_i(h)$ past its suffix, up to unitary factors. For permutation matrices $\|P - Q\|_2^2 = 2d(P, Q)$; hence

$$T^2 := \max_{i,h} \mathbb{E}_g \|F(gh) - F(g)u_i(h)\|_2^2 \leq 2(\Lambda_N + 40N)\epsilon.$$

Here h lies in factor i. There is no same-copy multiplication defect after rounding.

Entropy from this error. We give the sector argument to specify exactly what is used. Suppose $\|u_i(c) - I\|_2 \geq \Delta$ for a three-cycle c, $T \leq \Delta$ and $T^2 \leq 5/6$. On $\ell^2(S_3^N) \otimes \mathbb{C}^n$ use the regular action R and the isometry

$$Wv = |S_3^N|^{-1/2} \sum_g |g\rangle \otimes F(g)v, \qquad \Pi = WW^*, \quad \rho = \Pi/n.$$

The ordered-error identity gives $\|\mathsf{R}(h_i)W - Wu_i(h)\|_{HS} \leq \sqrt{n}\, T$. Thus, for $Z_i = 2I - \mathsf{R}(c_i) - \mathsf{R}(c_i)^*$, $\mathrm{Tr}(\rho Z_i) \geq (\Delta - T)^2$. Let $\mathcal{E}_i$ average conjugation by the six regular operators in factor i. Its leakage outside Π is at most T^2. Pinching into Π and $I - \Pi$ gives ranks at most n and $5n$, because one of the six states being averaged is ρ. Therefore $S(\mathcal{E}_i\rho) \leq \log n + h_2(T^2) + T^2 \log 5$. The $\mathcal{E}_i$ commute. Monotonicity of relative entropy, together with $S(\mathcal{E}\sigma) - S(\sigma) = D(\sigma \| \mathcal{E}\sigma)$, bounds each successive entropy increment by its increment on ρ. For $\bar{\rho} = \mathcal{E}_N \cdots \mathcal{E}_1 \rho$ this yields

$$S(\bar{\rho}) \leq \log n + N\{h_2(T^2) + T^2 \log 5\}.$$

On each irreducible sector of S_3^N, Z_i is $3I$ if factor i is standard and zero otherwise. If a sector has k standard factors, its dimension is 2^k. Since $\bar{\rho}$ commutes with the regular action,

its entropy is at least the mean of k. The twirls preserve the expectation of each central Z_i, so $N(\Delta - T)^2 \le 3S(\bar{\rho})$. Combining gives

$$\log n \ge N\{(\Delta - T)^2/3 - h_2(T^2) - T^2 \log 5\}.$$

No pairwise estimate entered this argument.

Constants. The stated hypothesis gives $T \le 1/100$ and $\epsilon \le 1/(800000N)$. The preceding displacement estimate gives $\Delta \ge 49/100$ (a deliberately conservative Hilbert–Schmidt lower bound). Using $\log_2(10000) < 14$ and $1/\ln 2 < 3/2$, one has

$$h_2(1/10000) \le 31/20000, \qquad (1/10000)\log 5 \le 6/20000.$$

Thus the entropy per copy is at least $(48/100)^2/3 - 37/20000 = 1499/20000 > 1/16$. For elementary-letter sweeps, split h into at most three letters and use the free commutator product identity. The full sweep area is at most $3\Lambda_N$, and $3\Lambda_N + 40N \le 3(\Lambda_N + 40N)$, proving the $1/60000$ version. □

The $40N$ term retains the counting method's linear scale in r even if sweep areas are very small. This is a restriction on lower bounds produced by the criterion, not an upper bound on any group's actual profile. The transported finite-group and area-to-defect strategy follows Slofstra and Slofstra–Vidick [36, 37]; metric approximation growth with growing balls, as studied by Arzhantseva–Cherix and Bradford [2, 39], is distinct from the fixed-chunk accuracy profile used here.

8.3 Block coherence and the conditional square-root bound

All areas in this subsection use Johnson's five-relator presentation P of (1). For a monotone list $L = ((q_1, c_1), \dots, (q_m, c_m))$, with $0 \le q_1 \le \cdots \le q_m \le K$ and $c_j \in \{0, 1\}$, put

$$e_0 = \alpha, \quad e_1 = a\alpha a^{-1}, \qquad A(L) = \prod_{j=1}^{m} a^{q_j} e_{c_j} a^{-q_j},$$

and let $B(L)$ be the word obtained by replacing every a by b, including those inside e_1. Every positive word of length d in a, b conjugates α to the ray-1 transposition at positions $d+1, d+2$. Thus $A(L) = B(L)$ in H_3. The freely telescoped boundary $A(L)B(L)^{-1}$ has length at most $4K + 6m$. Define the finite maximum

$$\mathrm{BC}(K; \mu) = \max_{L} \mathrm{Area}_P(A(L)B(L)^{-1}), \qquad \mathrm{BC}(K) = \mathrm{BC}(K; 3),$$

where at most μ entries occur at any position. In particular $m \le 3(K+1)$ for $\mathrm{BC}(K)$. The multiplicity restriction is essential to the definition used here.

Proposition 8.4 (Reduction to block coherence)**.** *For the copies generated by $a^{3i}\alpha a^{-3i}$ and $a^{3i+1}\alpha a^{-3i-1}$, full normal-form sweep areas satisfy*

$$\Lambda_N \le 2\mathrm{BC}(3N) + CN$$

with an absolute constant C. If H_3 has a quadratic Dehn function, or if $\mathrm{BC}(K) = O(K^2)$, then, for the existing chunk E_H,

$$\exp(c\sqrt{r}) \le \sigma_{E_H}(r) \le \exp(C'\sqrt{r}\log r) \qquad (r \ge r_0). \tag{5}$$

The existing unconditional bounds of Theorem 1.1 remain in force.

Proof. For a sweep whose first copy is at $3i$, conjugate by $a^{-(3i+3)}$. The first copy's two generators become $a^{-3}\alpha a^3$ and $a^{-2}\alpha a^2$, both supported wholly on ray 2. The suffix becomes $A(L)$ at positions $0, 3, 6, \ldots, 3(N-i)-3$, with at most three letters per position. Changing its two occurrences to $B(L)$ costs at most $2\mathrm{BC}(3N)$. Both fixed ray-2 transpositions commute with b and α in the ray action. The four resulting fixed null commutators therefore have fixed finite P-fillings, by Johnson's presentation. Commuting a word of at most three such transpositions through the telescoped $B(L)$ costs $O(|B(L)|) = O(N)$. This proves the reduction. The shift by three ensures that neither transposition straddles rays 1 and 2.

If $\mathrm{BC}(K) = O(K^2)$, then $\Lambda_N = O(N^2)$ and the native rounding proof of Theorem 8.3 applies directly to E_H. Indeed the two square defects are conjugates of r_1, the braid defect is a conjugate of r_2, and the displaced three-cycle is γ. No extra square relator or chunk element is required for the second generator $a\alpha a^{-1}$: it is evaluated as an exact conjugate. The prefix argument for E_H gives relator defect at most $23/r$ and $d(\phi(\gamma), 1) \geq 1 - 5/r$. Taking N a sufficiently small constant times $\sqrt{r}$ in the sweep estimate yields $\log \sigma_{E_H}(r) \geq N/16$. Proposition 6.2 supplies the upper side of (5).

