The dynamics and geometry of free group endomorphisms

Jean Pierre Mutanguha*

Abstract

We prove that ascending HNN extensions of free groups are word-hyperbolic if and only if they have no Baumslag-Solitar subgroups. This extends the theorem of Brinkmann that free-by-cyclic groups are word-hyperbolic if and only if they have no free abelian subgroups of rank 2. The paper is split into two independent parts:

1) We study the dynamics of injective nonsurjective endomorphisms of free groups. We prove a canonical structure theorem that initializes the development of improved relative train tracks for endomorphisms; this structure theorem is of independent interest since it makes many open questions about injective endomorphisms tractable.

2) As an application of the structure theorem, we are able to (relatively) combine Brinkmann's theorem with our previous work and obtain the main result stated above. In the final section, we further extend the result to HNN extensions of free groups over free factors.

Overview

Word-hyperbolic groups are groups that act geometrically (properly and cocompactly) on proper δ-hyperbolic spaces and these spaces are the coarse equivalents of negatively curved geometric spaces. Word-hyperbolic groups are an important class of groups introduced by Misha Gromov [15] and one of the fundamental problems in geometric group theory is determining necessary and sufficient algebraic conditions for a group to be word-hyperbolic.

Question (Coarse hyperbolization [1, Question 1.1]). Let G be a group of finite type. If G contains no Baumslag-Solitar subgroups, then must it be word-hyperbolic?

Finite type can be thought of as a strengthening of finite presentation for torsion-free groups. Baumslag-Solitar groups are two-generator one-relator groups with the presentation $BS(m, n) = \langle a, t \mid t^{-1}a^m t = a^n \rangle$ for $m, n \neq 0$. It is now a classical fact that a word-hyperbolic group cannot have subgroups isomorphic to Baumslag-Solitar groups. The question, historically attributed to Gromov, asks if Baumslag-Solitar subgroups are the only essential obstruction to word-hyperbolicity for finite type groups. It has been

---
*Email: jpmutang@uark.edu, *Web address:* https://mutanguha.com
Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR

answered in the affirmative for certain classes of groups: the fundamental groups of closed 3-manifolds (Thurston [31, 32], Perelman) and free-by-cyclic groups (Brinkmann [8]). The main result of this paper extends the latter to a larger class of HNN extensions.

Main Theorem. *Let $A \leq F$ be a free factor of a finite rank free group F and $\phi : A \to F$ be an injective homomorphism. The HNN extension $F*_A$ of F over A and ϕ is word-hyperbolic if and only if it contains no $BS(1, n)$ subgroups for $n \geq 1$.*

The HNN extension has the presentation $F*_A = \langle F, t \mid t^{-1}at = \phi(a), \ \forall a \in A \rangle$. If $A = F = \phi(A)$, then $F*_A = F \rtimes_\phi \mathbb{Z}$ is a free-by-cyclic group and this case was proven by Peter Brinkmann. Our proof of the main theorem above uses Brinkmann's result. If $A = F \neq \phi(A)$, then $F*_A = F*_\phi$ is a strictly ascending HNN extension and coarse hyperbolization for this class of groups was our original motivation. Finally, when $A \neq F$, we more or less reduce this to the case $A = F$.

The case $A = F$ of the main theorem (Theorem 6.7) proved to be difficult since it required an understanding of the dynamics of injective nonsurjective endomorphisms of free groups and no such complete study had been carried out yet. Patrick Reynolds studied the dynamics of irreducible nonsurjective endomorphisms of free groups [28] and we previously used this work to prove an instance of the main theorem where $\phi : F \to F$ is irreducible but not surjective [24, 25]. In this case, we showed that the strictly ascending HNN extension $F*_\phi$ is always word-hyperbolic. Unfortunately, it remained unclear how Reynolds' work could be generalized to all injective nonsurjective endomorphisms and no further progress on the problem had been made.

The first part of this paper (Sections 1–4) is a self-contained systematic study of all injective nonsurjective endomorphisms of free groups and is the main novel contribution of the paper. We hereby present a summary of the important results from these sections. Note that the statements given here are not the complete statements of the cited results.

Summary. *If $\phi : F \to F$ is injective but not surjective, then there is:*

1. *a unique maximal proper $[\phi]$-fixed free factor system $\mathcal{A}$;*   *(Proposition 2.3)*

2. *a unique minimal proper $[\phi]$-invariant free factor system $\mathcal{A}^*$ that carries $\mathcal{A}$ and is fixed under backward iteration, i.e. $\phi^{-1} \cdot \mathcal{A}^* = \mathcal{A}^*$;*   *(Proposition 2.4)*

3. *a unique expanding $\mathcal{A}^*$-relative immersion for ϕ.*   *(Corollary 4.7)*

The free factor systems $\mathcal{A}$ and $\mathcal{A}^*$ could be trivial.

This means any injective nonsurjective endomorphism $\phi : F \to F$ is induced by a graph map $f : \Gamma \to \Gamma$ such that: 1) some possibly empty or disconnected proper subgraph $\Gamma' \subset \Gamma$ is f-invariant and the restriction of f to Γ' is a homotopy equivalence; 2) some f-invariant proper subgraph $\Gamma'' \supset \Gamma'$ has an f^n-image contained in Γ' for some $n \geq 1$; and

3) collapsing the lifts of the subgraph Γ'' in the universal cover $\tilde{\Gamma}$ induces an *expanding immersion* $\bar{f} : T \to T$ on a simplicial F-tree. In particular, if Γ' is empty, then f is an expanding graph immersion.

In our previous work [24], we extended a result of Ilya Kapovich [18] and proved the special case of the main theorem with $A = F$ and $\phi : F \to F$ induced by an expanding graph immersion. The expanding immersion was crucial to the proof and so we developed expanding relative immersions with the intent of, in a sense, relativizing our previous proof and extending it to all injective nonsurjective endomorphisms. This application of the structure theorem is carried out in the second part of the paper (Sections 5–7).

Expanding relative immersions are interesting in their own right as they can potentially be applied to many currently open problems about nonsurjective endomorphisms. To name a few: relative immersions may be needed to extend Hagen-Wise's cubulation of word-hyperbolic free-by-cyclic groups [17, 16] to the groups $F*_A$; we suspect that they could be used to extend the main theorem to a characterization of the possible Dehn/isoperimetric functions of the groups $F*_A$ following work by Bridson-Groves [7] (See also Problem 2); they provide a framework for generalizing Bestvina-Feighn-Handel's construction of improved relative train tracks for automorphisms [5], which were used in Brinkmann's, Hagen-Wise's, and Bridson-Groves' results; for a different direction of generalization, the main theorem may be extended to HNN extensions $G*_A$ of torsion-free word-hyperbolic groups G over free factors $A \leq G$; and finally, the *virtual fibering* question is a particularly interesting problem that was first posed to us by Dawid Kielak and our structure theorem along with the cubulation problem could be significant steps towards its resolution:

Question (Virtual fibering)**.** Suppose $\phi : F \to F$ is injective but not surjective. If $F*_\phi$ is word-hyperbolic, then does it have a free-by-cyclic finite index subgroup? What if it has a quadratic Dehn function?

Besides Proposition 4.6 (uniqueness of expanding relative immersion) and Section 7 (case $A \neq F$), the results in this paper were the author's doctoral thesis. With future applications of the structure theorem in mind, we emphasize in this paper the independence of the "theory" (first part) and "application" (second part). The two parts can be read separately. If you choose to skip the first part, then we encourage you to read the prologue and interlude; these optional sections contain several examples that capture the key points. The main results of the first part are summarized again in the interlude with a bit more detail since we are assuming you will, by then, be familiar with the terms defined in the definitions and conventions section. The summary is all you will need to use expanding relative immersions in your own work. The epilogue is a brief discussion of some questions about endomorphisms in the context of expanding relative immersions.

The main tools used in the first part are Bass-Serre theory, Stallings graphs, bounded cancellation, and some Bestvina-Handel train track theory. The second part uses the Bestvina-Feighn combination theorem and some Bass-Serre theory.

Contents

Prologue

We will start with a few motivating examples to illustrate the main construction of the first part of this paper. Let $F = F(a,b)$ be the free group on two generators and $\phi, \varphi, \psi : F \to F$ be injective nonsurjective endomorphisms given by

$$\phi : (a,b) \mapsto (ab, ba), \quad \varphi : (a,b) \mapsto (a, bab^{-1}), \quad \text{and} \quad \psi : (a,b) \mapsto (a, abab).$$

The *standard rose* is a rose R with two petals and an identification $F = \pi_1(R)$ such that the basis $\{a,b\}$ corresponds to the petals; it will be graphically represented by a rose with two oriented petals labelled by a and b respectively.

For each integer $k \geq 1$, let $\hat{R}_k$ be the cover of R corresponding to the subgroup $\phi^k(F)$, i.e. it is the quotient of the universal cover $\tilde{R}$ by the action of the subgroup $\phi^k(F)$. The Stallings graph $S(\phi^k(F))$ is the core of $\hat{R}_k$, i.e. the smallest connected subgraph of $\hat{R}_k$ with rank 2. Alternatively, we can define $\tilde{R}(\phi^k(F)) \subset \tilde{R}$ to be the smallest subtree invariant
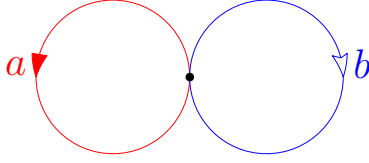
4

Figure 1: The standard rose.

under the $\phi^k(F)$-action and $S(\phi^k(F))$ to be the quotient $\phi^k(F)\backslash\tilde{R}(\phi^k(F))$. Evidently, there is a natural isomorphism $\phi^k(F) \cong \pi_1(S(\phi^k(F)))$. A Stallings graph S is graphically represented as an R-digraph, i.e. S will be an oriented graph whose oriented edges are labelled by a or b as shown in Figure 2.
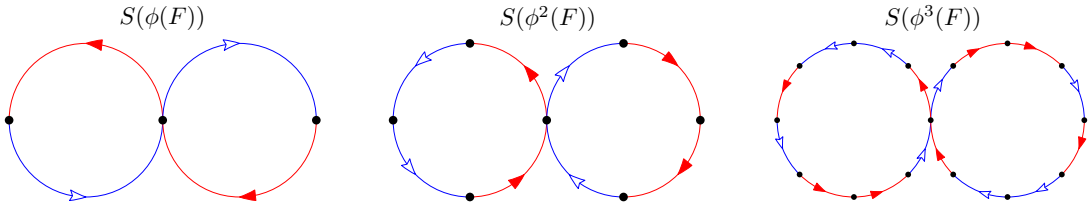

Figure 2: Stallings graphs $S(\phi^k(F))$ with respect to the standard rose for $k = 1, 2, 3$.

The Stallings graphs $S(\phi^k(F))$ are roses for all $k \geq 1$ and each petal doubles in size with each iteration. Since ϕ is injective, it is an isomorphism onto its image. In particular, we may use ϕ to get an isomorphism $F \cong \phi^k(F) \cong \pi_1(S(\phi^k(F)))$ for $k \geq 1$. Under this isomorphism, the basis $\{a, b\}$ corresponds to the petals of $S(\phi^k(F))$ and we recover the standard rose for all $k \geq 1$. This is equivalent to the existence of an immersion (locally injective map) on the rose $f : R \to R$ such that $\pi_1(f) = \phi$ [25, Lemma 3.2]. Indeed, the obvious map $f : R \to R$ that maps the a-petal to the path ab and the b-petal to the path ba is an immersion. The petals of $S(\phi^k(F))$ doubling in size with each iteration implies f is in fact an expanding immersion, i.e. all edges expand under f-iteration.

The Stallings graphs $S(\varphi^k(F))$ are barbells for all $k \geq 1$, the middle bars roughly double in size with each iteration, and the *plates* are labelled by a (See Figure 3). Under the isomorphism $F \cong \varphi^k(F) \cong \pi_1(S(\varphi^k(F)))$, the conjugacy classes $\{[a], [b]\}$ correspond to the two plates of the barbell $S(\varphi^k(F))$ and we recover the so called *standard barbell* for all $k \geq 1$. Just as in the first example, this is equivalent to the existence of an immersion on the barbell $g : B \to B$ such that $\pi_1(g) = \varphi$ under some identification $F = \pi_1(B)$. However, the immersion is not expanding since the plates are not growing in size with each iteration.

What we are now about to do will be the crux of the first part of the paper. Since the (nongrowing) plates of $S(\varphi^k(F))$ are all labelled a, we deduce that $\langle a \rangle$ is a free factor fixed by φ and $\langle b \rangle$ is mapped to a conjugate of $\langle a \rangle$. Thus, $\{\langle a \rangle, \langle b \rangle\}$ forms an *invariant free factor system* that contains a *fixed free factor*. Now consider $\tilde{g} : \tilde{B} \to \tilde{B}$ to be the lift of g
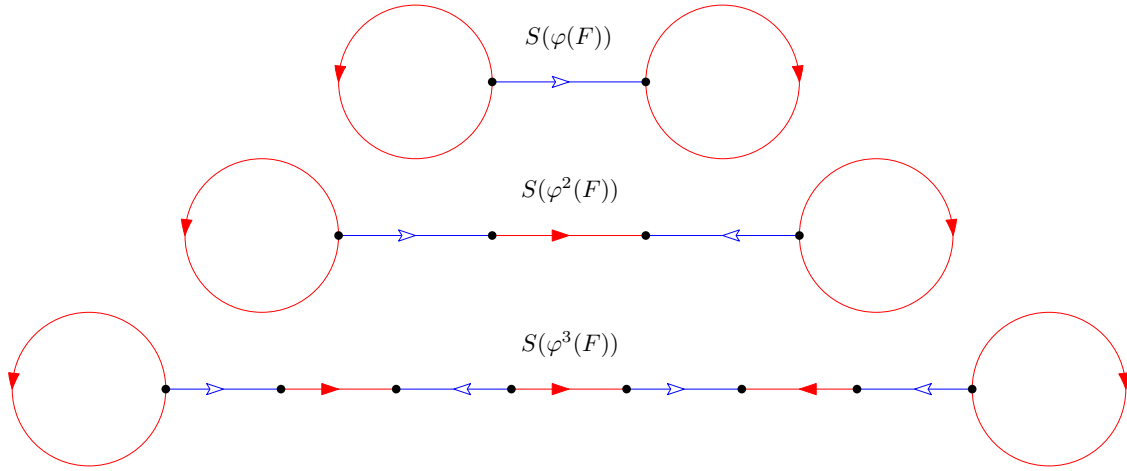
5

Figure 3: Stallings graphs $S(\varphi^k(F))$ for $k = 1, 2, 3$.

to the universal cover $\tilde{B}$. Collapsing all translates of axes of a and b in $\tilde{B}$, i.e. collapsing all edges of $\tilde{B}$ labelled by the plates of B, will produce a so-called $(F, \{\langle a \rangle, \langle b \rangle\})$-tree T with a nonfree F-action and $\tilde{g}$ induces an expanding immersion $\bar{g} : T \to T$ where each edge doubles in size under $\bar{g}$-iteration. The map $\bar{g}$ will be referred to as a *relative immersion*.
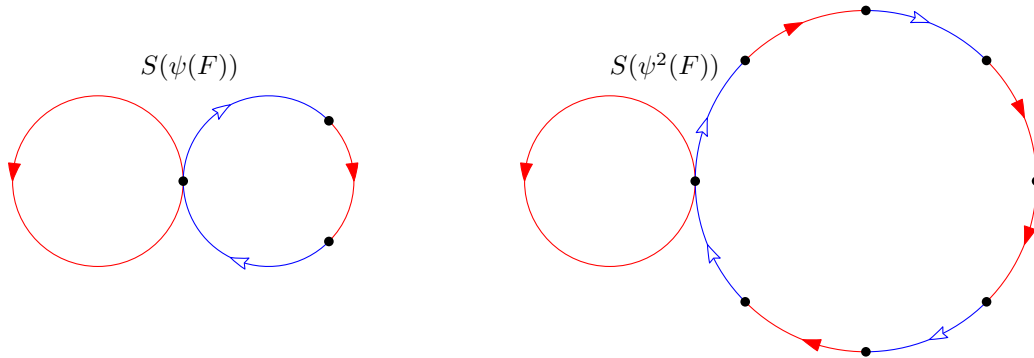
Figure 4: Stallings graphs $S(\psi^k(F))$ for $k = 1, 2$.