Alternatively, a quadratic Dehn function gives quadratic BC because the boundaries have length $O(K)$. One can also see the sweep length directly: each normal form in $\alpha, a\alpha a^{-1}$ has length at most 9 in the original generators, and the monotone suffix telescopes to length at most $15N$. The first transported normal form has length at most $6N + 9$, so the entire sweep has length at most $45N + 18$. Its area is then $O(N^2)$. Both routes are conditional; neither proves the required quadratic upper filling bound. □

Proposition 8.5 (An all-scale block-coherence lower bound)**.** *For every integer $K \geq 0$,*

$$\mathrm{BC}(K) \geq K(K-1).$$

For $K \geq 2$, a single-entry list $L = ((K, 1))$ witnesses this lower bound on the number of r_4-cells in every filling.

Proof. Signed cells. Project a, b, α to $e_1, e_2, 0$ in the square lattice. Only r_4 contributes a signed unit square; the other relators have zero projected edge chain. A coherence boundary $A(L)B(L)^{-1}$ has zero projected edge chain, so the finite integral 2-chain of its r_4-cells has zero boundary. There is no nonzero finite 2-cycle in the square lattice (an outermost nonzero square would expose an edge). Thus the signed count at every offset is zero, as is its sum in each residue class modulo M. In the corner model of Proposition 6.1, a failing clock fibre of the boundary forces an r_4-cell at an offset in its corresponding residue class: without one, every cell is exact on that fibre and so is the boundary. Distinct failing fibres give distinct residue classes. Each occupied class contains at least two cells with opposite signs. Therefore

$$\#\{r_4\text{-cells}\} \geq 2\#\{\text{failing clock fibres}\}. \tag{6}$$

It suffices that the boundary has nonidentity coordinate permutation on a fibre; some colorings then move.

Single-lamp endpoints. Use the one-based coordinate labels of Proposition 6.1. At clock (i, j) put $x = M - 1 - i$, $y = M - 1 - j$. For $3 \leq d \leq M - 3$, let A_d, B_d be the endpoint pairs of $a^d\alpha a^{-d}$ and $b^d\alpha b^{-d}$. Write $D_i = B_{(i,M-1)}$ for the top-row link in that proposition. Its recursion gives

$$D_i^{-1}(v) = H_x(v) := \begin{cases} v+1, & v \leq x+3, \\ v+M-x, & v \geq x+4, \end{cases} \qquad 1 \leq v \leq M+1.$$

For completeness, invert the top-row recursion and put $h_R = \tau c_R$, which fixes 1 and cycles $2, R, R-1, \ldots, 3$. Start at $v+i+1$ and apply $h_{M+3}, h_{M+2}, \ldots, h_{M+4-i}$. If $v \leq M+2-i = x+3$,

every step decreases by one, ending at $v+1$; otherwise every step fixes the point, ending at $v+i+1 = v+M-x$. This proves the displayed formula at every scale.

With $r = y+4$, define

$$F_{r,x}(v) = \begin{cases} v+x+1, & v < r, \\ r+x+1-v, & r \le v \le r+x, \\ v, & v > r+x. \end{cases}$$

Along the inverse a-path, consecutive increasing forward cycles have at most one wrap. Before the wrap an endpoint increases by one at each step; at the wrap the labels $r, \dots, r+x$ are sent in reverse order to $x+1, \dots, 1$ by the resulting cycle product. After it, labels above $r+x$ are fixed. These three cases give $F_{r,x}$ and hence

$$A_d = \begin{cases} (d+1, d+2), & d \le x+1, \\ (F_{r,x}(d-x), F_{r,x}(d-x+1)), & d > x+1. \end{cases}$$

For the b-path the same no-wrap case applies; after the wrap the top-row inverse just computed is followed by y ordinary steps. Thus

$$B_d = \begin{cases} (d+1, d+2), & d \le y+1, \\ (H_x(d-y)+y, H_x(d-y+1)+y), & d > y+1. \end{cases}$$

Substitution shows that if $x+y \ge d-2$, both pairs are $(d+1, d+2)$. If $x+y \le d-3$, the second pair has an endpoint at least $M+4$, while both endpoints of the first are at most $d+2 \le M-1$. Consequently the single-lamp coherence fails exactly at

$$x, y \ge 0, \qquad x+y \le d-3,$$

which consists of $(d-1)(d-2)/2$ clock fibres. For $L = ((K,1))$ take $d = K+1$ and $M \ge K+4$; (6) gives $K(K-1)$ cells. For $K = 0, 1$ the bound is zero and follows from nonnegativity of area. □

The maximum in $\mathrm{BC}(K)$ is important. There is no analogous lower bound for every list with deepest entry K: $L = ((K,1),(K,1))$ has a two-cell filling using conjugates of $r_1 = \alpha^2$, and its boundary is exact in every corner fibre. Likewise a quadratic total budget constrains only the total and average cell count; it imposes no constant multiplicity at each offset. Finite checks of the endpoint formulas and of both cancelling and noncancelling lists accompany the manuscript. They support the all-scale calculation above and do not establish a quadratic upper bound for BC.

8.4 Baumslag–Solitar groups

Theorem 8.6 (Baumslag–Solitar groups; Cornulier's Proposition 3.17 with a direct proof)**.** *Let $m, n \neq 0$ and $G = BS(m,n) = \langle a, t \mid ta^m t^{-1} = a^n \rangle$.*

(a) Every chunk E of G has $\sigma_E(r) \le C_E r$ for $r \ge 2$.

(b) If $|m|, |n| \ge 2$ and $|m| \neq |n|$, some chunk of G does not embed in a finite group, so its profile is at least $2r$ for every $r > 1$, which sharpens Cornulier's lower bound r.

Part (a) is the upper bound of [8, Proposition 3.17], whose proof there used Theorem 3.15. Part (b), with r in place of $2r$, is also in [8, Proposition 3.17], where it follows from [8, Fact 3.8] and Meskin's theorem; we include a direct proof of (b) with an explicit chunk. By

Meskin's theorem [30], the groups in (b) are exactly the Baumslag–Solitar groups that are not residually finite, and (b) shows directly that they are not LEF. In the remaining cases, $|m| = 1$, $|n| = 1$ or $|m| = |n|$, every chunk of G has bounded profile [8, Proposition 3.17]: these groups are residually finite by Meskin's theorem, and a finite quotient that separates the finitely many elements $g^{-1}h$, $g \neq h$ in a chunk E, gives an embedding of E in a finite group.

Proof. Write $M = |m|$, $N = |n|$ and $\mu = MN$. Permutations act on the right, as in Section 2, and d is the normalized Hamming distance. We use Britton's lemma [29]: a word in $a^{\pm 1}, t^{\pm 1}$ is *reduced* if it contains no subword ta^jt^{-1} with $m \mid j$ and no subword $t^{-1}a^jt$ with $n \mid j$; every word can be brought to reduced form by free reductions and the replacements $ta^{mj}t^{-1} \to a^{nj}$ and $t^{-1}a^{nj}t \to a^{mj}$, each using $|j|$ conjugates of the relator; and a reduced word containing $t^{\pm 1}$ is not the identity.