For the final example, the Stallings graphs $S(\psi^k(F))$ are roses again where one petal roughly doubles in size with each iteration and another is labelled by a (See Figure 4). Unlike the first example, the roses with the isomorphisms $F \cong \psi^k(F) \cong \pi_1(S(\psi^k(F)))$ are not standard roses since b does not correspond to a petal for $k \geq 1$. In fact, applying ψ to the labels of $S(\psi^k(F))$ forces a full fold with one petal to get $S(\psi^{k+1}(F))$ which means the Stallings graphs are all distinct *marked roses*. In particular, ψ cannot be induced by a graph immersion! But the observation that the nongrowing petals of $S(\psi^k(F))$ are

all labelled by a implies $\langle a \rangle$ is a fixed free factor. Despite ψ not being induced by an immersion, the obvious map on the standard rose $h : R \to R$ will induce an expanding immersion $\bar{h} : Y \to Y$ on an $(F, \{\langle a \rangle\})$-tree Y where every edge doubles in size.

The main result of the first part of this paper is that this construction always produces expanding relative immersions, Theorem 4.5: any injective nonsurjective endomorphism of a free group is induced by a graph map $f : \Gamma \to \Gamma$ such that collapsing the translates of axes in $\tilde{\Gamma}$ of some canonical invariant free factor system containing a fixed free factor system will induce an expanding relative immersion $\bar{f} : T \to T$. This can be summarized into two steps: first establishing the existence of a unique (possibly trivial) maximal fixed free factor system (Section 2); then showing that collapsing this free factor system and its preimages in the universal cover of an appropriately chosen graph Γ will produce a tree T on which we can define an expanding relative immersion (Section 4). The two guiding principles will be *bounded cancellation* (Lemmas 1.4 & 3.1) and the fact that Stallings graphs for iterated injective nonsurjective endomorphisms have unbounded size (Lemma 1.1 & Proposition 3.9).

Definitions and conventions

F will always be a free group with finite rank at least 2. A nontrivial subgroup system of F is a nonempty finite collection of nontrivial subgroups $\mathcal{A} = \{A_1, \dots, A_l\}$ of F. A nontrivial free factor system of F is a nonempty collection of nontrivial free factors $\mathcal{A} = \{A_1, \dots, A_l\}$ of F such that $A_i \cap A_j$ is trivial if $i \neq j$ and $\langle A_1, \dots, A_l \rangle$ is free factor of F. We define the trivial system to be the collection consisting of the trivial subgroup; this will allow us to treat *absolute* and *relative* cases simultaneously in our proofs. For any subgroup system $\mathcal{A}$, the subgroups $A_i \in \mathcal{A}$ are called the components of $\mathcal{A}$. A subgroup system $\mathcal{A}$ is finitely generated if its components are finitely generated. A free factor system $\mathcal{A}$ is proper if some component is proper, i.e. $\mathcal{A} \neq \{F\}$. Let $\mathcal{A}$ and $\mathcal{B}$ be subgroup systems of F. We shall say $\mathcal{A}$ carries $\mathcal{B}$ if for every component $B \in \mathcal{B}$, there is a component $A \in \mathcal{A}$ and element $x \in F$ such that $B \leq xAx^{-1}$ and, when $\mathcal{A}$ and $\mathcal{B}$ are free factor systems, we shall denoted it by $\mathcal{B} \preceq \mathcal{A}$. For a conjugacy class of elements $[g]$ in F, we shall also say $\mathcal{A}$ carries $[g]$ if there is a component $A \in \mathcal{A}$ and element $x \in F$ such that $g \in xAx^{-1}$. If $\mathcal{A} = \{A\}$ (or $\mathcal{B} = \{B\}$) is a singleton, then we will write, "A carries $[g] / \mathcal{B}$ (or B)." One can easily verify that the $\preceq$-relation on free factor systems is a preorder, i.e. it is reflexive and transitive. So it determines an equivalence relation on the set of free factor systems and a partial order on the set of equivalence classes. Free factor systems will always be considered up to this equivalence relation. In particular, we can replace components in a system with conjugates whenever convenient.

Remark. Suppose $F = \langle a, b \rangle$ be the free group of rank 2. With our definition, $\{\langle a \rangle, \langle b \rangle\}$ and $\{\langle bab^{-1} \rangle, \langle (ba)b(ba)^{-1} \rangle\}$ are equivalent free factor systems but $\{\langle a \rangle, \langle (ba)b(ba)^{-1} \rangle\}$

is not a free factor system. Thus, when we replace components in a free factor system with conjugates, we need to ensure the resulting subgroup system is still a free factor system.

Two group homomorphisms $h_1, h_2 : A \to B$ are equivalent if there is an inner automorphism $i_b : B \to B$ such that $h_2 = i_b \circ h_1$ and this is denoted by $[h_1] = [h_2]$. Outer endomorphisms of F are the equivalence classes on the set of endomorphisms of F. Typical properties such as being irreducible are not just properties of an endomorphism ϕ but also its outer class $[\phi]$.

Let $\phi : F \to F$ be an endomorphism. We say a subgroup system $\mathcal{A} = \{A_1, \ldots, A_l\}$ is [ϕ]-invariant if $\mathcal{A}$ carries the subgroup system $\phi(\mathcal{A}) = \{\phi(A_1), \ldots, \phi(A_l)\}$, i.e. there exists a set of elements $\{x_1, \ldots, x_l\} \subset F$ and a function $\sigma : \{1, \ldots, l\} \to \{1, \ldots, l\}$ such that $\phi(A_i) \leq x_i A_{\sigma(i)} x_i^{-1}$ for all i. A [ϕ]-invariant subgroup system $\mathcal{A}$ is [ϕ]-fixed if its components are permuted up to conjugacy, i.e. σ is a permutation and $[\phi(A_i)] = [A_{\sigma(i)}]$ for all i. The endomorphism ϕ is reducible if it has a nontrivial proper invariant free factor system and irreducible otherwise. One immediate consequence of Stallings folds is the injectivity of irreducible endomorphisms (Observation below). So we can drop the injectivity hypothesis when specializing results to the irreducible case.

Remark. When ϕ is an automorphism, all [ϕ]-invariant free factor systems are [ϕ]-fixed. However, for endomorphisms in general, being [ϕ]-invariant is weaker than being [ϕ]-fixed.

When ϕ is injective, then, for any $k \geq 1$ and free factor system $\mathcal{A}$ of F, $\phi^k(\mathcal{A})$ is a free factor system of $\phi^k(F)$; conversely, any free factor system of $\phi^k(F)$ gives us a free factor system of F via the isomorphism $\phi^k : F \to \phi^k(F)$. Finally, note that when $\mathcal{A}$ is a [ϕ]-fixed free factor system but ϕ is not surjective, $\phi^k(\mathcal{A})$ will typically not be a free factor system of F (See first remark). A subgroup $A \leq F$ is eventually [ϕ]-periodic if $[\phi^m(A)] = [\phi^n(A)]$ for some $m > n \geq 1$; it is [ϕ]-periodic if $[\phi^m(A)] = [A]$ for some $m \geq 1$.

For the topological perspective, graphs are 1-dimensional CW-complexes and a graph map $f : X \to X'$ will be a continuous map of graphs that sends vertices to vertices and any edge to a vertex or immersed path. An edge e of X is pretrivial if $f(e)$ is a vertex. For a graph map $f : \Gamma \to \Gamma'$ of finite graphs, let K be the maximum of the (combinatorial) length of the edge-path $f(e)$ as e varies over all the edges of Γ. Then f is K-Lipschitz, a fact that will be used throughout the paper. Generally, $K(f)$ will denote some convenient Lipschitz constant for f rather than the infimum. For the moment, X is used for arbitrary graphs but, for most of the paper, Γ will be used for finite connected noncontractible graphs and T for infinite simply-connected graphs (simplicial trees). A core graph is a graph with no proper deformation retract.

A direction at a vertex $v \in X$ is a *half-edge* attached to the vertex. A vertex is bivalent if it has exactly two directions. The set of directions at vertex v is denoted by $T_v X$. Branch points are vertices with at least three directions and natural edges are maximal edge-paths whose interior vertices are bivalent. We will say a graph map is natural if it maps branch points to branch points and any natural edge to a branch point or immersed path. If the graph map $f : X \to X'$ has no pretrivial edges, then the restriction to initial

segments induces the *derivative map* at v, $df_v : T_v X \to T_{f(v)} X'$. The graph map f is an immersion if it is locally injective, i.e. it has no pretrivial edges and the derivative maps df_v are injective for all vertices v; note that immersions are natural. An immersion f is expanding if the length of $f^n(e)$ is unbounded as $n \to \infty$ for every edge e of X.

With immersions defined, we preface the observation with a summary of Stallings' *folding theorem*. Let R be a rose whose edges are indexed by a basis $\{a_1, \ldots, a_n\}$ of F and let $f : R \to R$ be the map where $f(a_i)$ is the immersed edge-path in R labelled by $\phi(a_i)$. Stallings [30] showed that f factors as $\iota \circ f_l \circ \cdots \circ f_1$ where each f_i is a *fold* and ι is an immersion. We will use this factorization again to prove bounded cancellation in Lemma 1.4 and Proposition 3.1.

Observation. *If $\phi : F \to F$ is irreducible, then it is injective.*

Proof. If ϕ is not injective, then at least one of the folds in Stallings' factorization of ϕ collapses a subgraph of the domain. In particular, the kernel of ϕ contains a proper free factor $A \leq F$; therefore, $\phi(A) = \{1\} \leq A$ and $[\phi]$ is reducible. $\qquad \square$

Let $\mathcal{A} = \{A_1, \ldots, A_l\}$ be a nontrivial subgroup system of F, then an $\mathcal{A}$-marked graph (Γ_*, α_*) is a collection of graphs $\Gamma_* = \{\Gamma_1, \ldots, \Gamma_l\}$ where the finite connected core graphs Γ_i are indexed by markings: isomorphisms $\alpha_i : A_i \to \pi_1(\Gamma_i)$. When $\mathcal{A} = \{A\}$ is a singleton, we may write A-marked graph in place of $\mathcal{A}$-marked graph. A marked graph (Γ, α) is an F-marked graph. For any marked graph (Γ, α) and any conjugacy class of a nontrivial element $g \in F$, denoted by $[g]$, we define its length with respect to α, $\|g\|_\alpha$, to be the (combinatorial) length of the immersed loop in $\overline{\Gamma}$ representing $[g]$.

More generally, we want to consider pairs $\mathcal{A} \preceq \mathcal{B}$ of free factor systems of F. For each component $B_i \in \mathcal{B}$, let $\mathcal{A}_i$ be either the nonempty maximal subset of $\mathcal{A}$ carried by B_i or the trivial system if no such subset exists. Typically, we shall replace components of $\mathcal{A}$ with conjugates so that $\mathcal{A}$ is also a free factor system of $\mathcal{B}$, i.e. each $\mathcal{A}_i$ is a free factor system of B_i (next remark below). A $(\mathcal{B}, \mathcal{A})$-forest T_* is a simplicial forest of $(B_i, \mathcal{A}_i)$-trees T_i, i.e. a collection of simplicial trees $T_* = \{T_1, \ldots, T_k\}$ where each tree T_i is equipped with a minimal simplicial B_i-action whose edge stabilizers are trivial and point stabilizers are trivial or conjugates (in B_i) of $\mathcal{A}_i$-components; we note that a $(B, \{B\})$-tree is a point also known as a degenerate tree.

When $\mathcal{A}$ is the trivial system, an $(F, \mathcal{A})$-tree is a free minimal F-tree T. In that case, the quotient $F \backslash T$ is a marked graph. When $\mathcal{A}$ is a nontrivial proper free factor system, the quotient of an $(F, \mathcal{A})$-tree is a *graph of groups* decomposition of F with trivial edge groups and $\mathcal{A}$ as the nontrivial vertex groups [29]; such decompositions are also known as free splittings of F and they will be our relative analogues for marked graphs. Any given $(F, \mathcal{A})$-tree endowed with the combinatorial metric has an associated length function $l_T : F \to \mathbb{R}_{\geq 0}$. Precisely, any automorphism of a simplicial tree is either elliptic (fixes a point) or loxodromic (preserves an *axis* of least displacement). If $g \in F$ is elliptic, then $l_T(g) = 0$; otherwise, $l_T(g) > 0$ is the translation distance of g acting on its axis. When

9

$\mathcal{A}$ is trivial and $F \backslash T$ is the marked graph (Γ, α), then $l_T(\cdot) = \| \cdot \|_\alpha$ is a positive function $F \to \mathbb{R}_{>0}$. The $\underline{\text{minimal subtree}}$ $T(H) \subset T$ of a nontrivial subgroup $H \leq F$ is the smallest subtree with a minimal induced H-action, i.e. the union of all fixed points and axes of nontrivial elements in H.

Remark. Let $\mathcal{A} \preceq \mathcal{B}$ be free factor systems of F, $B_i \in \mathcal{B}$ be a free factor with nontrivial $\mathcal{A}_i$, and T be an $(F, \mathcal{A})$-tree. Choose a connected fundamental domain of B_i acting on $T(B_i)$ and the nontrivial stabilizers in B_i of vertices in the domain form a free factor system $\mathcal{A}'_i$ of B_i. Then $\mathcal{A}'_i$ and $\mathcal{A}_i$ are equivalent as free factor systems of F since $\mathcal{A}_i \preceq \{B_i\}$.

Suppose $\psi : F \to F'$ is an injective homomorphism. Then it determines a contravariant preimage function ψ^{-1} from the poset of free factor systems of F' to the poset of free factor systems of F. One way to define this function is through trees. Let $\mathcal{A}'$ be a free factor system of F', T be any $(F', \mathcal{A}')$-tree, and $T(\psi(F)) \subset T$ be the minimal subtree of $\psi(F) \leq F'$. In particular, the quotient $\psi(F) \backslash T(\psi(F))$ is a free splitting of $\psi(F)$ whose nontrivial vertex groups form a nontrivial free factor system $\mathcal{A}''$ of $\psi(F)$. If $T(\psi(F))$ has a free $\psi(F)$-action, then this condition is independent of the choice of $(F', \mathcal{A}')$-tree T and we set $\psi^{-1} \cdot \mathcal{A}'$ to be the trivial system. Otherwise, we have a nontrivial free factor system $\mathcal{A}''$ of $\psi(F)$ and a corresponding nontrivial free factor system of F via the isomorphism $\psi : F \to \psi(F)$; the equivalence class of the latter free factor system is independent of the $(F', \mathcal{A}')$-tree T and we will denote it by $\psi^{-1} \cdot \mathcal{A}'$.

Let T and T' be $(F, \mathcal{A})$- and $(F', \mathcal{A}')$-trees respectively and $\psi : F \to F'$ be an injective homomorphism such that $\mathcal{A}'$ carries $\psi(\mathcal{A})$ or, equivalently, $\psi^{-1} \cdot \mathcal{A}' \succeq \mathcal{A}$. This *carrying condition* ensures that elliptic elements in F (with respect to T) have elliptic ψ-images (with respect to T'). A tree map $f : T \to T'$ is $\underline{\psi\text{-equivariant}}$ if $f(g \cdot x) = \psi(g) \cdot f(x)$ for all $g \in F$ and $x \in T$. If we require $\psi^{-1} \cdot \mathcal{A}' = \mathcal{A}$, then loxodromic elements in F have loxodromic ψ-images; this is the relative analogue of a π_1-injective graph map.

Suppose $\mathcal{A}$ is a nontrivial $[\phi]$-invariant free factor system of F for some injective endomorphism $\phi : F \to F$; so there is a function $\sigma : \{1, \ldots, l\} \to \{1, \ldots, l\}$ and inner automorphisms $i_{x_i} : F \to F$ such that the restrictions $\phi_i = (i_{x_i} \circ \phi)|_{A_i}$ are homomorphisms $\phi_i : A_i \to A_{\sigma(i)}$. The collection $\{\phi_i\}$ is a $\underline{\text{restriction of } \phi \text{ to } \mathcal{A}}$, which we denote by $\phi|_{\mathcal{A}} : \mathcal{A} \to \mathcal{A}$. A $\underline{\text{weak representative for } [\phi_i]}$ is a $\underline{\text{graph map } f_i : \Gamma_i \to \Gamma_{\sigma(i)}}$ from an A_i-marked graph (Γ_i, α_i) to an $A_{\sigma(i)}$-marked graph $(\Gamma_{\sigma(i)}, \alpha_{\sigma(i)})$ such that $[\pi_1(f_i) \circ \alpha_i] = [\alpha_{\sigma(i)} \circ \phi_i]$ for all $i \in \{1, \ldots, l\}$. A $\underline{\text{weak representative for } [\phi|_{\mathcal{A}}]}$ is a graph map $f_* : \Gamma_* \to \Gamma_*$ of an $\mathcal{A}$-marked graph (Γ_*, α_*) that is a union of weak representatives for $[\phi_i]$. A $\underline{\text{natural representative}}$ is a weak representative that is also natural. A $\underline{\text{topological representative}}$ is a weak representative that has no pretrivial edges and whose underlying graph Γ_* has no bivalent vertices.

Recall $\mathcal{A} \preceq \mathcal{B}$. Suppose $\phi^{-1} \cdot \mathcal{A} = \mathcal{A}$, $\mathcal{B}$ is $[\phi]$-invariant, and $\phi|_{\mathcal{B}} = \{\phi'_i : B_i \to B_{\sigma'(i)}\}$ is a restriction of ϕ. An $\underline{\mathcal{A}\text{-relative weak representative}}$ for $\phi|_{\mathcal{B}}$ is a forest map $f_* : T_* \to T_*$ of a $(\mathcal{B}, \mathcal{A})$-forest T_* composed of ϕ'_i-equivariant tree maps $f_i : T_i \to T_{\sigma'(i)}$. As before, an $\underline{\mathcal{A}\text{-relative natural representative}}$ is an $\mathcal{A}$-relative weak representative that is also natural;

an $\mathcal{A}$-relative representative is an $\mathcal{A}$-relative weak representative with no pretrivial edges and whose underlying forest has no bivalent vertices. A relative weak representative is minimal if it has no *orbit-closed* invariant subforests whose components are bounded. An (expanding resp.) $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$ is an $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$ that is also an (expanding resp.) immersion.

In the first part of the paper, we show that any injective endomorphism of F has a canonical invariant free factor system $\mathcal{A}$ and an expanding $\mathcal{A}$-relative immersion f such that $\mathcal{A}$ is eventually mapped into a canonical fixed free factor system. In the second part of the paper, we use the canonical systems and the expanding relative immersion to prove the main theorem.

Dynamics of free group endomorphisms

1 Stallings graphs and bounded cancellation

Let (Γ, α) be a marked graph with no bivalent vertices. For any nontrivial subgroup system $\mathcal{H}$ of F, the Stallings (subgroup) graph for $\mathcal{H}$ with respect to a marked graph (Γ, α) is the smallest $\mathcal{H}$-marked graph $(S(\mathcal{H}), \beta_*)$ along with an immersion $\iota_* : S(\mathcal{H}) \to \Gamma$ such that $[\pi_1(\iota_i) \circ \beta_i] = [\alpha|_{H_i}]$ for every $H_i \in \mathcal{H}$. Alternatively, $S(\mathcal{H})$ is the collection of *cores* $S(H_i)$ of the covers $\hat{\Gamma}_{H_i}$ of Γ corresponding to $\alpha(H_i)$, i.e. the smallest deformation retract of $\hat{\Gamma}_{H_i}$, and ι_i is the restriction to $S(H_i)$ of the covering map $\hat{\Gamma}_{H_i} \to \Gamma$. In the notation, the marking and immersion for Stallings graphs will usually be omitted. If H and H' are in the same conjugacy class, $[H]$, then there is a homeomorphism $h : S(H) \to S(H')$ such that $\iota = \iota' \circ h$; the converse holds as well. So the Stallings graph $S[H]$ is uniquely determined by the conjugacy class $[H]$. Furthermore, the Stallings graph $S[\mathcal{H}]$ is a finite graph if and only if $\mathcal{H}$ is finitely generated. Suppose $\phi : F \to F$ is injective and $\mathcal{A}$ is a nontrivial free factor system of F. We will study the (iterated) Stallings graphs $(S[\phi^k(\mathcal{A})], \beta_{*,k})$ for $k \geq 1$.

Remark. For any free factor system $\mathcal{F}$ of F and nontrivial subgroup system $\mathcal{H}$ carried by $\mathcal{F}$, we can similarly define the Stallings graph for $\mathcal{H}$ with respect to $\mathcal{F}$-marked graphs (Γ_*, α_*). If $\phi_* : \mathcal{F} \to \mathcal{F}$ is an injective endomorphism, i.e. a collection of injective homomorphisms $\{\phi_i : F_i \to F_{\sigma(i)}\}$ that need not be a restriction of an endomorphism of F, then we can still consider the Stallings graphs $S[\phi_*^k(\mathcal{A})]$ for $k \geq 1$ and $\mathcal{A} \preceq \mathcal{F}$. The point of this remark is that the results of this paper hold when $\phi : F \to F$ is replaced with $\phi_* : \mathcal{F} \to \mathcal{F}$. In other words, the graph Γ need not be connected for our results.

Unlike the ambient graph Γ, we allow Stallings graphs $S = S[\mathcal{H}]$ to have bivalent vertices. More precisely, we subdivide S so that the immersion $\iota_* : S \to \Gamma$ is simplicial, i.e. maps edges to edges. With this subdivision, we get a combinatorial metric on (S, β_*) that is compatible with (Γ, α), i.e. $\|g\|_\alpha = \|g\|_{\beta_i}$ for any nontrivial element g in H_i.

Lemma 1.1. *Let $\phi : F \to F$ be injective and H be a finitely generated nontrivial subgroup of F. If H is not eventually $[\phi]$-periodic, then the length of the longest natural edge in $S[\phi^k(H)]$ is unbounded as $k \to \infty$.*

Proof. Suppose the length of the longest natural edge in $S[\phi^k(H)]$ with respect to some marked graph (Γ, α) was uniformly bounded as $k \to \infty$. We want to show that H is eventually $[\phi]$-periodic. The number of natural edges in $S[\phi^k(H)]$ is bounded above by $3 \cdot \operatorname{rank}(H) - 3$. Our assumption implies there is a bound on the volume of (i.e. number of edges in) the graphs $S[\phi^k(H)]$ as $k \to \infty$. So the sequence $S[\phi^k(H)]$ is eventually periodic, i.e. there are integers $m > n \geq 1$ and a simplicial isomorphism $h : S[\phi^m(H)] \to S[\phi^n(H)]$ such that $\iota_m = \iota_n \circ h$. Since a Stallings graph determines the conjugacy class of its defining subgroup, we have $[\phi^m(H)] = [\phi^n(H)]$, i.e. H is eventually $[\phi]$-periodic. $\qquad \square$

This lemma will be invoked on invariant free factor systems containing a component that is not eventually $[\phi]$-periodic. Conversely, the next lemma handles the case when an invariant free factor system consists entirely of eventually periodic components.

Lemma 1.2. *Let $\phi : F \to F$ be injective and $\mathcal{A}$ be a nontrivial $[\phi]$-invariant free factor system. If all components in $\mathcal{A}$ are eventually $[\phi]$-periodic, then some nonempty subset $\mathcal{B} \subset \mathcal{A}$ is a $[\phi]$-fixed free factor system and $\phi^k(\mathcal{A})$ is carried by $\mathcal{B}$ for some $k \geq 0$.*

E.g. if $\phi : F \to F$ is injective and F is eventually $[\phi]$-periodic, then ϕ is an automorphism.

Proof. Let $\sigma : \{1, \ldots, l\} \to \{1, \ldots, l\}$ be the function used to define the $[\phi]$-invariance of $\mathcal{A} = \{A_1, \ldots, A_l\}$. Then there is a nonempty subset $J \subset \{1, \ldots, l\}$ on which σ acts as a bijection and $\sigma^l(\{1, \ldots, l\}) = J$. Let $\mathcal{B} \preceq \mathcal{A}$ by the nontrivial $[\phi]$-invariant free factor system corresponding to J. Then $\phi^l(\mathcal{A})$ is carried by $\mathcal{B}$ since the image of σ^l is J. It remains to show that $\mathcal{B}$ is $[\phi]$-fixed. Set j to be the order of $\sigma|_J$, fix $B \in \mathcal{B}$, and let $i_x : F \to F$ be the inner automorphism such that $i_x \circ \phi^j(B) \leq B$. Define $\psi = i_x \circ \phi^j$. As B is eventually $[\phi]$-periodic and hence eventually ϕ^j-periodic, there are integers $m > n \geq 1$ such that $[\psi^m(B)] = [\phi^{jm}(B)] = [\phi^{jn}(B)] = [\psi^n(B)]$. Therefore, there is an element $y \in F$ such that $y\psi^n(B)y^{-1} = \psi^m(B) \leq \psi^n(B)$. But no finitely generated subgroup of F is conjugate to a proper subgroup of itself (Lemma 1.3 below). So $\psi^m(B) = \psi^n(B)$ and, by injectivity of ϕ, $\psi(B) = B$. Since $B \in \mathcal{B}$ was arbitrary, ϕ^j fixes the components of $\mathcal{B}$ up to conjugation; as $\mathcal{B}$ is $[\phi]$-invariant, it must be $[\phi]$-fixed. $\qquad \square$

The following fact will be used again in the proof of Proposition 2.4.

Lemma 1.3. *No finitely generated subgroup of F is conjugate to a proper subgroup of itself.*

Proof. By Marshall Hall's theorem, free groups are *subgroup separable/locally extended residual finite (LERF)*, i.e. for any finitely generated subgroup $H \leq F$ and any element

$g \in F \backslash H$, there is a finite group G and a homomorphism $\varphi : F \to G$ such that $\varphi(g) \notin \varphi(H)$. See [30] for a proof due to Stallings.

For a contradiction, suppose there is an element $y \in F$ such that $yHy^{-1} \leq H$ and $g \in H \setminus yHy^{-1}$. By subgroup separability, there is a finite group G and homomorphism $\varphi : F \to G$ such that $\varphi(g) \notin \varphi(yHy^{-1})$. But $g \in H$ implies $\varphi(g) \in \varphi(H)$ and, by finiteness of G, $yHy^{-1} \leq H$ implies $\varphi(yHy^{-1}) = \varphi(H)$ — a contradiction. $\qquad\square$

The next lemma, known as the <u>Bounded Cancellation Lemma</u>, will be used extensively in this paper. At the risk of overloading notation, for an edge-path p in a graph Γ, $[p]$ denotes the immersed edge-path that is homotopic to p *rel. endpoints*; for a loop ρ in Γ, $[\rho]$ will be the immersed loop that is *freely* homotopic to ρ.

Lemma 1.4 (Bounded Cancellation). *Let $g : \Gamma \to \Gamma'$ be a π_1-injective graph map. Then there is a constant $C(g)$ such that, for any natural edge-path decomposition $p_1 \cdot p_2$ of an immersed path in the universal cover $\tilde{\Gamma}$, the edge-path $[\tilde{g}(p_1)] \cdot [\tilde{g}(p_2)]$ is contained in the $C(g)$-neighborhood of $[\tilde{g}(p_1) \cdot \tilde{g}(p_2)]$.*

The following proof is due to Bestvina-Feighn-Handel [4, Lemma 3.1].

Proof. Any graph map $g : \Gamma \to \Gamma'$ factors as a pretrivial edge collapse and edge subdivision g_0, a composition of $r \geq 0$ folds $g_r \circ \cdots \circ g_1$, and a simplicial immersion g_{r+1}. The subdivisions and immersions always have cancellation constants 0; by π_1-injectivity, the collapse has cancellation constant 0 and each fold has cancellation constant 1. Thus we may set $C(g) = r$. $\qquad\square$

Let $f : \Gamma \to \Gamma$ be a topological representative for an injective endomorphism $\phi : F \to F$, $\mathcal{A}$ be a nontrivial $[\phi]$-invariant free factor system of F, and $\hat{\Gamma}_k$ be the disjoint union of covers of Γ corresponding to $\phi^k(\mathcal{A})$ for some $k \geq 1$. Then f lifts to a map $\hat{f}_k : \hat{\Gamma}_k \to \hat{\Gamma}_k$ and the deformation retraction $\hat{\Gamma}_k \to S[\phi^k(\mathcal{A})]$ induces a map $\bar{f}_k : S[\phi^k(\mathcal{A})] \to S[\phi^k(\mathcal{A})]$ with $K(\bar{f}_k) = K(f)$ and $C(\bar{f}_k) = C(f)$. We shall call $\bar{f}_k$ the <u>(k-th) homotopy lift</u> of f.

The Stallings graphs $S[\phi^k(\mathcal{A})]$ are $\phi^k(\mathcal{A})$-marked graphs by definition and the maps $\bar{f}_k$ are weak representatives for $[\phi|_{\phi^k(\mathcal{A})}]$ that might map branch points to bivalent vertices. Note that injectivity of ϕ allows us to also consider the graphs as $\mathcal{A}$-marked graphs and $\bar{f}_k$ as weak representatives for $[\phi|_{\mathcal{A}}]$. We hope to replace the weak representatives $\bar{f}_k$ with homotopic natural representatives while maintaining uniform control on the Lipschitz and cancellation constants. The next lemma allows us to measure how close the homotopy lift $\bar{f}_k$ is to mapping branch points to branch points.

Lemma 1.5. *Suppose $g : \Gamma \to \Gamma'$ be a π_1-injective graph map with cancellation constant $C = C(g)$. Then g maps branch points to the C-neighborhood of branch points.*

Proof. Set $C = C(g)$. If Γ' is the C-neighborhood of its branch points, then there is nothing to prove. Suppose $\nu \in \Gamma'$ is a bivalent vertex whose distance to the nearest branch point

is greater than C. We need to show that ν is not the g-image of any branch point. Set ϵ_1 and ϵ_2 to be the distinct oriented directions originating from ν and $\bar{\epsilon}_1, \bar{\epsilon}_2$ are the same directions with opposite orientation.