(a) Tiles. Fix integers $L \geq 1$ and $\ell \geq 1$. Since free groups are residually finite, there are a finite group Φ and elements $s_1, \dots, s_\mu \in \Phi$ such that no nonempty freely reduced word of length at most L in them is trivial. Let $X = \mathbb{Z}/\mu\ell$ and $\Omega = X \times \Phi$, and let A act by $(x, h) \mapsto (x + 1, h)$. The permutation A^m of X has M cycles, the residue classes modulo M, each of length $N\ell$. Cut each cycle into N arcs of ℓ consecutive points $p_i(0), \dots, p_i(\ell - 1)$, with $p_i(k + 1) = p_i(k) + m$; this gives tiles $P_1, \dots, P_\mu$. In the same way the cycles of A^n give tiles $Q_1, \dots, Q_\mu$ with points $q_i(k)$ and $q_i(k + 1) = q_i(k) + n$. Let Z_m move each $p_i(k)$ to $p_i(k + 1 \bmod \ell)$ and Z_n each $q_i(k)$ to $q_i(k + 1 \bmod \ell)$, keeping the Φ-coordinate. Define

$$(q_i(k), h) \cdot T = (p_i(k), hs_i).$$

Then $TZ_mT^{-1} = Z_n$ exactly. The cycle Z_m differs from A^m only at the last point of each tile, so $d(Z_m, A^m) \leq 1/\ell$, and likewise $d(Z_n, A^n) \leq 1/\ell$. By bi-invariance the relator $ta^mt^{-1}a^{-n}$ has defect $d(TA^mT^{-1}, A^n) \leq 2/\ell$.

Reduced words move every point. Let $v = a^{k_0}t^{\epsilon_1}a^{k_1} \cdots t^{\epsilon_j}a^{k_j}$ be reduced with $1 \leq j \leq L$. Along the action of v, each T multiplies the Φ-coordinate on the right by the s_i of the tile Q_i it leaves, and each T^{-1} by s_i^{-1} for the tile P_i it leaves. For a subword ta^kt^{-1}, the point x reached after t lies in some tile P_i, which lies in one residue class modulo M; since $m \nmid k$, the point $x + k$ lies in another residue class, hence in another tile $P_{i'}$, and $s_is_{i'}^{-1}$ does not cancel. The case $t^{-1}a^kt$ is the same modulo N, and letters of equal sign do not cancel. So v multiplies the Φ-coordinate by a nonempty reduced word of length $j \leq L$, a nonidentity element, and $v(A, T)$ fixes no point. A power a^k with $0 < |k| < \mu\ell$ also fixes no point.

From words to chunks. For $g \in E$ fix a word w_g, with w_1 empty. A reduced word without stable letters is a power of a, which is trivial only when the exponent is zero, since a has infinite order. So for $g, h, gh \in E$ Britton's lemma reduces the null word $w_gw_hw_{gh}^{-1}$ to the empty word, and for $g \neq h$ it turns $w_g^{-1}w_h$ into a reduced word $v_{g,h}$ that either contains a stable letter or is a nonzero power of a. Let A_E bound the number of relator conjugates used in these finitely many reductions, L_E the largest number of stable letters in a $v_{g,h}$, and K_E the largest absolute exponent of those $v_{g,h}$ that are powers of a. Take $L = \max\{1, L_E\}$ and $\ell = \lceil 2A_Er \rceil + K_E + 1$, and put $f(g) = w_g(A, T)$. Each relator conjugate costs at most $2/\ell$, so $d(f(g)f(h), f(gh)) \leq 2A_E/\ell \leq 1/r$, and $d(f(g), f(h)) \geq d(v_{g,h}(A, T), \mathrm{id}) - 2A_E/\ell \geq 1 - 1/r$. The degree $\mu\ell|\Phi|$ is at most C_Er, because Φ depends only on L_E.

(b) Let $d_0 = \gcd(M, N)$, and let $\bar{a}, \bar{t}$ be the images of a, t in a finite quotient, with $\bar{a}$ of order q. Since $\bar{a}^m$ and $\bar{a}^n$ are conjugate, they have the same order, so $k = \gcd(q, M) = \gcd(q, N)$ divides d_0, and $\bar{t}$ conjugates $\langle \bar{a}^m \rangle = \langle \bar{a}^k \rangle$ onto $\langle \bar{a}^n \rangle = \langle \bar{a}^k \rangle$, so it normalizes this subgroup. Its subgroup $\langle \bar{a}^{d_0} \rangle$ is characteristic, so $\bar{t}$ and $\bar{t}^{-1}$ normalize it, and $\bar{t}^{\pm 1}\bar{a}^{d_0}\bar{t}^{\mp 1}$ commutes with $\bar{a}$. Put $w = [ta^{d_0}t^{-1}, a]$ if $d_0 < M$, and $w = [t^{-1}a^{d_0}t, a]$ otherwise, when $d_0 = M < N$. Then w is trivial in every finite quotient, and it is reduced and contains t, so $w \neq 1$. Let E_*

consist of $1, a^{\pm 1}, t^{\pm 1}$, the prefixes of $ta^m t^{-1} a^{-n}$ and of w, and their inverses. An embedding of E_* in a finite group Q sends a, t to elements satisfying the defining relation, hence defines a homomorphism $G \to Q$, and it sends w to a nonidentity element, which is impossible. Lemma 2.1 completes the proof. □

Remark 8.7 (Several stable letters). Every chunk of $\langle a, t_1, \dots, t_k \mid t_i a^{m_i} t_i^{-1} = a^{n_i} \rangle$ also has at most linear profile. The same proof applies: take $X = \mathbb{Z}/C\ell$ with C divisible by every $|m_i|$ and $|n_i|$, cut it into tiles for each i, and give each pair of a stable letter and a tile its own generator of the finite group. In $BS(m, n)$ the obstruction to exact finite models is a mismatch of cyclic scales, a^m against a^n, and the construction pays for it with one defect per tile. Displaced copies of a non-abelian group, as in H_3, cost far more.

9 Open problems

(1) *The exponents.*

(a) The sofic exponent of E_H in r lies between $1/(\log 107 + 1) \approx 1/7.74$ and $1/2$ (up to the logarithm). For unitary approximations with relator defect e in the normalized Hilbert–Schmidt norm, $\log D$ lies between order $e^{-2\beta_H}$, with $2\beta_H \approx 0.138$ (Theorem 1.4), and $O(e^{-2/3})$ (Corollary 6.7). What are the true exponents?

(b) Since the area of the far commutators along a ray is at least quadratic (Proposition 3.4), transport by the powers of a along a ray, through Theorem 8.1's pairwise estimate, cannot give more than $1/3$ in r for permutations and $2/5$ in $1/e$ for unitaries. The orbits of H_3 grow linearly, so the transport of many copies used in [11] is not available as it stands. Given a finite set R of relations of H_3 and a pair u, v as in Theorem 8.1, is there $c > 0$ such that for all words $w_1, \dots, w_N$ some commutator $[w_i x w_i^{-1}, w_j y w_j^{-1}]$, $i \neq j$, $x, y \in \{u, v\}$, has area at least cN^2? If so, these values are the limits of Theorem 8.1 for H_3.

(c) A quadratic Dehn function for H_3 would reach these limits of the pairwise theorem, since every word in that argument has linear length. The sweep strengthening now gives the stronger conditional bound (5), with permutation exponent $1/2$, superseding the pairwise conclusion. Theorem 3.3 gives the polynomial upper bound $O(n^{\log 107 + 4})$, and the Dehn function is at least quadratic. Is it quadratic? For Thompson's group F a quadratic Dehn function is a theorem [21], proved with planar diagram groups, which are torsion-free [22]. The group H_3, which has torsion, is a braided diagram group [15, Example 4.3] but not a planar one, and the method does not apply to it directly; torsion alone is no obstruction, since Thompson's group T has a quadratic Dehn function [31].