For a contradiction, let $v \in \Gamma$ be a branch point and $g(v) = \nu$. As v is a branch point, there are at least three distinct oriented directions originating from v: e_1, e_2, and e_3. Let p_{12} be an immersed path that starts and ends with e_1 and $\bar{e}_2$ respectively and define p_{23} similarly. Set $p_{13} = [p_{12} \cdot p_{23}]$ and $p'_{13} = p_{12} \cdot \bar{p}_{23}$, where $\bar{p}_{23}$ is the reversal of the path p_{23}. See Figure 5 for an illustration. Although the paths are loops, we still treat them as paths, i.e. tightening is done rel. the endpoints. Without loss of generality, assume $[g(p_{12})]$ starts with ϵ_1.
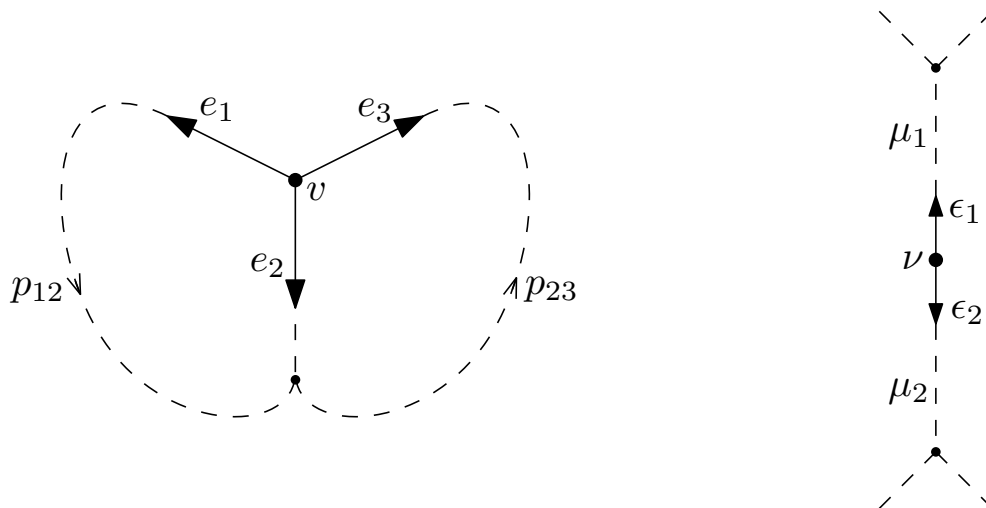

Figure 5: Schematic for paths p_{12}, p_{23}, p_{13}, and p'_{13}. The path p_{13} starts with e_1 follows the dashed path and ends with $\bar{e}_3$. The path p'_{13} is the "figure 8" path traced by p_{12} then $\bar{p}_{23}$.

If $[g(p_{12})]$ ends with $\bar{\epsilon}_1$, then $[g(p_{12})] = \mu_1 \cdot \rho \cdot \bar{\mu}_1$, where μ_1 is an extension of ϵ_1 to an embedded path from ν to a branch point and ρ is an immersed nontrivial loop. By hypothesis, μ_1 is longer than C. Since p_{12} starts and ends with e_1 and $\bar{e}_2$ respectively, the concatenation $p_{12} \cdot p_{12}$ is an immersed path such that $[g(p_{12})] \cdot [g(p_{12})]$ has $\bar{\mu}_1 \cdot \mu_1$ as a subpath, violating bounded cancellation. So we must assume $[g(p_{12})]$ starts and ends with ϵ_1 and $\bar{\epsilon}_2$.

If $[g(p_{23})]$ starts and ends with ϵ_2 and $\bar{\epsilon}_1$, then $[g(p_{13})] = [g(p_{12}) \cdot g(p_{23})]$ starts and ends with ϵ_1 and $\bar{\epsilon}_1$ respectively, which violates bounded cancellation for the same reason given in the previous paragraph. Similarly, if $[g(p_{23})]$ starts and ends with ϵ_1 and $\bar{\epsilon}_2$, we rule out this possibility by considering $[g(p'_{13})]$. We have ruled out all cases and reached a contradiction; therefore, no branch point v of Γ is mapped to ν. $\qquad\square$

14

Corollary 1.6. *Let $g : \Gamma \to \Gamma'$ be a K-Lipschitz π_1-injective graph map with cancellation constant C. Then g is homotopic to a $(K+C)$-Lipschitz natural graph map with cancellation constant $2C$.*

Proof. By Lemma 1.5, g maps branch points to the C-neighborhood of branch points. So we can find a graph map g' homotopic to g that maps branch points to branch points. Since the homotopy is moving images of branch points a distance at most C, we can use $K(g') = K + C$ and $C(g') = 2C$.

The bounded cancellation lemma only considers natural edge-paths, so homotopies that are supported in the interior of natural edges will not affect the cancellation constant. Using a tightening homotopy supported in the interior of natural edges in Γ, we may assume g' is a natural graph map. The homotopy will not affect the Lipschitz and cancellation constants. $\square$

So we can replace the homotopy lift $\bar{f}_k$ with a homotopic natural representative that has Lipschitz and cancellation constants $K(f)+C(f)$ and $2C(f)$ respectively. It is tempting to collapse the pretrivial edges and forget bivalent vertices to get a topological representative but we will not do that since we want compatibility: $\|\cdot\|_{\beta_{*,k}}$ is the restriction of $\|\cdot\|_\alpha$ to $\phi^k(\mathcal{A})$. To summarize the properties of homotopy lifts that will be used in the next sections:

Proposition 1.7. *Suppose $\phi : F \to F$ is injective, $f : \Gamma \to \Gamma$ is a topological representative for $[\phi]$, and $\mathcal{A}$ is a nontrivial $[\phi]$-invariant free factor system. For any $k \geq 1$, there is a natural representative $\bar{f}_k : S[\phi^k(\mathcal{A})] \to S[\phi^k(\mathcal{A})]$ for $[\phi|_{\phi^k(\mathcal{A})}]$ with Lipschitz and cancellation constants $K(\bar{f}_k) = K(f) + C(f)$ and $C(\bar{f}_k) = 2C(f)$ respectively.*

The crucial point is that the Lipschitz and cancellation constants are independent of k.

2   Canonical free factor systems

In this section, we will construct canonical fixed and invariant free factor systems for any given injective endomorphism of F. The invariant free factor system, called the *elliptic free factor system*, is crucial for the construction of *(expanding) relative immersions* later in the paper.

Suppose $\mathcal{A}$ is a subgroup system and $\phi : F \to F$ is an injective endomorphism. We shall say that a conjugacy class $[g]$ in F has an infinite $[\phi]$-tail if for every $n \geq 1$, there is a conjugacy class $[g_n]$ in F such that $[\phi^n(g_n)] = [g]$. The system $\mathcal{A}$ carries an infinite $[\phi]$-tail of a conjugacy class $[g]$ in F if for every $n \geq 1$, there is a conjugacy class $[g_n]$ carried by $\mathcal{A}$ such that $[\phi^n(g_n)] = [g]$. We now state and prove the main technical result of this section.

Theorem 2.1. *If $\phi : F \to F$ is injective, then $[\phi]$ has a nontrivial fixed free factor system if and only if some nontrivial conjugacy class has an infinite $[\phi]$-tail.*

15

Remark. Reynolds defined *expansive* endomorphisms [28, Definition 3.8] and $[\phi]$ is expansive in this sense exactly when only the trivial conjugacy class has an infinite $[\phi]$-tail. Under this equivalence, Reynolds' Remark 3.12 in [28] is a weaker form of this theorem.

Proof. Let $\phi : F \to F$ be an injective endomorphism. The forward direction is obvious: if $[\phi]$ has a nontrivial fixed free factor system $\mathcal{A}$, then any nontrivial conjugacy class $[g]$ carried by $\mathcal{A}$ has an infinite $[\phi]$-tail. The main content of the theorem is in the reverse direction. Fix a nontrivial conjugacy class $[g]$ in F with an infinite $[\phi]$-tail. We proceed by descending down the poset of free factor systems. The following claim is the key idea:

Claim (Descent). *Let $\mathcal{D}$ be a $[\phi]$-invariant free factor system that carries an infinite $[\phi]$-tail of $[g]$. If $\mathcal{D}$ contains a component that is not eventually $[\phi]$-periodic, then some $[\phi]$-invariant free factor system $\mathcal{D}' \prec \mathcal{D}$ carries an infinite $[\phi]$-tail of $[g]$.*

Since there are no infinite chains in the poset of free factor systems, the descent (proven below) starts with the free factor system $\{F\}$ and then finds a $[\phi]$-invariant free factor sytem $\mathcal{A}$ that carries an infinite $[\phi]$-tail of $[g]$ and whose components are eventually $[\phi]$-periodic; such a free factor system contains a nontrivial $[\phi]$-fixed free factor system by Lemma 1.2 and we are done. $\qquad\square$

Note that the assumption "$\mathcal{D}$ contains a component that is not eventually $[\phi]$-periodic" in the descent claim is more general than we really need for our conclusion. We could have worked with a more natural assumption "$\mathcal{D}$ has no periodic components" and reached the same conclusion. However, we give the general argument instead since we will invoke a variation of it in the next proposition.

Proof of descent. Let (Γ, α) be a marked graph, $f : \Gamma \to \Gamma$ be a topological representative for $[\phi]$, and set $K = K(f) + C(f)$ and $C = 2C(f)$. Suppose $\mathcal{D}$ is a $[\phi]$-invariant free factor system that carries an infinite $[\phi]$-tail $[g_n]_{n \geq 1}$ of $[g]$ and contains a component that is not eventually $[\phi]$-periodic. Then, for all $k \geq 1$ and $n \geq k$, there is an immersed loop $\rho_k(g_n)$ in $\Delta_k = S[\phi^k(\mathcal{D})]$ corresponding to $[\phi^k(g_n)]$, where $\Delta_k = S[\phi^k(\mathcal{D})]$ is the Stallings graph with respect to (Γ, α). Set $L = \max\{\|g\|_\alpha, C\}$.

For all $k \geq 1$, let $\bar{f}_k : \Delta_k \to \Delta_k$ be the natural representatives for $[\phi|_{\phi^k(\mathcal{D})}]$ given by Proposition 1.7. In particular, these representatives have Lipschitz and cancellation constants K and C respectively. As $[g_n]_{n \geq 1}$ is an infinite $[\phi]$-tail of $[g]$ carried by $\mathcal{D}$, we get, for any fixed $k \geq 1$, an infinite sequence of immersed loops $(\rho_k(g_n))_{n \geq k}$ in Δ_k such that the free homotopy classes of $\bar{f}_k^{n-k}(\rho_k(g_n))$ have length $\|g\|_\alpha$ for all $n \geq k$.

Form a directed graph $\mathbb{G}_k$ whose vertices are the natural edges of Δ_k and there is a directed edge $E_i \to E_j$ if $\bar{f}_k$ maps natural edge E_i over E_j. Note that the number of natural edges of Δ_k is at most $N = 3 \cdot \text{rank}(F) - 3$ and so $\mathbb{G}_k$ has at most N vertices.

Since $\mathcal{D}$ contains a component that is not eventually $[\phi]$-periodic, the length of natural edges in Δ_k is unbounded as $k \to \infty$ by Lemma 1.1. Fix $k \gg 0$ such that the longest natural edge in Δ_k is longer than $L \cdot K^{N-1}$. Let $\mathcal{L}_0$ be the natural edges of Δ_k longer

than $L \cdot K^{N-1}$ and $\mathcal{L}$ be the union of $\mathcal{L}_0$ and all natural edges on a directed path to $\mathcal{L}_0$ in $\mathbb{G}_k$. Since $\bar{f}_k$ is K-Lipschitz and the shortest directed path in $\mathbb{G}_k$ from a natural edge in $\mathcal{L}$ to $\mathcal{L}_0$ has at most N natural edges on it, every natural edge in $\mathcal{L}$ is longer than L. The natural edges in $\mathcal{L}$ will be referred to as the *long* natural edges and the remaining natural edges as *short*.

Set $\Delta' \subset \Delta_k$ to be the union of short natural edges, which will be a proper subgraph since long natural edges exist. The subgraph is automatically $\bar{f}_k$-invariant by construction of $\mathcal{L}$. Since $\rho_k(g_k)$ is an immersed loop in Δ_k with length $\|g\|_\alpha \leq L$, its natural edges are short and Δ' is a nonempty, noncontractible proper subgraph of Δ_k. Therefore, Δ' determines a nontrivial $[\phi]$-invariant proper free factor system $\mathcal{D}' \prec \mathcal{D}$. Technically, it determines an invariant free factor system of $\phi^k(F)$ but, as ϕ is injective, this corresponds to an invariant free factor system of F. It remains to show that $\mathcal{D}'$ carries an infinite $[\phi]$-tail of $[g]$.

Let $\mathbb{L} \subset \mathbb{G}_k$ be the full subgraph generated by the long natural edges $\mathcal{L}$. If there are no directed cycles in $\mathbb{L}$, then $\bar{f}_k^N(\Delta_k) \subset \Delta'$; in this case, the sequence of nontrivial loops $(\bar{f}_k^N(\rho_k(g_n)))_{n \geq k+N}$ in Δ' determines an infinite $[\phi]$-tail of $[g]$ carried by $\mathcal{D}'$ and we are done. Now suppose there are directed cycles in $\mathbb{L}$ and let ρ be an immersed loop in Δ_k that contains a long natural edge in such a cycle. Then, by bounded cancellation and the fact long natural edges are longer than $L \geq C$, $[\bar{f}_k^m(\rho)]$ contains a long natural edge in the same directed cycle in $\mathbb{L}$ for all $m \geq 1$. Consequently, none of the immersed loops $\rho_k(g_n)$ in Δ_k contain a long natural edge that is in a directed cycle of $\mathbb{L}$. Therefore, as far as the sequence of loops $(\rho_k(g_n))_{n \geq k}$ is concerned, we may assume there are no directed cycles in $\mathbb{L}$ and, as before, the sequence $(\bar{f}_k^N(\rho_k(g_n)))_{n \geq k+N}$ determines an infinite $[\phi]$-tail of $[g]$ carried by $\mathcal{D}'$. $\square$

The following dichotomy is (equivalent to) a result in Reynolds' thesis.

Corollary 2.2 ([28, Proposition 3.11]). *If $\phi : F \to F$ is irreducible, then either ϕ is an automorphism or only the trivial conjugacy class has an infinite $[\phi]$-tail.*

Proof. Suppose ϕ is irreducible and there is a nontrivial conjugacy class with an infinite $[\phi]$-tail. By Theorem 2.1, there is a nontrivial $[\phi]$-fixed free factor system $\mathcal{A}$. Since ϕ is irreducible, $\mathcal{A} = \{F\}$ and ϕ is an automorphism. $\square$

The fixed free factor system given by Theorem 2.1 may depend on the chosen conjugacy class $[g]$ with an infinite tail or the marked graphs (Γ, α) chosen in the descent. The next proposition constructs a canonical fixed free factor system for $[\phi]$; this system carries all conjugacy classes with an infinite tail as well as all finitely generated fixed subgroup system. The proof will use both descent and ascent (like a losing game of Tetris) in the poset of free factor systems!