(d) The quadratic lower bound of Proposition 3.4 counts only cells of type r_4; do $O(k^2)$ cells of type r_4 suffice to fill C_k? For the bounded-multiplicity block problem defined in Section 8, Proposition 8.5 gives $\mathrm{BC}(K) \geq K(K-1)$ at every scale. Does the matching all-cell upper bound $\mathrm{BC}(K) = O(K^2)$ hold? By Proposition 8.4 it would imply (5).

Whole-suffix fillings and the native Hamming error estimate now improve the pairwise implication. The quadratic upper filling bound needed to make that improvement unconditional remains open.

(2) *Permutation versus unitary approximation.* Relator errors add up in measure for permutations and in norm for unitaries, so the two profiles may differ. In common units, a permutation matrix has squared normalized Hilbert–Schmidt defect twice

its normalized Hamming defect. At squared defect $\delta = e^2$, our permutation models give $\log D = O(\delta^{-1/2})$ and unitary models $O(\delta^{-1/3})$ (Section 6.2). That our two constructions differ does not show that the optimal profiles differ. A permutation lower exponent for E_H strictly above $1/3$ would separate them using this unitary upper bound. Can a different method prove a separation?

(3) *Beyond the counting bound.* Theorem 8.1 never gives more than $2^{O(r)}$, and [11] comes within a polylogarithmic factor in the exponent of that ceiling, in an elementary amenable group. Does some sofic group have a chunk with profile $2^{r^{1+c}}$ for some $c > 0$? A different lower-bound mechanism would be needed.

(4) *Solvable groups.* Cornulier's assertion that every solvable group has polynomial sofic profile [8, Example 3.16] rests on Theorem 3.15, and H_3 is not solvable. Is it true?

(5) *Products of Baumslag–Solitar groups.* For a product of k non-residually-finite Baumslag–Solitar groups, the product of the models of Theorem 8.6, each at accuracy $1/(kr)$, gives chunks with profile $O(r^k)$. This proves directly the upper bound $\sigma \preceq r^k$ that [8, Example 3.19] states for $BS(2,3)^k$ as a test case, and which follows from his Proposition 3.17, hence from Theorem 3.15. Is there a matching lower bound?

(6) *Extensions.* Proposition 7.1 bounds the profile of an extension by the Følner size of the quotient and the profile of a kernel chunk that grows with r through the cocycle of a section; the kernel chunk is fixed when the cocycle takes finitely many values. The groups G_0 and H_3 of Section 7 show that the kernel and the quotient alone do not determine the profile. Which data of an extension do? In particular, does the outer action of the quotient on the kernel?

Use of AI

This work was carried out with extensive use of generative AI systems. They proposed constructions and proof strategies, wrote and checked proofs, carried out computations, searched the literature, wrote the verification code, and drafted and edited the text. The authors directed the research, assessed and selected the results, and take full responsibility for the content. The proofs are given in full in the text; repeated adversarial checking and the accompanying programs support them but do not replace them. The verification notes state which statements are checked by machine. New sessions developed the whole-suffix sweep argument and its applications, and an independent blind check tested the proofs and finite certificates before this revision.

A Fillings for the far-commutator bound

Areas are taken in P; free conjugation and inversion do not change them, and freely $[x, yz] = [x, y] \cdot y[x, z]y^{-1}$. Write $\tau_j = b^j \alpha b^{-j}$, $t = b^{-1}a$, $\rho = a^{-1}b^{-1}ab$, $t_i = a^i t a^{-i}$, $X_n = b^{-n}a^n$, and

$$A(d) = \mathrm{Area}[\sigma_0, \sigma_d], \quad E(i) = \mathrm{Area}[\sigma_i, t], \quad U(j) = \mathrm{Area}(\tau_j \sigma_j^{-1}), \quad B(k) = \mathrm{Area}[\sigma_{-k}, b].$$

Since $b^{-n}(\tau_n \sigma_n^{-1})b^n = [\alpha, X_n]$ freely, $U(n) = \mathrm{Area}[\alpha, X_n]$. Since $[x^{-1}, y] = x^{-1}[y, x]x$, the signs in the definition of A^* do not matter and $A^*(M) = \max_{2 \le d \le M} A(d)$. Every relation below with a numerical area is certified in the ancillary files (Section 3).

Base relations. $[\alpha, t]$ is a rotation of r_5^{-1}; $aba^{-1} = \alpha b$ and $bab^{-1} = \alpha^{-1}a$ are rotations of $r_4^{\pm 1}$; $[\sigma_i, \sigma_{i+2}]$ is a conjugate of r_3; all have area 1. The relation $B_0 := \mathrm{Area}[\alpha, \sigma_1\sigma_0 b] \le 5$ follows by replacing $b\alpha^{-1}b^{-1}$ with σ_1^{-1} (one r_5), which leaves $\sigma_0\sigma_1\sigma_0\sigma_1^{-1}\sigma_0^{-1}\sigma_1^{-1}$, then three r_1 cells and one r_2 cell.

Lemma A.1 (Transfer and halving)**.** *For $i \ge 1$, $j \ge 1$, $k \ge 2$:*

(i) $E(i) \le 2i + 1 + \sum_{d=2}^{i+1} A(d)$, *and* $E(0) = 1$;

(ii) $U(j) \le \sum_{i=0}^{j-1} E(i)$;

(iii) $B(k) \le 2k + 5 + \sum_{d=2}^{k-1} A(d)$;

(iv) $A(k + j) \le A(k) + 2jB(k) + 2U(j)$.

Proof. (i) By $bab^{-1} = \alpha^{-1}a$, $t_k = t_{k+1}\sigma_{k-1}$ with one cell, so $t = t_i\sigma_{i-2}\cdots\sigma_0\sigma_{-1}$ with i cells, and i more for t^{-1}. Then $[\sigma_i, t_iS] = [\sigma_i, t_i]\cdot t_i[\sigma_i, S]t_i^{-1}$ with $S = \sigma_{i-2}\cdots\sigma_{-1}$; the first factor is $a^i[\alpha, t]a^{-i}$, and the second is a product of conjugates of $[\sigma_i, \sigma_l]$ for $l = i-2, \dots, -1$, of widths $2, \dots, i+1$. (ii) Freely $\tau_{i+1} = b\tau_i b^{-1}$; replace τ_i by σ_i at cost $U(i)$, and use that $b\sigma_i b^{-1}\sigma_{i+1}^{-1} = b[\sigma_i, t]b^{-1}$ freely. (iii) Freely $[\sigma_{-k}, b] = a^{-k}[\alpha, a^k b a^{-k}]a^k$, and $a^k b a^{-k} = \sigma_{k-1}\cdots\sigma_1\sigma_0 b$ with k cells (twice, for the element and its inverse); then move α past $\sigma_{k-1}, \dots, \sigma_2$ and finish with B_0. (iv) Conjugating by a^{-k} turns $[\sigma_0, \sigma_{k+j}]$ into $[\sigma_{-k}, \sigma_j]$. Replace $\sigma_j^{\pm 1}$ by $\tau_j^{\pm 1}$ ($2U(j)$), move $\sigma_{-k}^{\pm 1}$ across the $2j$ letters $b^{\pm 1}$ ($2jB(k)$), and what remains is $b^j[\sigma_{-k}, \alpha]b^{-j}$, a conjugate of $[\sigma_0, \sigma_k]$. □

The lemma gives, in this order: $A(2) = 1$, $E(0) = U(1) = 1$, $E(1) \le 4$, $U(2) \le 5$, $B(2) \le 9$, $A(3) \le 21$ ($k = 2, j = 1$), $A(4) \le 47$ ($k = j = 2$), $E(2) \le 27$, $U(3) \le 32$, $B(3) \le 12$, $A(5) \le 79$ ($k = 3, j = 2$), $A(6) \le 157$ ($k = j = 3$) and $E(3) \le 76$. Part (iv) with $k = 2$ gives $A(m) \le 2U(m-2) + 18m - 35$ for $m \ge 3$.

The cheap letters. We use $g^h = h^{-1}gh$. Freely, $\psi(t) = b^{-2}a^2 = t\rho t$, $t^b = b^{-2}ab = t\rho$, and $\psi(\rho) = \rho^a\rho^{ba}\rho\rho^b = t^{-1}\rho^b\rho^{b^2}t\cdot\rho\cdot\rho^b$ (from $[x^2, y^2] = (xcx^{-1})(xy\,c\,y^{-1}x^{-1})\,c\,(ycy^{-1})$ with $c = [x, y]$, $x = a^{-1}$, $y = b^{-1}$, and $a = bt$). The one relation needed is

$$\rho^b = R' := t\rho t^{-1}\cdot\rho\cdot t\rho t^{-1}, \qquad \text{area} \le 14.$$

Replace each ρ by α^{ba} (four r_4 cells); freely the relation becomes $\alpha^{bab} = \alpha^{b^2}\alpha^{ba}\alpha^{b^2}$, and after conjugating by $(bab)^{-1}$ it reads $\alpha = \alpha^g\alpha^{g'}\alpha^g$ with $g = ba^{-1}b^{-1}$ and $g' = bab^{-1}a^{-1}b^{-1}$. Six cells of $bab^{-1} = \alpha^{-1}a$ turn this into $\alpha = \alpha^{-1}\sigma_1\alpha\cdot b\alpha b^{-1}\cdot\alpha^{-1}\sigma_1\alpha$; one r_5 cell, two r_1 cells and one r_2 cell finish. Also $[\alpha, t]$ has area 1 and $[\alpha, \rho]$ area at most 5: $\rho\alpha^{-1}\rho^{-1}$ becomes $a^{-1}b^{-1}\sigma_2^{-1}ba$ after one r_5 cell, and $b^{-1}\sigma_2^{-1}b \to \sigma_1^{-1}$ costs $E(1) \le 4$. Consequently, for a word W in $t^{\pm 1}, \rho^{\pm 1}$, $W^b = W(t\rho, R')$ at area $14\,\#_\rho(W)$, and $\psi(W) = W(t\rho t, R)$ at area $84\,\#_\rho(W)$, where

$$R = t^{-1}\cdot R'\cdot R'(t\rho, R')\cdot t\cdot\rho\cdot R'$$

uses the relation six times per letter ρ; $|R'| = 7$ and $|R| = 46$.

Lemma A.2 (The doubling lifts)**.** $\psi(\alpha) = \Pi := \sigma_1\sigma_2\sigma_0\sigma_1$ *with* 6 *cells, and* $\mathrm{Area}\,\psi(r_i) \le 17,359,864,0,182$ *for* $i = 1, \dots, 5$. *Hence* $\mathrm{Area}\,\psi(w) \le 864\,\mathrm{Area}(w)$ *for every null word* w.

Proof. The identity for $[x^2, y^2]$ with $x = a$, $y = b$ and four r_4 cells ($[a, b] = \alpha$) give $\psi(\alpha) = \sigma_1\cdot a(b\alpha b^{-1})a^{-1}\cdot\alpha\cdot b\alpha b^{-1}$, and two r_5 cells give Π. $\psi(r_4)$ is freely trivial. $\psi(r_1) = \psi(\alpha)^2$: after two conversions, Π^2 costs five cells (σ_1^2, one r_3 swap, then $\sigma_2^2, \sigma_0^2, \sigma_1^2$). $\psi(r_5) = a^2\psi(\alpha)a^{-2}\cdot b^2\psi(\alpha)^{-1}b^{-2}$: after two conversions, $a^2\Pi a^{-2} = \sigma_3\sigma_4\sigma_2\sigma_3$ freely, and $b^2\sigma_i^{-1}b^{-2} \to \sigma_{i+2}^{-1}$

costs $E(i)+E(i+1)$; over the letters of Π this is $E(0)+3E(1)+3E(2)+E(3)\le 170$. $\psi(r_3)=[\psi(\alpha),a^4\psi(\alpha)a^{-4}]$: after four conversions it is $[\Pi,\sigma_5\sigma_6\sigma_4\sigma_5]$, which splits freely into sixteen conjugates of far commutators of widths $2,3,3,3,3,4,4,4,4,4,4,5,5,5,5,6$, costing 840. $\psi(r_2)=(\psi(\alpha)\,a^2\psi(\alpha)a^{-2})^3$: after six conversions it is the positive word $(\Pi\,\sigma_3\sigma_4\sigma_2\sigma_3)^3$ of length 24 in $\sigma_0,\dots,\sigma_4$, which represents the identity of $\mathrm{Sym}(6)$. Reduce it by the coset procedure: keep the word in the form $h\cdot c_k$ with $c_k=\sigma_4\sigma_3\cdots\sigma_k$ and h a word in $\sigma_0,\dots,\sigma_3$; absorbing a letter σ_i costs far commutations for $i\le k-2$, nothing for $i=k-1$, one cancellation for $i=k$, and far commutations plus one braid move for $i\ge k+1$, and then recurse on h. A braid move $\sigma_i\sigma_{i-1}\sigma_i=\sigma_{i-1}\sigma_i\sigma_{i-1}$ costs four cells (three r_1, one r_2). Charging each far commutation its actual width, the procedure uses 47 moves at total cost 323. The last claim follows by applying ψ to a product of conjugates of relators. □

Proof of Theorem 3.1. Define words in $t^{\pm1},\rho^{\pm1}$ by $W_1=t$, $W_{2n}=W_n(t\rho t,R)$ and $W_{2n+1}=W_{2n}(t\rho,R')\,t$, and let $\mathrm{Rw}(n)=\mathrm{Area}(X_nW_n^{-1})$. Then $\mathrm{Rw}(1)=0$. Since $X_{2n}=\psi(X_n)$ literally, Lemma A.2 and the cheap-letter relations give $\mathrm{Rw}(2n)\le 864\,\mathrm{Rw}(n)+84\,|W_n|$ and $|W_{2n}|\le 46|W_n|$. Since $X_{2n+1}=X_{2n}^b\,t$ freely, $\mathrm{Rw}(2n+1)\le\mathrm{Rw}(2n)+14\cdot 46|W_n|$ and $|W_{2n+1}|\le 7|W_{2n}|+1$. Per binary digit, $\mathrm{Rw}\le 864\,\mathrm{Rw}'+728|W'|$ and $|W|\le 323|W'|$, so with $s=\lfloor\log n\rfloor$,

$$|W_n|\le 323^s,\qquad \mathrm{Rw}(n)\le 728\sum_{i<s}864^{s-1-i}323^i\le\frac{728}{541}\,864^s\le 1.346\,n^{\log 864}.$$