Proposition 2.3. *If $\phi : F \to F$ is injective, then there is a unique maximal $[\phi]$-fixed free factor system $\mathcal{A}$. Precisely, $\mathcal{A}$ carries every conjugacy class with an infinite $[\phi]$-tail and every finitely generated $[\phi]$-fixed subgroup system.*

Proof. Let $\phi : F \to F$ be an injective endomorphism. If the trivial conjugacy class is the only conjugacy class with an infinite $[\phi]$-tail, then the trivial system is the only $[\phi]$-fixed subgroup system. In this case, the trivial system is the unique maximal $[\phi]$-fixed free factor system and we are done as it vacuously carries all conjugacy classes with infinite $[\phi]$-tails and all $[\phi]$-fixed subgroup systems. We can now assume some nontrivial conjugacy class has an infinite $[\phi]$-tail. By Theorem 2.1, ϕ has a nontrivial fixed free factor system $\mathcal{D}_0$. We proceed by ascending up the poset of free factor systems:

Claim (Ascent). *Let $\mathcal{B}$ be a finitely generated $[\phi]$-fixed subgroup system, $[g]$ a conjugacy class with an infinite $[\phi]$-tail, and $\mathcal{D}$ a nontrivial $[\phi]$-fixed free factor system. If $\mathcal{D}$ does not carry both $\mathcal{B}$ and $[g]$, then some $[\phi]$-fixed free factor system $\mathcal{D}' \succ \mathcal{D}$ carries both $\mathcal{B}$ and $[g]$.*

Once again, as there are no infinite chains in the poset of free factor systems, the ascent (proven below) starts with the nontrivial $[\phi]$-fixed proper free factor system $\mathcal{D}_0$ and stops at a necessarily unique maximal $[\phi]$-fixed free factor system $\mathcal{A}$ that carries all finitely generated $[\phi]$-fixed subgroup systems and all conjugacy classes with infinite $[\phi]$-tails. $\qquad\square$

Proof of ascent. Let $\mathcal{B}$ be a finitely generated nontrivial $[\phi]$-fixed subgroup system of F, $[g]$ be a nontrivial conjugacy class in F with an infinite $[\phi]$-tail, and $\mathcal{D}$ be a nontrivial $[\phi]$-fixed free factor system of F that does not carry both $\mathcal{B}$ and $[g]$. We now describe the descent:

Claim (Descent). *Let $\mathcal{D}''$ be a $[\phi]$-invariant free factor system that carries $\mathcal{D}$, $\mathcal{B}$, and an infinite $[\phi]$-tail of $[g]$. If $\mathcal{D}''$ contains a component that is not eventually $[\phi]$-periodic, then some $[\phi]$-invariant free factor system $\mathcal{D}''' \prec \mathcal{D}''$ carries $\mathcal{D}$, $\mathcal{B}$, and an infinite $[\phi]$-tail of $[g]$.*

Starting with $\{F\}$, the descent (proven below) will find a nontrivial $[\phi]$-invariant free factor system $\mathcal{D}^*$ that carries $\mathcal{D}$, $\mathcal{B}$, and an infinite $[\phi]$-tail of $[g]$ and whose components are eventually $[\phi]$-periodic. By Lemma 1.2, $\mathcal{D}^*$ contains a $[\phi]$-fixed free factor subsystem $\mathcal{D}' \subset \mathcal{D}^*$ such that $\phi^k(\mathcal{D}^*)$ is carried by $\mathcal{D}'$ for some $k \geq 0$. As $\mathcal{D}$ and $\mathcal{B}$ are $[\phi]$-fixed, they are carried by $\mathcal{D}'$. Similarly, $\mathcal{D}'$ carries $[g]$ since $\mathcal{D}^*$ carries an infinite $[\phi]$-tail of $[g]$. So $\mathcal{D}'$ is a $[\phi]$-fixed free factor system that carries $\mathcal{D}$, $\mathcal{B}$, and $[g]$ as needed for ascent. $\qquad\square$

Proof of descent. Let (Γ, α) be a marked graph, $f : \Gamma \to \Gamma$ be a topological representative for $[\phi]$, and set $K = K(f) + C(f)$ and $C = 2C(f)$. Suppose $\mathcal{D}''$ is a $[\phi]$-invariant free factor system that carries $\mathcal{D}$, $\mathcal{B}$, and an infinite $[\phi]$-tail of $[g]$. Let $S[\phi^k(\mathcal{D})]$ and $S[\phi^k(\mathcal{B})]$ be the Stallings graphs with respect to (Γ, α). Since $\mathcal{D}$ and $\mathcal{B}$ are finitely generated and $[\phi]$-fixed, the length of the longest immersed loop in $S[\phi^k(\mathcal{D})]$ or $S[\phi^k(\mathcal{B})]$ that covers any edge at most twice is uniformly bounded by some L_0 for all $k \geq 1$. Set $L = \max\{L_0, \|g\|_\alpha, C\}$. The proof now mimics that of the descent in Theorem 2.1 and we only give a sketch.

For all $k \geq 1$, let $\Delta''_k = S[\phi^k(\mathcal{D}'')]$ and, by Proposition 1.7, there is a K-Lipschitz natural representative for $[\phi|_{\mathcal{D}''}]$, $\bar{f}_k : \Delta''_k \to \Delta''_k$, that has cancellation constant C. As some component in $\mathcal{D}''$ is not eventually $[\phi]$-periodic, we can fix $k \gg 0$ so that the longest natural edge in Δ''_k is longer than $L \cdot K^{N-1}$ by Lemma 1.1. Define the long and short natural edges as before and deduce long natural edges are longer than L. Set $\Delta''' \subset \Delta''_k$

to be the union of the short natural edges, which is necessarily proper and $\bar{f}_k$-invariant. Recall that immersed loops of $S[\phi^k(\mathcal{D})]$ and $S[\phi^k(\mathcal{B})]$ that cover any edge at most twice have length bounded by $L_0 \leq L$ and these Stallings graphs have simplicial immersions into Δ_k''. Hence the images of these immersions lie in the subgraph of short natural edges Δ'''. So Δ''' is neither empty nor contractible. The subgraph Δ''' determines a $[\phi]$-invariant proper free factor system $\mathcal{D}''' \prec \mathcal{D}''$ that carries both $\mathcal{D}$ and $\mathcal{B}$ since both $S[\phi^k(\mathcal{D})]$ and $S[\phi^k(\mathcal{B})]$ have immersions into Δ'''. From the proof of descent in Theorem 2.1, $L \geq C$ implies $\mathcal{D}'''$ carries an infinite $[\phi]$-tail of $[g]$. $\qquad\square$

Although this proposition produces a canonical fixed free factor system for an injective endomorphism, we shall enlarge the system again to get a better $[\phi]$-invariant free factor system that gives us some control of the *relative dynamics* of $[\phi]$; we do this by taking iterated preimages. We then show that the resulting invariant free factor system is a disjoint union of the maximal fixed free factor system with a free factor system that eventually gets mapped into the fixed system.

Proposition 2.4. *If $\phi : F \to F$ is injective and $\mathcal{A}$ is the maximal $[\phi]$-fixed free factor system, then there is a unique maximal $[\phi]$-invariant free factor system $\mathcal{A}^* \succeq \mathcal{A}$ such that $\phi^k(\mathcal{A}^*)$ is carried by $\mathcal{A}$ for some $k \geq 0$. After replacing the components of $\mathcal{A}$ with conjugates if necessary, we can assume $\mathcal{A} \subset \mathcal{A}^*$.*

The free factor system given by this proposition will be called $[\phi]$-elliptic free factor system as it will be carried by the vertex groups in free splittings of F in the subsequent sections.

Proof. Let $\phi : F \to F$ be an injective endomorphism and $\mathcal{A}$ be the maximal $[\phi]$-fixed free factor system of F given by Proposition 2.3. If $\mathcal{A}$ is $\{F\}$ or trivial, then $\mathcal{A}^* = \mathcal{A}$ and we are done. Thus, we assume that $\mathcal{A}$ is a nontrivial proper free factor system. For all $k \geq 1$, define $\mathcal{A}_k = \phi^{-k} \cdot \mathcal{A}$.

Since $\mathcal{A}$ is $[\phi]$-invariant, we get that $\mathcal{A} \preceq \mathcal{A}_k \preceq \mathcal{A}_{k+1}$ for all $k \geq 1$ and, consequently, all $\mathcal{A}_k$ are $[\phi]$-invariant. As there are no infinite chains in the poset of free factor systems of F, the chain of $[\phi]$-invariant free factor systems $\mathcal{A}_k$ ($k \geq 1$) stabilizes and we can set $\mathcal{A}^*$ to be the maximal free factor system in the chain. By construction, $\mathcal{A}^*$ carries a subgroup system $\mathcal{B}$ if and only if $\phi^k(\mathcal{B})$ is carried by $\mathcal{A}$ for some $k \geq 0$ and this implies the uniqueness of $\mathcal{A}^*$ amongst $[\phi]$-invariant free factor system that have ϕ-iterates carried by $\mathcal{A}$. It remains to show that $\mathcal{A} \subset \mathcal{A}^*$ after replacing the components of $\mathcal{A}$ with conjugates if necessary.

As in the proof of Lemma 1.2, suppose $\sigma : \{1, \ldots l\} \to \{1, \ldots, l\}$ is the function associated to the $[\phi]$-invariance of $\mathcal{A}^* = \{A_1, \ldots A_l\}$. Then there is a maximal nonempty subset $J \subset \{1, \ldots, l\}$ on which σ is a bijection. Let $\mathcal{A}_J = \{A_j \in \mathcal{A}^* : j \in J\}$. Since $\mathcal{A}^*$ carries the maximal $[\phi]$-fixed free factor system $\mathcal{A}$, it follows that $\mathcal{A} \preceq \mathcal{A}_J$. Replace components of $\mathcal{A}$ with conjugates if necessary and assume $\mathcal{A}$ is a free factor system of $\mathcal{A}_J$; in particular, each $A \in \mathcal{A}$ is a free factor of some $A_j \in \mathcal{A}_J$. We want to show that $\mathcal{A} \subset \mathcal{A}_J$. Choose a component $A \in \mathcal{A}$ and let $A_j \in \mathcal{A}_J$ be the component such that $A \leq A_j$. Furthermore,

19

fix an inner automorphism $i_x : F \to F$ such that $i_x \circ \phi^s(A_j) \leq A_j$ for some $s \geq 0$. Set $\psi = i_x \circ \phi^s$.

By construction, for some $k \geq 0$, $A_j \in \mathcal{A}^* = \phi^{-ks} \cdot \mathcal{A}$ implies $\psi^k(A_j)$ is conjugate to a subgroup of a $[\phi]$-periodic, and hence $[\psi]$-periodic, component $A' \in \mathcal{A}$. We must have $A' \leq A_j$ since $A' \leq A_{j'} \in \mathcal{A}_J$, $\{A'\}$ carries $\{\psi^k(A_j)\}$, $\psi(A_j) \leq A_j$, and $\sigma|_J$ is a bijection. So $i_y \circ \psi^k(A_j) \leq A' \leq A_j$ for some inner automorphism $i_y : F \to F$. The $[\psi]$-periodicity of A' implies $(i_y \circ \psi^k)^m(A') \leq A'$ is conjugate to A' for some $m \geq 1$. But Lemma 1.3 says no finitely generated subgroup of F is conjugate to a proper subgroup of itself. Therefore, $(i_y \circ \psi^k)^m(A') = A'$ and, by injectivity of ψ, $i_y \circ \psi^k(A_j) = A' = A_j$. In particular, $A = A' = A_j$. As this holds for arbitrary components $A \in \mathcal{A}$, we get $\mathcal{A} \subset \mathcal{A}_J$. $\qquad \square$

It is obvious that the maximal $[\phi]$-fixed free factor system is proper exactly when ϕ is not surjective. The same holds for the $[\phi]$-elliptic free factor system:

Observation. *Let $\phi : F \to F$ be an injective endomorphism. The $[\phi]$-elliptic free factor system is*

1. *proper exactly when ϕ is not surjective.*

2. *trivial exactly when only the trivial conjugacy class has an infinite $[\phi]$-tail;*

3   Relative representatives

We will now use relative (weak) representatives as the basis for inductively studying dynamical properties of free group endomorphisms. For the whole section, we suppose $\phi : F \to F$ is an injective endomorphism and $\phi^{-1} \cdot \mathcal{A} = \mathcal{A}$, i.e. $\mathcal{A}$ is a $[\phi]$-invariant proper free factor system such that there is no free factor system $\mathcal{A}' \succ \mathcal{A}$ with an image $\phi(\mathcal{A}')$ carried by $\mathcal{A}$; e.g. consider an injective nonsurjective ϕ and its elliptic free factor system.

The first step is to establish the relative version of the bounded cancellation.

Lemma 3.1 (Bounded Cancellation)**.** *Let T and T' be $(F, \mathcal{A})$- and $(F', \mathcal{A}')$-trees respectively and $\psi : F \to F'$ be an injective homomorphism such that $\psi^{-1} \cdot \mathcal{A}' = \mathcal{A}$. If $g : T \to T'$ is a ψ-equivariant tree map, then there is a constant $C(g)$ such that for every natural edge-path decomposition $p_1 \cdot p_2$ of an immersed path in T, the edge-path $[g(p_1)] \cdot [g(p_2)]$ is contained in the $C(g)$-neighborhood of $[g(p_1) \cdot g(p_2)]$.*

Proof. The proof is the same as before. Since the trees are simplicial, the map g factors as an equivariant pretrivial edge collapse and subdivision, a composition of $r \geq 0$ equivariant folds, and an equivariant simplicial embedding. As ψ is injective and $\psi^{-1} \cdot \mathcal{A}' = \mathcal{A}$, no fold or collapse identifies vertices in the same orbit and, hence, the collapse has cancellation constant 0 and each fold has cancellation constant 1. We may choose $C(g) = r$. $\qquad \square$

Let $\mathcal{A} \preceq \mathcal{B}$ be a chain of $[\phi]$-invariant free factor systems, T_* be a $(\mathcal{B}, \mathcal{A})$-forest with no bivalent vertices, and $f_* : T_* \to T_*$ be a $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$. For all

$k \geq 1$, set $T_*(\phi^k(\mathcal{B})) \subset T_*$ to be the minimal subforest for $\phi^k(\mathcal{B})$; minimal subforests are the relative analogues of iterated Stallings graphs. We will assume the minimal subforests $T_*(\phi^k(\mathcal{B}))$ inherit their simplicial structure from the *ambient forest* T_* and so they might have bivalent vertices unlike T_*. For a graph of groups decomposition with bivalent vertices, branch points are vertices that are images of branch points of the *Bass-Serre* tree and natural edges are images of natural edges of the tree.

For any $k \geq 1$, let $f_{*,k} : T_*(\phi^k(\mathcal{B})) \to T_*$ be the restriction of f_* to $T_*(\phi^k(\mathcal{B}))$ and then replace it with an equivariantly homotopic map $T_*(\phi^k(\mathcal{B})) \to T_*(\phi^k(\mathcal{B}))$ that is induced by the deformation retraction of $f_{*,k}(T_*(\phi^k(\mathcal{B})))$ to $T_*(\phi^k(\mathcal{B}))$, which we call the (*k*-th) homotopy restriction of f_*. Note that if $X \subset T_*(\phi^k(\mathcal{B}))$ is an axis such that $f_*|_X$ is an immersion, then $f_{*,k}|_X$ is still an immersion.

These homotopy restrictions are the relative analogues of homotopy lifts. The proof of the following lemma is also almost the same as that of Lemma 1.5.