Expanding $[\alpha,W_n]$ letter by letter, $U(n)\le 2\,\mathrm{Rw}(n)+5|W_n|\le 2.70\,n^{\log 864}+5\,n^{\log 323}$. For $m\ge 5$, $A(m)\le 2U(m-2)+18m-35\le 5.39\,m^{9.755}+10\,m^{8.336}+18m\le 0.87\,m^{11}$, the ratio being largest at $m=5$; and $A(2),A(3),A(4)\le 1,21,47$. □

Proof of Theorem 3.2. Put $q=\sigma_{-2}$ and add the auxiliary relation $r_6=[q,b]$. It has a nine-cell filling in P, with counts $(3,1,0,4,1)$ in the original five types: two r_4 cells give $a^2ba^{-2}=\sigma_1\alpha b$, and commuting α through this word costs one adjacent braid (three r_1 cells and one r_2 cell) and one r_5 cell. Replacing the word on both sides costs four r_4 cells in total. We keep r_6 atomic until the final expansion. All recursive certificates use only the types $I=(1,3,4,5,6)$, so no image of r_2 is needed.

Here are finite certificates and their derivation, with vectors ordered by I. For $i\in I$, $\mathbf{e}_i$ denotes the unit vector for relator type r_i (thus $\mathbf{e}_3$ is the second coordinate vector, not the third). Let F_d fill $[\alpha,\sigma_d]$, C_m certify $\sigma_m=\tau_m$, and J_j certify $b\sigma_jb^{-1}=\sigma_{j+1}$. Starting from $\mathbf{f}_2=\mathbf{e}_3$ and $\mathbf{c}_0=0$,

$$\mathbf{j}_j=2j\mathbf{e}_4+\mathbf{e}_5+\sum_{d=2}^{j+1}\mathbf{f}_d,\qquad \mathbf{c}_m=\mathbf{c}_{m-1}+\mathbf{j}_{m-1},\qquad \mathbf{f}_{m+2}=2\mathbf{c}_m+2m\mathbf{e}_6+\mathbf{e}_3.$$

Indeed, $b\sigma_jb^{-1}=Q_j\sigma_{j+1}Q_j^{-1}$ with $Q_j=\sigma_0^{-1}\cdots\sigma_{j-1}^{-1}$ costs $2j$ copies of r_4 and one of r_5; commute past Q_j using distances $2,\dots,j+1$. For the last recurrence conjugate to $[q,\sigma_m]$, replace σ_m by τ_m, and move q past the $2m$ letters $b^{\pm1}$ and then past α. This is noncircular. It gives

d	r_1	r_3	r_4	r_5	r_6
2	0	1	0	0	0
3	0	1	0	2	2
4	0	3	4	4	4
5	0	7	12	10	10
6	0	17	32	24	24

For $N_m=[\sigma_{-m},b]$, $\mathbf{n}_2=\mathbf{e}_6$ and $\mathbf{n}_m=\mathbf{n}_{m-1}+\mathbf{f}_{m-1}+2\mathbf{e}_4$: conjugate by a^{-1}, replace $a^{-1}ba=\sigma_{-1}^{-1}b$, and remove the first factor using F_{m-1}. Thus $\mathbf{n}_3=(0,1,2,0,1)$ and $\mathbf{n}_4=(0,2,4,2,3)$.

Use the free substitution $D(a) = a^2, D(b) = b^2, D(\alpha) = T = [a^2, b^2]$. The equality $T = S := \sigma_1\alpha\sigma_2\sigma_1^{-1}$ has vector $\mathbf{v} = (1, 1, 4, 2, 0)$: expand into four commutators $[a, b]$, use r_5 twice, commute σ_2 past α, and invert the final σ_1. The image counts are

$$\begin{aligned} \mathbf{d}_1 &= 2\mathbf{v} + \mathbf{e}_3 + 2\mathbf{e}_1, & \mathbf{d}_3 &= 4\mathbf{v} + \mathbf{f}_2 + 4\mathbf{f}_3 + 6\mathbf{f}_4 + 4\mathbf{f}_5 + \mathbf{f}_6, \\ \mathbf{d}_4 &= 0, & \mathbf{d}_5 &= 2\mathbf{v} + \mathbf{H}_0 + 2\mathbf{H}_1 + \mathbf{H}_2, \\ \mathbf{H}_j &= 2\mathbf{e}_4 + 2\mathbf{j}_j + \mathbf{f}_{j+2}, & \mathbf{d}_6 &= 2\mathbf{v} + 2(2\mathbf{n}_3 + \mathbf{n}_4 + \mathbf{n}_2). \end{aligned}$$

For $\mathbf{d}_1$, fill $(\alpha\sigma_2)^2$ inside its conjugator. For $\mathbf{d}_3$, the sixteen crossings between S and a^4Sa^{-4} have gap multiplicities $1, 4, 6, 4, 1$. The fourth image is free. For the fifth, $h_2 := b^{-2}a^2 = t^2q$ costs one r_4 cell, so $[h_2, \sigma_j]$ has vector $\mathbf{H}_j$, and $D(r_5) = b^2[h_2, T]b^{-2}$ freely. For the sixth, replace $D(q)$ by $\sigma_{-3}\sigma_{-4}q\sigma_{-3}^{-1}$ and commute its four factors through b^2. Consequently the substitution matrix is

$$M = \begin{pmatrix} 4 & 3 & 8 & 4 & 0 \\ 4 & 72 & 120 & 104 & 96 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 16 & 36 & 24 & 12 \\ 2 & 10 & 24 & 8 & 12 \end{pmatrix}.$$

Its rows are the counts of the chosen fillings of $D(r_i)$, not a substitution of the expanded nine-cell proof. For $w = (25, 616, 1, 133, 78)^T$, direct integer arithmetic gives

$$Mw = (2488, 65892, 0, 14070, 8234)^T \le 107w.$$

Every weight dominates its original expansion cost, including $w_6 \ge 9$.

Put $u_i = t^{-i}qt^i$. The identities $\sigma_{-3} = u_1$, $\sigma_{-4} = u_2$ cost one and seven auxiliary cells: use r_6 for the first, then replace σ_{-3} twice in N_3 and conjugate the first identity by a^{-1}. On the formal t, q alphabet define

$$d(t) = t^2q, \quad d(q) = u_1u_2qu_1^{-1}, \qquad E(t) = t^2qt^{-1}, \quad E(q) = q, \qquad o = E \circ d.$$

These certify $D(t) = d(t)$, $D(q) = d(q)$, $b^{-1}D(t)b = o(t)$ and $b^{-1}D(q)b = o(q)$. Their respective cell vectors and reduced formal lengths are

rule	cell vector	length
$d(t)$	$(0, 0, 1, 0, 0)$	3
$d(q)$	$(1, 2, 6, 2, 6)$	10
$o(t)$	$(0, 0, 3, 0, 1)$	7
$o(q)$	$(1, 2, 12, 2, 10)$	26

The even rules follow from $h_2 = t^2q$ and the block identity; E expresses conjugation by b^{-1}, giving the odd rules. All four weighted costs are at most 2315.

Set $W_1 = t$, $W_{2m} = d(W_m)$ and $W_{2m+1} = o(W_m)t$, freely reducing in the formal alphabet. Substitution of cells, followed by these rules, certifies $h_k = b^{-k}a^k = W_k$: freely $h_{2m} = D(h_m)$ and $h_{2m+1} = b^{-1}D(h_m)bt$. If A_k is its weighted cost and $n_k = |W_k|$, then, with $m = \lfloor k/2 \rfloor$,

$$A_k \le 107A_m + 2315n_m, \qquad n_k \le 26n_m + 1.$$

The t-exponent of W_m is m, so $n_m \ge 1$. Since $2315 + 29 \cdot 26 + 29 \le 107 \cdot 29$,

$$A_k + 29n_k \le 107(A_m + 29n_m) \le 29 \cdot 107^{\lfloor \log k \rfloor} \le 29k^{\log 107}.$$

Now $[t, \alpha]$ and $[q, \alpha]$ each cost one original cell. Replacing h_k twice and commuting its formal letters therefore fills $[h_k, \alpha]$ with at most $2A_k + n_k \le 58k^{\log 107}$ original cells. The free identity $\sigma_m\tau_m^{-1} = b^m[h_m, \alpha]b^{-m}$ and the preceding far-commutator reduction give, for $k = m + 2 \ge 3$,

$$A(k) \le 116m^{\log 107} + 18m + 1 \le 135m^{\log 107}.$$

Together with $A(2) = 1$, this proves the theorem. Literal certificates for every image and rule above, the bounded-distance inputs and the recursive identities through $k = 8$ are distributed in `certificates107.json.gz`; an independent verifier recomputes their target words, their original-relator expansions and their cell histograms. The all-depth conclusion follows from the displayed recurrence, whose scalar induction is also checked in Lean. □

Remark A.3 (Asymptotic refinement of this certificate scheme). Direct calculation gives $\det(xI - M) = x(x^4 - 112x^3 + 572x^2 + 3328x - 14208)$. The quartic changes sign on $(-6,-5),(3,4),(7,8),(106,107)$, so the five eigenvalues are distinct and the spectral radius is $\rho \in (106,107)$. Thus $\|M^j\| = O(\rho^j)$. Unweighted rule vectors are bounded by $(1,2,12,2,10)$, and $n_k \le 27 n_{\lfloor k/2 \rfloor}$. The resulting convolution is $O(\sum_{j<s} 27^j \rho^{s-1-j}) = O(\rho^s)$, since $\rho > 27$. Expanding auxiliary cells and commuting formal letters gives $A^*(M) = O(M^{\log \rho})$, where $\log \rho \approx 6.73252$. The explicit constants elsewhere use 107.

B From far commutators to the full Dehn function

We follow Lee's scheme [27, Section 4.2]. Lee reduced the Dehn function of H_3 to the areas of far commutators and of the words $\tau_j \sigma_j^{-1}$, which compare the conjugates of α by powers of b and of a, and bounded these areas exponentially [27, Lemma 4.7 and Remark 4.8]. We feed the polynomial bounds of Section 3 and Lemma A.1 for these words into a buffered form of his collection scheme.

Proof of Theorem 3.3. Fix an integer $n \ge 1$ and put $F_n = 1 + A^*(4n+2)$ and $U(0) = 0$. Lemma A.1 gives

$$E(i) \le 2F_n i + 1, \qquad U(q) \le \sum_{i=0}^{q-1} E(i) \le F_n q(q-1) + q \le F_n q^2 \quad (0 \le q \le 2n).$$

Here $F_n \ge 1$; all costs below count cells of the original five-relator presentation.

Buffered collection. Let w be a null word of length at most n and put $h = a^n b^n$. The two quotient coordinates while reading w after h stay in $[0,2n]^2$. Write the current word as $v\, s(m,q)$, where $s(m,q) = a^m b^q$ and v is a word in the $\sigma_j^{\pm 1}$. Using $bab^{-1} = \alpha^{-1}a$ exactly q times gives $b^q a b^{-q} = \tau_{q-1}^{-1} \cdots \tau_0^{-1} a$. Replace each τ_j by σ_j at cost $U(j)$. Thus

$$s(m,q)a = (\sigma_{m+q-1}^{-1} \cdots \sigma_m^{-1}) s(m+1,q),$$
$$s(m,q)a^{-1} = (\sigma_{m-1} \cdots \sigma_{m+q-2}) s(m-1,q)$$

each costs at most $q + \sum_{j<q} U(j) \le 2n + (8/3)F_n n^3 \le 5F_n n^3$; an empty product is omitted. The second rule is the inverse rearrangement of the first at $(m-1,q)$, and is used only when the next coordinate is nonnegative. The rules $s(m,q)\alpha^{\pm 1} = \sigma_{m+q}^{\pm 1} s(m,q)$ cost $U(q)$, while multiplication by $b^{\pm 1}$ merely changes q and costs no cells. The buffer ensures that all indices produced are nonnegative.

After at most n rules the quotient coordinates return to (n,n), and hence hwh^{-1} has been changed to a word v at cost at most $5F_n n^4$ (the $\alpha^{\pm 1}$ rule costs at most $4F_n n^2$ per letter). It has at most $2n^2 + n \le 3n^2$ letters, with $0 \le j \le 4n$. Since $w = 1$, the word v is the identity permutation on the first $4n+2$ points of ray 1. It remains to fill this permutation word with a cost that is uniform in the size of that symmetric group.

For example, take the null word $w = aba^{-1}b^{-1}\alpha^{-1}$, $n = 5$ and $h = a^5b^5$. The complete collection trace is

letter	(m, q) after the step	letters appended to v	cell cost at most
a	$(6, 5)$	$\sigma_9^{-1} \cdots \sigma_5^{-1}$	$5 + \sum_{j<5} U(j)$
b	$(6, 6)$	1	0
a^{-1}	$(5, 6)$	$\sigma_5 \cdots \sigma_{10}$	$6 + \sum_{j<6} U(j)$
b^{-1}	$(5, 5)$	1	0
α^{-1}	$(5, 5)$	σ_{10}^{-1}	$U(5)$

The resulting v freely cancels. Its collection costs at most $11 + 110F_5$. This trace illustrates the inverse-letter rules; general null words require the insertion argument below.

Coxeter insertion. In S_{d+1}, with adjacent transpositions $s_0, \dots, s_{d-1}$, use the recursive normal form uc_k, where $u \in S_d$, $c_k = s_{d-1}s_{d-2} \cdots s_k$ for $0 \le k < d$, and $c_d = 1$. With right actions c_k sends the last point to $k + 1$; this determines k, after which u is determined recursively. To insert a generator on the right use

$$c_k s_i = \begin{cases} s_i c_k, & i \le k - 2, \\ c_{k-1}, & i = k - 1, \\ c_{k+1}, & i = k, \\ s_{i-1} c_k, & i \ge k + 1. \end{cases}$$

The first case uses far commutations, the middle cases concatenate or cancel, and the last case moves the final s_i left across $s_k, \dots, s_{i-2}$, applies one braid, and moves the resulting s_{i-1} left across $s_{i+1}, \dots, s_{d-1}$. There are at most $d + 1$ Coxeter moves at a level and at most d levels per insertion, since only one generator is passed into u. Consequently a word of length l reduces to its unique normal form in at most $ld(d + 1)$ moves, and the identity reduces to the empty word.