Lemma 3.2. *Let T and T' be $(F, \mathcal{A})$- and $(F', \mathcal{A}')$-trees respectively and $\psi : F \to F'$ be an injective homomorphism such that $\psi^{-1} \cdot \mathcal{A}' = \mathcal{A}$. If $g : T \to T'$ is a ψ-equivariant tree map with cancellation constant $C = C(g)$, then g maps branch points to the C-neighborhood of branch points.*

Proof. Set $C = C(g)$ and let ν be a bivalent vertex in T' whose distance to the nearest branch point is greater than C. In particular, ν has a trivial stabilizer. We denote by ϵ_1, ϵ_2 the 2 distinct directions at ν oriented away from the vertex. Suppose v is a branch point of T such that $g(v) = \nu$. As ψ is injective, v has a trivial stabilizer under the action of F. Choose 3 distinct directions at v: e_1, e_2, and e_3. Let p_{12} be an embedded path in T that starts with e_1 and ends with a translate of $\bar{e}_2$. Since v has a trivial stabilizer, the path determines a unique loxodromic element x_{12} in F with axis a_{12} such that p_{12} is a fundamental domain of the axis under the translation action of x_{12}. Without loss of generality, $[g(p_{12})]$ starts with ϵ_1.

If $[g(p_{12})]$ ends with the translate $\phi(x_{12})\bar{\epsilon}_1$, then $[g(p_{12})] = \mu \cdot \rho \cdot (\phi(x_{12})\bar{\mu})$, where μ is an extension of ϵ_1 to an embedded path from ν to the axis of $\phi(x_{12})$ and ρ is a fundamental domain of the axis of loxodromic element $\phi(x_{12})$. By assumption, μ is longer than C. Decompose the axis $a_{12} = a_- \cdot a_+$ at v, then $[g(a_-)] \cdot [g(a_+)]$ has $\bar{\mu} \cdot \mu$ as a subpath, violating bounded cancellation. The remaining cases are handled similarly. Upon ruling out all cases, we conclude that no branch point v of T is mapped to ν. $\square$

As before, we get a corollary whose proof is essentially the same as that of Corollary 1.6.

Corollary 3.3. *Let T and T' be $(F, \mathcal{A})$- and $(F', \mathcal{A}')$-trees respectively and $\psi : F \to F'$ be an injective homomorphism such that $\psi^{-1} \cdot \mathcal{A}' = \mathcal{A}$. If $g : T \to T'$ is a ψ-equivariant K-Lipschitz tree map with cancellation constant C, then g is equivariantly homotopic to a ψ-equivariant $(K + C)$-Lipschitz natural tree map with cancellation constant $2C$.*

The corollary allows us to replace $f_{*,k}$ with an equivariantly homotopic $\phi^k(\mathcal{A})$-relative natural representative that has Lipschitz constant $K(f_{*,k}) = K(f_*) + C(f_*)$ and cancellation constant $C(f_{*,k}) = 2C(f_*)$. If $X \subset T_*(\phi^k(\mathcal{B}))$ is an axis and $f_{*,k}|_X$ is an immersion before the homotopy, then $f_{*,k}(X)$ is an immersed path after the homotopy; however, the restriction $f_{*,k}|_X$ may fail to be an immersion due to pretrivial edges. The following summary is a relative analogue of Proposition 1.7.

Proposition 3.4. Let $f_* : T_* \to T_*$ be an $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$. For any $k \geq 1$, there is an $\phi^k(\mathcal{A})$-relative natural representative $f_{*,k} : T_*(\phi^k(\mathcal{B})) \to T_*(\phi^k(\mathcal{B}))$ for $\phi|_{\phi^k(\mathcal{B})}$ such that:

1. $K(f_{*,k}) = K(f_*) + C(f_*)$ and $C(f_{*,k}) = 2C(f_*)$.

2. If f_* restricted to the axes of some conjugacy class $[b]$ in $\phi^k(\mathcal{B})$ is an immersion, then $f_{*,k}$ restricted to the axes of $[b]$ is also an immersion modulo pretrivial edges.

Collapsing a maximal (orbit-closed) $f_{*,k}$-invariant subforest of $T_*(\phi^k(\mathcal{B}))$ with bounded components and forgetting the bivalent vertices induces a minimal $\phi^k(\mathcal{A})$-relative representative $g_{*,k} : Y_{*,k} \to Y_{*,k}$ for $\phi|_{\phi^k(\mathcal{B})}$ defined on a $(\phi^k(\mathcal{B}), \phi^k(\mathcal{A}))$-forest $Y_{*,k}$. Note that the collapsed maximal subforest contains the pretrivial edges as $\phi^{-1} \cdot \mathcal{A} = \mathcal{A}$. Since $g_{*,k}$ is induced by equivariantly collapsing a forest and forgetting bivalent vertices, we have $K(g_{*,k}) = K(f_*) + C(f_*)$, $C(g_{*,k}) = 2C(f_*)$, and $l_{Y_{*,k}} \leq l_{T_*}|_{\phi^k(\mathcal{B})}$. If X is an axis of b in $T_*(\phi^k(\mathcal{B}))$ and $f_{*,k}|_X$ is an immersion modulo pretrivial edges, then $g_{*,k}|_{X'}$ is an immersion, where X' is the axis of b in $Y_{*,k}$.

Proposition 3.5. Let $f_* : T_* \to T_*$ be an $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$. For any $k \geq 1$, there is a $(\phi^k(\mathcal{B}), \phi^k(\mathcal{A}))$-forest $Y_{*,k}$ and a minimal $\phi^k(\mathcal{A})$-relative representative $g_{*,k} : Y_{*,k} \to Y_{*,k}$ for $\phi|_{\phi^k(\mathcal{B})}$ such that:

1. $K(g_{*,k}) = K(f_*) + C(f_*)$ and $C(g_{*,k}) = 2C(f_*)$.

2. $l_{Y_{*,k}} : \phi^k(\mathcal{B}) \to \mathbb{R}$ is dominated by $(\leq)$ the restriction $l_{T_*}|_{\phi^k(\mathcal{B})} = l_{T_*(\phi^k(\mathcal{B}))}$;

3. If f_* restricted to the axes of some conjugacy class $[b]$ in $\phi^k(\mathcal{B})$ is an immersion, then $g_{*,k}$ restricted to the axes of $[b]$ is also an immersion.

For any $\mathcal{A}$-relative weak representative f_*, we define the <u>transition matrix</u> $A(f_*)$: it is a nonnegative integer square matrix whose rows and columns are indexed by the orbits of edges in T_*; the entry of $A(f_*)$ in row-i and column-j, $A(f_*)(i,j)$, is given by the number of translates of e_i that are contained in the immersed edge-path $f_*(e_j)$, where e_i is a orbit representative for the i-th orbit of edges. An $\mathcal{A}$-relative weak representative f_* is <u>irreducible</u> if the matrix $A(f_*)$ is irreducible, i.e. for any pair (i,j), there is a positive integer n_{ij} such that $A(f_*)^{n_{ij}}(i,j) > 0$. In this case, the <u>stretch factor</u> of f_*, $\lambda(f_*) \geq 1$, is the *Perron-Frobenius eigenvalue* of $A(f_*)$. An irreducible $\mathcal{A}$-relative weak representative is <u>expanding</u> if $\lambda(f_*) > 1$. Note that irreducible $\mathcal{A}$-relative representatives are minimal.

We say $\phi|_{\mathcal{B}}$ is <u>irreducible relative to $\mathcal{A}$</u> if there is no $[\phi]$-invariant free factor system $\mathcal{C}$ such that $\mathcal{A} \prec \mathcal{C} \prec \mathcal{B}$. If $\mathcal{A}$ is trivial and $\mathcal{B} = \{F\}$, then we recover the definition of ϕ's irreducibility. The next lemma is the most useful property of an irreducible restriction for our purposes.

Lemma 3.6. *If $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$, then every minimal $\mathcal{A}$-relative weak representative for $\phi|_{\mathcal{B}}$ is irreducible.*

Proof. Suppose a minimal $\mathcal{A}$-relative weak representative for $\phi|_{\mathcal{B}}$ has a reducible transition matrix; in particular, it has an invariant $\mathcal{B}$-equivariant proper subforest (with unbounded components) that determines a $[\phi]$-invariant free factor system $\mathcal{C}$ such that $\mathcal{A} \prec \mathcal{C} \prec \mathcal{B}$. So $\phi|_{\mathcal{B}}$ is not irreducible relative to $\mathcal{A}$. $\qquad\square$

Remark. Bestvina-Handel give the absolute version of this property as the definition of irreducibility and then prove that it is equivalent to the definition of irreducibility given in this paper [6, Lemma 1.16]. The relative version of this equivalence holds as well but we will not prove it as it is not needed.

Bestvina-Handel used the next proposition to construct *train tracks* [6, Theorem 1.7].

Proposition 3.7. *If $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$, then there is an irreducible $\mathcal{A}$-relative representative $f_* : T_* \to T_*$ for $\phi|_{\mathcal{B}}$ with the minimal stretch factor, i.e. if $f'_* : T'_* \to T'_*$ is an irreducible $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$, then $\lambda(f'_*) \geq \lambda(f_*)$.*

The minimal stretch factor will be denoted by $\lambda([\phi], \mathcal{B}, \mathcal{A})$.

Proof. Let $g_* : Y_* \to Y_*$ be a minimal $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$ and suppose $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$. Then g_* is an irreducible $\mathcal{A}$-relative representative by Lemma 3.6 with stretch factor $\lambda(g_*)$. By the lack on bivalent vertices, any irreducible $\mathcal{A}$-relative representative has a transition matrix of size $\leq N = 3 \cdot \text{rank}(F) - 3$. Suppose B is an irreducible integer square matrix with Perron-Frobenius eigenvalue $\lambda(B) \leq \lambda(g_*)$. Then B has a positive right eigenvector $\vec{v}$ associated with $\lambda(B)$. So for all $k \geq 1$, B^k has right eigenvector $\vec{v}$ associated with eigenvalue $\lambda(B)^k$. Assuming the smallest entry of $\vec{v}$ is 1 (rescale if necessary), we get that the minimum row-sum of B^k is at most $\lambda(B)^k$ for any $k \geq 1$. If B has no more than N rows, then the largest entry of B is at most the minimum row-sum of $B^{N!}$, which we know is at most $\lambda(B)^{N!} \leq \lambda(g_*)^{N!}$. So there are finitely many irreducible integer square matrices with size $\leq N$ and Perron-Frobenius eigenvalue $\leq \lambda(g_*)$. Thus, there is a finite set of stretch factors $\leq \lambda(g_*)$ for irreducible $\mathcal{A}$-relative representatives for $\phi|_{\mathcal{B}}$. In particular, there is an irreducible $\mathcal{A}$-relative representative $f_* : T_* \to T_*$ for $\phi|_{\mathcal{B}}$ with the minimal stretch factor. $\qquad\square$

Bestvina-Handel's work [6] can be adapted to show that an irreducible $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$ with the minimal stretch factor is an *$\mathcal{A}$-relative train track* (Appendix A) and, conversely, bounded cancellation implies all irreducible $\mathcal{A}$-relative train tracks for $\phi|_{\mathcal{B}}$

have the minimal stretch factor. We do not prove this converse as it is not needed. The next lemma is an application of train track theory that will be invoked once, in the second half of the proof of Proposition 3.9.

Lemma 3.8 (Train Track Theory). *If $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$ and $f_* : T_* \to T_*$ is an irreducible $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$ with the minimal stretch factor, then there is an element g in $\mathcal{B}$ with an axis a_g such that the restriction of f_*^k to a_g is an immersion for all $k \geq 1$.*

Such an axis will be known as an $\underline{f_*\text{-legal axis}}$.

Proof. If $\lambda(f_*) = 1$, then f_* is a simplicial embedding and we are done. So we may assume $\lambda(f_*) > 1$. By minimality of its stretch factor, f_* is an expanding irreducible $\mathcal{A}$-relative *train track* for $\phi|_{\mathcal{B}}$ (Theorem A.1), i.e. for any edge e in T_*, $f_*^k(e)$ is an expanding immersed path for all $k \geq 1$. A 2-edge path $e_1 \cdot e_2$ is f_*-*legal* if it is a translate of a subpath of $f_*^k(e)$ for some edge e and integer $k \geq 1$. By irreducibility of f_*, every edge e is contained in a 3-edge path $e_- \cdot e \cdot e_+$ whose 2-edge subpaths are both f_*-legal. This means we can form an axis a_g whose 2-edge subpaths are all f_*-legal. By the train track property, the restriction of f_*^k to a_g is an immersion for all $k \geq 1$. $\square$

The main tools from Section 1 that were used in the previous section were Lemma 1.1, bounded cancellation, and Proposition 1.7. The relative analogues of the latter two have already been established in this section. We now state the main technical result of this section, an analogue of Lemma 1.1 — analogous in the sense that both give sufficient conditions for iterated subgroup graphs to have arbitrarily long natural edges.

Proposition 3.9. *Let $\mathcal{A} \prec \mathcal{B}$ be a chain of $[\phi]$-invariant free factor systems with $\phi|_{\mathcal{B}}$ irreducible relative to $\mathcal{A}$ and $\lambda([\phi], \mathcal{B}, \mathcal{A}) > 1$. If $\mathcal{A}$ carries the maximal $[\phi]$-fixed free factor system, then the length of the longest natural edge in $T_*(\phi^k(\mathcal{B}))$ is unbounded as $k \to \infty$.*

Before starting the proof, we will first describe the *(absolute) vertex blow-up* construction. Let $\mathcal{A} \prec \mathcal{B}$ be a chain of free factor systems and T_* be some $(\mathcal{B}, \mathcal{A})$-forest. Recall that we assume $\mathcal{A}_i \subset \mathcal{A}$ is also a free factor system of $B_i \in \mathcal{B}$. Fix some $\mathcal{A}$-marked roses $(R_{\mathcal{A}}, \alpha_{\mathcal{A}})$. Define $\Gamma_{\mathcal{B}}$ to be the graph formed by identifying the appropriate vertices of the graph of groups $\mathcal{B} \backslash T_*$ with the basepoints of roses $(R_{\mathcal{A}}, \alpha_{\mathcal{A}})$. If $c : R_{\mathcal{A}} \to \Gamma_{\mathcal{B}}$ is the inclusion map, then *Bass-Serre theory* gives markings $\alpha_{\mathcal{B}} = \{\alpha_i : B_i \to \pi_1(\Gamma_i)\}$ such that $[\pi_1(c) \circ \alpha_{\mathcal{A}}] = [(\alpha_{\mathcal{B}})|_{\mathcal{A}}]$. Thus, $(\Gamma_{\mathcal{B}}, \alpha_{\mathcal{B}})$ is a $\mathcal{B}$-marked graph. This construction and, in general any pair of graphs $\Gamma'_{\mathcal{A}} \subset \Gamma'_{\mathcal{B}}$ with collections of markings $\alpha'_{\mathcal{A}}, \alpha'_{\mathcal{B}}$ such that $[\pi_1(c') \circ \alpha'_{\mathcal{A}}] = [(\alpha'_{\mathcal{B}})|_{\mathcal{A}}]$ will be referred to as $\underline{\text{vertex blow-up}}$.