Apply this with $d = 4n + 1$ to v. First replace inverse letters by positive ones using $\sigma_j^2 = 1$, at one cell per letter. A Coxeter square costs one cell in P, a braid costs at most four, and a far commutation costs at most $A^*(4n + 2)$. Each move therefore costs at most $4F_n$. Since $d \le 5n$ and $d + 1 \le 6n$, insertion costs at most $3n^2 \cdot 30n^2 \cdot 4F_n = 360F_n n^4$. Inverse conversion costs at most $3n^2 \le 3F_n n^4$. Adding collection gives at most $(5 + 3 + 360)F_n n^4 \le 400F_n n^4$. Invariance of area under free conjugation proves the asserted bound for $\delta_P(n)$. The identification of null words with identity ray permutations uses Johnson's presentation, not the finite regression tests. Substituting Theorem 3.2, or Theorem 3.1, gives the two polynomial exponents. □

References

[1] V. Alekseev and H. Bradford, Sofic actions, halo products, and metric approximations of groups, arXiv:2601.18742 (2026).

[2] G. Arzhantseva and P.-A. Cherix, Quantifying metric approximations of discrete groups, Ann. Fac. Sci. Toulouse Math. **33** (2024), 831–877; arXiv:2008.12954.

[3] C. Birkar, Singularities of linear systems and boundedness of Fano varieties, Ann. of Math. (2) **193** (2021), 347–405.

[4] H. Bradford, Controlling LEF growth in some group extensions, arXiv:2201.05052 (2022).

[5] K. S. Brown, Finiteness properties of groups, J. Pure Appl. Algebra **44** (1987), 45–75.

[6] J. Burillo, S. Cleary, A. Martino and C. E. Röver, Commensurations and metric properties of Houghton's groups, Pacific J. Math. **285** (2016), 289–301; arXiv:1403.0026.

[7] P. Burton, M. Chaudkhari, K. Juschenko and K. Muliarchyk, Hyperlinear approximations to amenable groups come from sofic approximations, arXiv:2311.09202 (2023).

[8] Y. Cornulier, Sofic profile and computability of Cremona groups, Michigan Math. J. **62**(4) (2013), 823–841; doi:10.1307/mmj/1387226167.

[9] Y. de Cornulier, L. Guyot and W. Pitsch, On the isolated points in the space of groups, J. Algebra **307** (2007), 254–277; arXiv:math/0511714.

[10] M. De Chiffre, N. Ozawa and A. Thom, Operator algebraic approach to inverse and stability theorems for amenable groups, Mathematika **65** (2019), 98–118; arXiv:1706.04544.

[11] S. Douglas and N. Mghirbi, Amenable groups with nearly exponential sofic profile, and quantum channels that need nearly linear memory, version 1.0, Zenodo (2026); doi:10.5281/zenodo.23050303.

[12] S. Douglas and N. Mghirbi, Causal quantum-channel simulation: memory beyond entropy, version 1.1, Zenodo (2026); doi:10.5281/zenodo.23048998.

[13] A. Duncan, B. Heath and C. Urech, Representations of finite subgroups of Cremona groups, arXiv:2507.04474 (2025).

[14] G. Elek and E. Szabó, On sofic groups, J. Group Theory **9** (2006), 161–171; arXiv:math/0305352.

[15] D. S. Farley and B. Hughes, Braided diagram groups and local similarity groups, in: Geometric and Cohomological Group Theory, London Math. Soc. Lecture Note Ser. **444**, Cambridge Univ. Press, 2017, 15–33; arXiv:1406.4587.

[16] F. Fournier-Facio, A torsion-free non-sofic group, arXiv:2608.02025 (2026).

[17] A. Genevois, Embeddings into Thompson's groups from quasi-median geometry, Groups Geom. Dyn. **13** (2019), 1457–1510; arXiv:1709.03888.

[18] F. Glebe, A cup product obstruction to Frobenius stability, Münster J. Math. **18** (2025), 27–38; https://www.uni-muenster.de/FB10/mjm/vol_18/mjm_vol_18_02.pdf.

[19] W. T. Gowers and O. Hatami, Inverse and stability theorems for approximate representations of finite groups, Sb. Math. **208** (2017), 1784–1817; arXiv:1510.04085.

[20] V. S. Guba, Polynomial isoperimetric inequalities for Richard Thompson's groups F, T, and V, in: Algorithmic Problems in Groups and Semigroups (Lincoln, NE, 1998), Trends Math., Birkhäuser, Boston, 2000, 91–120.

[21] V. S. Guba, The Dehn function of Richard Thompson's group F is quadratic, Invent. Math. **163** (2006), 313–342.

[22] V. Guba and M. Sapir, Diagram groups, Mem. Amer. Math. Soc. **130** (1997), no. 620.

[23] A. Guld, Finite subgroups of the birational automorphism group are 'almost' nilpotent of class at most two, arXiv:2004.11715 (2020).

[24] C. H. Houghton, The first cohomology of a group with permutation module coefficients, Arch. Math. **31** (1978), 254–258.

[25] D. L. Johnson, Embedding some recursively presented groups, in: Groups St Andrews 1997 in Bath, II, London Math. Soc. Lecture Note Ser. **261**, Cambridge Univ. Press, 1999, 410–416.

[26] G. Kun and A. Thom, Nonsofic wreath products of residually finite groups, arXiv:2608.06222 (2026).

[27] S. R. Lee, Geometry of Houghton's groups, PhD thesis, University of Oklahoma (2012); arXiv:1212.0257v1. Numbered references use this arXiv version.

[28] E. H. Lieb and M. B. Ruskai, Proof of the strong subadditivity of quantum-mechanical entropy, J. Math. Phys. **14** (1973), 1938–1941.

[29] R. C. Lyndon and P. E. Schupp, Combinatorial Group Theory, Springer, Berlin, 1977.

[30] S. Meskin, Nonresidually finite one-relator groups, Trans. Amer. Math. Soc. **164** (1972), 105–114.

[31] M. Migliorini, Thompson's group T has quadratic Dehn function, Forum Math. Sigma **13** (2025), e109; arXiv:2410.23088.

[32] A. Yu. Ol'shanskii, Hyperbolicity of groups with subquadratic isoperimetric inequality, Internat. J. Algebra Comput. **1** (1991), 281–289.

[33] OpenAI, Non-sofic groups exist, `https://openai.com/index/ten-advances-in-mathematics/` (2026).

[34] Yu. Prokhorov and C. Shramov, Jordan property for Cremona groups, Amer. J. Math. **138** (2016), 403–418; arXiv:1211.3563.

[35] C. E. Röver, Subgroups of finitely presented simple groups, PhD thesis, University of Oxford (1999).

[36] W. Slofstra, A group with at least subexponential hyperlinear profile, arXiv:1806.05267v1 (2018).

[37] W. Slofstra and T. Vidick, Entanglement in non-local games and the hyperlinear profile of groups, Ann. Henri Poincaré **19** (2018), 2979–3005; arXiv:1711.10676.

[38] A. M. Vershik and E. I. Gordon, Groups that are locally embeddable in the class of finite groups, Algebra i Analiz **9**(1) (1997), 71–97; English transl. St. Petersburg Math. J. **9**(1) (1998), 49–67.

[39] H. Bradford, Quantifying local embeddings into finite groups, arXiv:2104.07111v2 (2021).

[40] M. Chaudkhari, K. Juschenko and K. Muliarchyk, Thompson's groups are sofic, arXiv:2610.25415v1 (2026).