We note that the Stallings graph $S[\phi^k(\mathcal{B})]$ with respect to $(\Gamma_{\mathcal{B}}, \alpha_{\mathcal{B}})$, as a $\phi^k(\mathcal{B})$-marked graph, is a vertex blow-up of $\phi^k(\mathcal{B}) \backslash T_*(\phi^k(\mathcal{B}))$: let $\iota : S[\phi^k(\mathcal{B})] \to \Gamma_{\mathcal{B}}$ be the Stallings graph's immersion and $S_{\mathcal{A}} \subset S[\phi^k(\mathcal{B})]$ the core of subgraph $\iota_{\mathcal{B}}^{-1}(R_{\mathcal{A}})$. Since $\phi^{-1} \cdot \mathcal{A} = \mathcal{A}$, $S_{\mathcal{A}} = S[\phi^k(\mathcal{A})]$ is marked by an isomorphism $\alpha'_{\mathcal{A}} : \phi^k(\mathcal{A}) \to \pi_1(S_{\mathcal{A}})$ and $\alpha'_{\mathcal{A}}$ is the

restriction of the marking $\alpha'_{\mathcal{B}} : \phi^k(\mathcal{B}) \to \pi_1(S[\phi^k(\mathcal{B})])$ to $\phi^k(\mathcal{A})$ with respect to the inclusion $S_{\mathcal{A}} \subset S[\phi^k(\mathcal{B})]$. Therefore, $S[\phi^k(\mathcal{B})]$ is also a vertex blow-up of $\phi^k(\mathcal{B})\backslash T_*(\phi^k(\mathcal{B}))$. The noncontractible components of the subgraph $\iota_{\mathcal{B}}^{-1}(R_{\mathcal{A}})$ will be known as the *lower stratum* and the rest of the graph as the *top stratum*.

Now suppose $\mathcal{A} \prec \mathcal{B}$ are also $[\phi]$-invariant and let $f_* : T_* \to T_*$ be a $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$ defined on some $(\mathcal{B}, \mathcal{A})$-forest T_* and $f_{\mathcal{A}} : R_{\mathcal{A}} \to R_{\mathcal{A}}$ be a topological representative for $\phi|_{\mathcal{A}}$. Construct a topological representative $f_{\mathcal{B}} : \Gamma_{\mathcal{B}} \to \Gamma_{\mathcal{B}}$ for $[\phi|_{\mathcal{B}}]$ that agrees with $f_{\mathcal{A}}$ on the $\mathcal{A}$-marked roses $R_{\mathcal{A}}$ and induces f_* on the Bass-Serre forest T_* upon collapsing the roses $R_{\mathcal{A}}$.

For any $k \geq 1$, we let $g_{*,k} : Y_{*,k} \to Y_{*,k}$ be the minimal $\phi^k(\mathcal{A})$-relative representative for $\phi|_{\phi^k(\mathcal{B})}$ given by Proposition 3.5 using f_*; also let $\bar{f}_k : S[\phi^k(\mathcal{B})] \to S[\phi^k(\mathcal{B})]$ be the natural representative for $[\phi|_{\phi^k(\mathcal{B})}]$ given by Proposition 1.7 using $f_{\mathcal{B}}$. By Proposition 3.5(3), if an element b in $\mathcal{B}$ has an f_*-legal axis, then $\phi^k(b)$ has a $g_{*,k}$-legal axis. It can be arranged for $S[\phi^k(\mathcal{A})] \subset S[\phi^k(\mathcal{B})]$ to be $\bar{f}_k$-invariant and $\bar{f}_k$ to induce $g_{*,k}$ on the $(\phi^k(\mathcal{B}), \phi^k(\mathcal{A}))$-forest $Y_{*,k}$ upon collapsing a maximal invariant proper subgraph of $S[\phi^k(\mathcal{B})]$ containing $S[\phi^k(\mathcal{A})]$ and forgetting bivalent vertices.

Here is the idea behind the proof. By irreducibility of the restriction $\phi|_{\mathcal{B}}$, we may assume the map $g_{*,k}$ is an expanding irreducible representative for $\phi|_{\phi^k(\mathcal{B})}$. For the contrapositive, suppose the forests $T_*(\phi^k(\mathcal{B}))$ had uniformly bounded natural edges. There is a sequence of loxodromic elements b_k in $\phi^k(\mathcal{B})$ with uniformly bounded translation lengths $l_{T_*}(g_k)$. Now suppose that the vertex blow-up $S[\phi^k(\mathcal{B})]$ had natural edges with aribtrarily long top stratum subpaths. Bounded cancellation, the fact $\bar{f}_k$ induces $g_{*,k}$, and the irreducibility of $g_{*,k}$ imply $g_{*,k}$ is an expanding irreducible *immersion*. However, this contradicts the first assumption since $l_{Y_{*,k}} \leq l_{T_*(\phi^k(\mathcal{B}))}$. So the second supposition is false and the natural edges of $S[\phi^k(\mathcal{B})]$ have top stratum subpaths with uniformly bounded length. Using the Lipschitz property, expanding irreducibility of $g_{*,k}$, and existence of a $g_{*,k}$-legal axis (train track theory), we find uniformly bounded lower stratum paths in $S[\phi^k(\mathcal{B})]$ connecting the origin of any oriented top stratum subpath of a natural edge to another top stratum subpath of a natural edge. Consequently, we are able to build uniformly bounded immersed loops in $S[\phi^k(\mathcal{B})]$ that contain top stratum subpaths. This implies some loxodromic conjugacy class in $\mathcal{B}$ has an infinite $[\phi]$-tail. By Proposition 2.3, any conjugacy class in F with an infinite $[\phi]$-tail is carried by the maximal $[\phi]$-fixed free factor system. Thus the maximal $[\phi]$-fixed free factor system cannot be carried by $\mathcal{A}$ as it carries a loxodromic conjugacy class.

Proof of Proposition 3.9. Suppose $\mathcal{A} \prec \mathcal{B}$ are $[\phi]$-invariant free factor systems, $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$, $f_{\mathcal{A}} : R_{\mathcal{A}} \to R_{\mathcal{A}}$ is a topological representative for $\phi|_{\mathcal{A}}$ defined on $\mathcal{A}$-marked roses $(R_{\mathcal{A}}, \alpha_{\mathcal{A}})$, and $f_* : T_* \to T_*$ is an expanding irreducible $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$ with the minimal stretch factor $\lambda(f_*) > 1$ (Proposition 3.7). By Lemma 3.8, there is an element b in $\mathcal{B}$ with an f_*-legal axis. Set $(\Gamma_{\mathcal{B}}, \alpha_{\mathcal{B}})$ to be the vertex

blow-up of $\mathcal{B}\backslash T_*$ with respect to the $\mathcal{A}$-marked roses $(R_\mathcal{A}, \alpha_\mathcal{A})$. The discussion preceding the proof gives minimal $\phi^k(\mathcal{A})$-relative representatives $g_{*,k} : Y_{*,k} \to Y_{*,k}$ for $\phi|_{\phi^k(\mathcal{B})}$ and natural representatives $\bar{f}_k : S[\phi^k(\mathcal{B})] \to S[\phi^k(\mathcal{B})]$ for $[\phi|_{\phi^k(\mathcal{B})}]$ that have these properties: for all $k \geq 1$,

1. $\bar{f}_k$ induces $g_{*,k}$ on $Y_{*,k}$ upon collapsing the $f_\mathcal{B}$-invariant subgraph $R_\mathcal{A} \subset \Gamma_\mathcal{B}$;

2. $K = K(\bar{f}_k) = K(f_\mathcal{B}) + C(f_\mathcal{B})$ and $C = C(\bar{f}_k) = 2C(f_\mathcal{B})$;

3. $l_{Y_{*,k}} : \phi^k(\mathcal{B}) \to \mathbb{R}$ is dominated by the restrictions $l_{T_*}|_{\phi^k(\mathcal{B})} = l_{T_*(\phi^k(\mathcal{B}))}$; and

4. $\phi^k(b)$ has a $g_{*,k}$-legal axis.

The collection $\phi|_{\phi^k(\mathcal{B})}$ is conjugate to $\phi|_\mathcal{B}$ by injectivity of ϕ. So $\phi|_{\phi^k(\mathcal{B})}$ is irreducible relative to $\phi^k(\mathcal{A})$ and $\lambda(f_*)$ is the minimal stretch factor for $\phi|_{\phi^k(\mathcal{B})}$ relative to $\phi^k(\mathcal{A})$. Furthermore, the minimal $\phi^k(\mathcal{A})$-relative representatives $g_{*,k}$ are irreducible (Lemma 3.6) and $\lambda(g_{*,k}) \geq \lambda(f_*) > 1$ by the minimality of $\lambda(f_*)$.

Suppose for the contrapositive that there is a bound $L \geq 1$ such that all natural edges in $T_*(\phi^k(\mathcal{B}))$ are shorter than L for all $k \geq 1$. Then, for all $k \geq 1$, there is a loxodromic element b_k in $\phi^k(\mathcal{B})$ such that $l_{T_*}(b_k) \leq NL$, where $N = 3 \cdot \text{rank}(F) - 3$. Every edge E in $\Gamma_{*,k} = \phi^k(\mathcal{B})\backslash Y_{*,k}$ lifts to a $\phi^k(\mathcal{B})$-orbit of a natural edge E' in $Y_{*,k}$, which corresponds to a top stratum subpath $\bar{E}$ of a natural edge in $S[\phi^k(\mathcal{B})]$.

Claim. *The subpath $\bar{E}$ in $S[\phi^k(\mathcal{B})]$ has length $\leq C \cdot K^{N-1}$ for all edges E in $\Gamma_{*,k}$ and $k \geq 1$.*

Suppose, the graph $\Gamma_{*,k}$ has an edge E_0 whose corresponding subpath $\bar{E}_0$ in $S[\phi^k(\mathcal{B})]$ is longer than $C \cdot K^{N-1}$ for some $k \geq 1$. As we did in the proof of Theorem 2.1, we construct the set of *long* edges $\mathcal{E}$ by looking at all the edges of $\Gamma_{*,k}$ that are eventually mapped over E_0. Here, an edge E_1 in $\Gamma_{*,k}$ mapped over E_0 if there are lifts E_1' and E_0' in $Y_{*,k}$ such that $g_{*,k}$ maps E_1' over E_0'. Since $\bar{f}_k : S[\phi^k(\mathcal{B})] \to S[\phi^k(\mathcal{B})]$ is K-Lipschitz and it induces $g_{*,k}$ on $Y_{*,k}$, each long edge in $\Gamma_{*,k}$ corresponds to a top statrum subpath in $S[\phi^k(\mathcal{B})]$ longer than C. Since $g_{*,k}$ is an irreducible $\phi^k(\mathcal{A})$-relative representative, all edges eventually map over E_0 and hence are long. The long natural edges of $S[\phi^k(\mathcal{B})]$ will be the natural edges in $S[\phi^k(\mathcal{B})]$ containing top stratum subpaths.

Suppose an edge E of $\Gamma_{*,k}$ had a lift E' in $Y_{*,k}$ that is the initial segment of the $g_{*,k}$-image of two edges that share an initial vertex. Then the top stratum subpath $\bar{E}$ is in a long natural edge of $S[\phi^k(\mathcal{B})]$ that is the initial segment of $\bar{f}_k$-images of natural edges that share an initial vertex; this violates bounded cancellation since long natural edges of $S[\phi^k(\mathcal{B})]$ longer than $C = C(\bar{f}_k)$. Hence, there is no folding in $g_{*,k}$, i.e. $g_{*,k}$ is an expanding irreducible $\phi^k(\mathcal{A})$-relative *immersion*. We may now find an $m \geq 1$ such that all loxodromic elements b in $\phi^k(\mathcal{B})$ have $l_{Y_{*,k}}(\phi^m(b)) > NL$. Since $l_{Y_{*,k}}$ is dominated by $l_{T_*}|_{\phi^k(\mathcal{B})}$, we get that $l_{T_*}(b') > NL$ for all loxodromic elements b' in $\phi^{k+m}(\mathcal{B})$. This contradicts the assumption that $l_{T_*}(b_{k+m}) \leq NL$ for some loxodromic b_{k+m} in $\phi^{k+m}(\mathcal{B})$.

So the top stratum subpath $\bar{E}$ in $S[\phi^k(\mathcal{B})]$ has length $\leq C \cdot K^{N-1}$ for all natural edges E of $\Gamma_{*,k}$ and $k \geq 1$. This ends the proof of the claim.

Next, we prove the existence of paths in the lower stratum of $S[\phi^k(\mathcal{B})]$ with uniformly bounded lengths connecting top stratum paths. Suppose E_0, E_1, and E_2 are edges of $\Gamma_{*,k}$ with lifts E_0', E_1', and E_2' in $Y_{*,k}$ such that $E_1' \cdot E_2'$ is a subpath of the immersed path $g_{*,k}(E_0')$. Then $\bar{E}_1 \cdot P_{12} \cdot \bar{E}_2$ is a subpath of immersed path $\bar{f}_k(\bar{E}_0)$ for some lower stratum path P_{12} in $S[\phi^k(\mathcal{B})]$. Since $\bar{E}_0$ has length bounded by $C \cdot K^{N-1}$ and $\bar{f}_k$ is K-Lipschitz, the path P_{12} has length bounded by $C \cdot K^N$. We say the 2-edge path $E_1 \cdot E_2$ in $\Gamma_{*,k}$ has a *nondegenerate turn bounded by* $C \cdot K^N$.

As $g_{*,k}$ is an expanding irreducible relative representative that has a legal axis (this is where the argument invokes train track theory), every edge E' in $Y_{*,k}$ can be extended to an immersed 3-edge path $E_-' \cdot E' \cdot E_+'$ that is a translate of a subpath of $g_{*,k}^n(E')$ and $n \leq 2 \cdot N!$. In particular, any edge in $\Gamma_{*,k}$ can be extended to a 3-edge path whose 2-edge subpaths both have nondegenerate turns bounded by $C \cdot K^{N-1} \cdot K^{2 \cdot N!}$, i.e. every top stratum subpath $\bar{E}$ can be extended to an immersed path $\bar{E}_- \cdot P_- \cdot \bar{E} \cdot P_+ \cdot \bar{E}_+$ with top stratum subpaths $\bar{E}_-, \bar{E}_+$ and lower stratum paths P_-, P_+ whose lengths are bounded by $C \cdot K^{N-1} \cdot K^{2 \cdot N!}$.

Using this bound on lower stratum paths and the bound on top stratum subpaths given by the claim, we can now form an immersed loop ρ_k in $\Gamma_{*,k}$ with the properties:

1. ρ_k lifts to an axis in $Y_{*,k}$ for some loxodromic conjugacy class $[b_k']$;

2. ρ_k passes any edge of $\Gamma_{*,k}$ at most twice and only takes short turns (including the turn at the endpoint), which implies it has at most $2N$ edges and (short) turns; and

3. ρ_k corresponds to a loop in $S[\phi^k(\mathcal{B})]$ with length bounded by $2N \cdot C(1 + K^{2 \cdot N!}) K^{N-1}$.

In summary, for each $k \geq 1$, we construct a loxodromic conjugacy class $[b_k']$ in $\phi^k(\mathcal{B})$ whose $\alpha_{\mathcal{B}}$-length is bounded by a constant independent of k. As there are finitely many conjugacy classes with $\alpha_{\mathcal{B}}$-length bounded by any given constant, the sequence of conjugacy classes $[b_k']_{k=1}^\infty$ has a constant infinite subsequence. Thus, some loxodromic conjugacy class $[b']$ has an infinite $[\phi]$-tail carried by $\mathcal{B}$. Recall that the maximal $[\phi]$-fixed free factor system carries all conjugacy classes with an infinite $[\phi]$-tail (Proposition 2.3); on the other hand, $\mathcal{A}$ does not carry loxodromic conjugacy classes in $\mathcal{B}$ by definition. Therefore, $\mathcal{A}$ cannot carry the maximal $[\phi]$-fixed free factor system. □

4 Canonical expanding relative immersions

The main result of this section is the existence and uniqueness of expanding immersions for nonsurjective endomorphisms relative to their elliptic free factor systems.

Let $\phi : F \to F$ be an injective endomorphism, $\mathcal{A} \prec \mathcal{B}$ be a pair of $[\phi]$-invariant free factor systems such that $\phi^{-1} \cdot \mathcal{A} = \mathcal{A}$, and $\phi|_{\mathcal{B}} : \mathcal{B} \to \mathcal{B}$ be a restriction of ϕ to $\mathcal{B}$.

Recall that a relative weak representative is minimal if it has no orbit-closed invariant subforests with bounded components and an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$ is an $\mathcal{A}$-relative representative $f_* : T_* \to T_*$ for $\phi|_{\mathcal{B}}$ that is an immersion whose edges expand under f_*-iteration.

There will be two possible ways of obtaining a relative immersion from a relatively irreducible restriction with a minimal stretch factor λ. If $\lambda = 1$, then an irreducible representative with stretch factor λ is automatically a simplicial embedding. The next proposition shows how to construct an immersion when $\lambda > 1$. This construction is unique to nonsurjective endomorphisms because we require that the restriction be irreducible relative to a free factor system that carries the $[\phi]$-elliptic free factor system — when ϕ is an automorphism, the $[\phi]$-elliptic free factor system is $\{F\}$ and no such restriction exists.

Proposition 4.1. *Let $\phi : F \to F$ be injective and $\mathcal{A} \prec \mathcal{B}$ be a chain of $[\phi]$-invariant free factor systems that carry the $[\phi]$-elliptic free factor system. If $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$ and $\lambda([\phi], \mathcal{B}, \mathcal{A}) > 1$, then there is an expanding irreducible $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$.*

Proof. Suppose $\phi : F \to F$ is injective, $\mathcal{A} \prec \mathcal{B}$ are $[\phi]$-invariant free factor systems that carry the $[\phi]$-elliptic free factor system, $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$, and $f_* : T_* \to T_*$ is an expanding irreducible $(\mathcal{B}, \mathcal{A})$-relative representative for $\phi|_{\mathcal{B}}$ with minimal stretch factor $\lambda(f_*) > 1$. Set $K = K(f_*) + C(f_*)$ and $C = 2C(f_*)$. By Proposition 3.4, there is a natural representative $f_{*,k} : T_*(\phi^k(\mathcal{B})) \to T_*(\phi^k(\mathcal{B}))$ for $\phi|_{\phi^k(B)}$ with Lipschitz and cancellation constants $K(f_{*,k}) = K$ and $C(f_{*,k}) = C$ respectively for all $k \geq 1$.

The first part of the proof proceeds as a relativized version of the proof of Theorem 2.1. By Proposition 3.9, we may fix $k \gg 0$ such that the set of natural edges $\mathcal{L}_0$ in $T_*(\phi^k(\mathcal{B}))$ longer than $C \cdot K^{N-1}$ is not empty, where $N = 3 \cdot \mathrm{rank}(F) - 3$. Choose $\mathcal{L}$ to be the set of all natural edges that eventually get mapped over those in $\mathcal{L}_0$ by $f_{*,k}$ and call $\mathcal{L}$ the long natural edges. As $f_{*,k}$ is K-Lipschitz and there are at most N orbits of natural edges in $T_*(\phi^k(\mathcal{B}))$, the long natural edges are longer that C. Injectivity of ϕ implies $\phi|_{\phi^k(\mathcal{B})}$ is conjugate to $\phi|_{\mathcal{B}}$. So $\phi|_{\phi^k(\mathcal{B})}$ is irreducible relative to $\phi^k(\mathcal{A})$, $\lambda(f_*)$ is the minimal stretch factor for $\phi|_{\phi^k(\mathcal{B})}$ relative to $\phi^k(\mathcal{A})$, and the short natural edges of $T_*(\phi^k(\mathcal{B}))$ form an orbit-closed $f_{*,k}$-invariant subforest with bounded components.

Collapse a maximal $f_{*,k}$-invariant subforest of $T_*(\phi^k(\mathcal{B}))$ that has bounded components and contains the short natural edges then forget the bivalent vertices; this induces a minimal $\phi^k(\mathcal{A})$-relative representative $g_{*,k} : Y_{*,k} \to Y_{*,k}$ for $\phi|_{\phi^k(\mathcal{B})}$. The map $g_{*,k}$ is an irreducible $\phi^k(\mathcal{A})$-relative representative for $\phi|_{\phi^k(\mathcal{B})}$ (Lemma 3.6) and $\lambda(g_{*,k}) \geq \lambda(f_*)$ by the minimality of $\lambda(f_*)$. So $g_{*,k}$ is an expanding irreducible $\phi^k(\mathcal{A})$-relative representative.

Since the lifts in $T_*(\phi^k(\mathcal{B}))$ of all edges in $Y_{*,k}$ are longer than the cancellation constant C, there is no folding in $g_{*,k}$ — otherwise, there would be folding in $f_{*,k}$ identifying paths longer than its cancellation constant, absurd. Thus, $g_{*,k}$ is an immersion. By injectivity of ϕ, we can view $Y_{*,k}$ as a $(\mathcal{B}, \mathcal{A})$-forest and $g_{*,k}$ as an expanding irreducible $\mathcal{A}$-relative

immersion for $\phi|_{\mathcal{B}}$. $\square$

We are now ready to state and prove our base case for the construction. In light of the previous proposition, the point is that a restriction $\phi|_{\mathcal{B}}$ that is irreducible relative to the $[\phi]$-elliptic free factor system $\mathcal{A}$ will satisfy $\lambda([\phi], \mathcal{B}, \mathcal{A}) > 1$.

Proposition 4.2. *Let $\phi : F \to F$ be injective and $\mathcal{A} \prec \mathcal{B}$ be a chain of $[\phi]$-invariant free factor systems where $\mathcal{A}$ is the $[\phi]$-elliptic free factor system. If $\phi|_{\mathcal{B}}$ is irreducible relative to $\mathcal{A}$, then there is an expanding irreducible $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$.*

Proof. Let $\phi : F \to F$ be injective and $\phi|_{\mathcal{B}}$ be irreducible relative to the $[\phi]$-elliptic free factor system $\mathcal{A}$. Then there is an irreducible $\mathcal{A}$-relative representative $f_* : T_* \to T_*$ for $\phi|_{\mathcal{B}}$ with stretch factor $\lambda(f_*) = \lambda([\phi], \mathcal{B}, \mathcal{A}) \geq 1$ (Proposition 3.7). We say $B_i \in \mathcal{B}$ is loxodromic if $T_i \in T_*$ is not a point, i.e. B_i contains a loxodromic element; similarly, the component $B_i \backslash T_i$ of the graph of groups $\mathcal{B} \backslash T_*$ is loxodromic if B_i is loxodromic. If $\lambda(f_*) = 1$, then the induced map on the loxodromic components of $\mathcal{B} \backslash T_*$ is a graph isomorphism. So for some $k \geq 1$, if $A \in \mathcal{A}$ is carried by a loxodromic $B \in \mathcal{B}$, then A is ϕ^k-invariant. By Proposition 2.4, the subset of all $A \in \mathcal{A}$ carried by the loxodromic component of $\mathcal{B}$ form a $[\phi]$-fixed free factor subsystem. As f_* induces a graph isomorphism on the loxodromic components of $\mathcal{B} \backslash T_*$ and these components' vertex groups form a $[\phi]$-fixed free factor system, we get that f_* is surjective when restricted to the unbounded components of the forest T_* and the loxodromic components of $\mathcal{B}$ form a $[\phi]$-fixed free factor system. This is a contradiction since $[\phi]$-periodic free factors are elliptic (Propositions 2.3 and 2.4). Therefore, $\lambda(f_*) > 1$ and the result follows from Proposition 4.1. $\square$

Specializing this proposition to the case where ϕ is irreducible and nonsurjective yields an alternate proof to a result due to Reynolds.

Corollary 4.3 ([28, Corollary 3.23])**.** *If $\phi : F \to F$ is irreducible but not surjective, then ϕ is induced by an expanding irreducible graph immersion.*

Remark. This proof of Reynolds' result is a variation of our previous proof [25, Theorem 4.5] with two crucial differences: 1) it makes no use of *limit trees* in the *compactification of outer space*; 2) the specialization of Proposition 4.1 to irreducible nonsurjective endomorphisms need not invoke train track theory since we can use Lemma 1.1 in place of Proposition 3.9.

The next proposition is the induction step for our construction.

Proposition 4.4. *Let $\phi : F \to F$ be injective, $\mathcal{A}$ be the $[\phi]$-elliptic free factor system, and $\mathcal{A} \prec \mathcal{B} \prec \mathcal{C}$ be a chain of $[\phi]$-invariant free factor systems. If there is an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$ and a $\mathcal{B}$-relative immersion for $\phi|_{\mathcal{C}}$, then there is an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{C}}$.*

Although the proof gets a bit technical, the idea is rather simple: a $\mathcal{B}$-relative immersion for $\phi|_{\mathcal{C}}$ (top stratum) and an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$ (lower stratum) can be patched together via a *(relative) vertex blow-up* to get a minimal $\mathcal{A}$-relative representative $g_* : Y_* \to Y_*$ whose only possible folds would have to happen between adjacent oriented of Y_* whose g_*-images have the same lower-stratum initial segment. As the restriction of g_* the lower stratum is an expanding immersion, we may assume the edges in the lower stratum are longer than the cancellation constant. This means no lower stratum edge is the initial segment of edges identified by a fold and so no folding in g_* is possible. Thus, g_* is a minimal $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{C}}$, which will be expanding if $\mathcal{A}$ is the $[\phi]$-elliptic free factor system.

Let $T_{\mathcal{C}}$ be a $(\mathcal{C}, \mathcal{B})$-forest and $T_{\mathcal{B}}$ be a $(\mathcal{B}, \mathcal{A})$-forest. For any component $C_i \in \mathcal{C}$, let $\mathcal{B}_i$ the maximal subset of $\mathcal{B}$ that is carried by C_i. Replace the components of $\mathcal{B}$ with conjugates if necessary and assume $\mathcal{B}$ is also a free factor system of $\mathcal{C}$. In particular, the components $B \in \mathcal{B}_i$ are free factors of the component $C_i \in \mathcal{C}$. Identifying the appropriate vertices of the graph of groups $\mathcal{C} \backslash T_{\mathcal{C}}$ with basepoints on the graph of groups $\mathcal{B} \backslash T_{\mathcal{B}}$ results in a graph of groups decomposition for $\mathcal{C}$ whose Bass-Serre forest T_* is a $(\mathcal{C}, \mathcal{A})$-forest that contains $T_{\mathcal{B}}$. We call T_* the vertex blow-up of $T_{\mathcal{C}}$ with respect to $T_{\mathcal{B}}$.

Proof of Proposition 4.4. Let $\phi : F \to F$ be injective, $\mathcal{B} \prec \mathcal{C}$ be a chain of $[\phi]$-invariant free factor systems that carry the $[\phi]$-elliptic free factor system $\mathcal{A}$. Suppose $f_{\mathcal{B}} : T_{\mathcal{B}} \to T_{\mathcal{B}}$ is an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}}$ and $f_{\mathcal{C}} : T_{\mathcal{C}} \to T_{\mathcal{C}}$ is a $\mathcal{B}$-relative immersion for $\phi|_{\mathcal{C}}$ then define T_* to be the vertex blow-up of $T_{\mathcal{C}}$ with respect to $T_{\mathcal{B}}$. The edges of T_* are of two types: the lower stratum, which are edges that are contained in the $\mathcal{C}$-orbit of $T_{\mathcal{B}}$, and the top stratum, which are the remaining edges.

Let $f_* : T_* \to T_*$ be a minimal $\mathcal{A}$-relative representative for $\phi|_{\mathcal{C}}$ such that $T_{\mathcal{B}}$ is an f_*-invariant subforest, the restriction of f_* to $T_{\mathcal{B}}$ agrees with $f_{\mathcal{B}}$, and f_* induces $f_{\mathcal{C}}$ upon collapsing the lower stratum. For all $k \geq 1$, set $T_*(\phi^k(\mathcal{C}))$ and $T_*(\phi^k(\mathcal{B}))$ to be the minimal subforests of T_* for $\phi^k(\mathcal{C})$ and $\phi^k(\mathcal{B})$ respectively. Similarly, define the minimal subforest $T_{\mathcal{B}}(\phi^k(\mathcal{B})) \subset T_{\mathcal{B}}$. By the inclusion of $T_{\mathcal{B}}$ in T_*, we get a simplicial identification of $T_*(\phi^k(\mathcal{B}))$ with $T_{\mathcal{B}}(\phi^k(\mathcal{B}))$. However, we want to consider these two forests differently with respect to their branch points and natural edges. In particular, there may be branch points of $T_*(\phi^k(\mathcal{C}))$ that are bivalent when considered as points on the subforest $T_*(\phi^k(\mathcal{B}))$. So by *"natural edges of $T_*(\phi^k(\mathcal{B}))$"*, we mean those inherited from the parent forest $T_*(\phi^k(\mathcal{C}))$; on the other hand, by *"natural edges of $T_{\mathcal{B}}(\phi^k(\mathcal{B}))$"*, we do mean exactly that. Under the identification of the two forests, the natural edges of $T_*(\phi^k(\mathcal{B}))$ partition any natural edge of $T_{\mathcal{B}}(\phi^k(\mathcal{B}))$ into at most $2\mathcal{X}$ segments, where $\mathcal{X} = \operatorname{rank}(F) - 1$.

Since $f_{\mathcal{C}} : T_{\mathcal{C}} \to T_{\mathcal{C}}$ is a $\mathcal{B}$-relative immersion for $\phi|_{\mathcal{C}}$, the restrictions of $f_{\mathcal{C}}$ to $T_{\mathcal{C}}(\phi^k(\mathcal{C}))$ are $\phi^k(\mathcal{B})$-relative immersions $f_{\mathcal{C},k} : T_{\mathcal{C}}(\phi^k(\mathcal{C})) \to T_{\mathcal{C}}(\phi^k(\mathcal{C}))$ for $\phi|_{\phi^k(\mathcal{C})}$ that are conjugate to $f_{\mathcal{C}}$. As f_* induces $f_{\mathcal{C}}$ upon collapsing the lower stratum, any edges in $f_*(T_*(\phi^k(\mathcal{C})))$ but not $T_*(\phi^k(\mathcal{C}))$ must be in the lower stratum and the restriction of f_* to $T_*(\phi^k(\mathcal{C}))$ induces

$f_{\mathcal{C},k}$ upon collapsing the lower stratum, i.e. the $\phi^k(\mathcal{C})$-orbit of $T_*(\phi^k(\mathcal{B}))$. By Proposition 3.4, there is an $\phi^k(\mathcal{A})$-relative natural representative $f_{*,k} : T_*(\phi^k(\mathcal{C})) \to T_*(\phi^k(\mathcal{C}))$ for $\phi|_{\phi^k(\mathcal{C})}$ with Lipschitz and cancellation constants $K = K(f_*) + C(f_*)$ and $C = 2C(f_*)$ respectively. Furthermore, $T_*(\phi^k(\mathcal{B}))$ is an $f_{*,k}$-invariant subforest and $f_{*,k}$ still induces $f_{\mathcal{C},k}$ upon collapsing the lower stratum. As f_* agrees with the immersion $f_\mathcal{B}$ on $T_\mathcal{B}$, $f_{*,k}$ differs from $f_\mathcal{B}$ on $T_*(\phi^k(\mathcal{B}))$ by a homotopy supported in the natural edges of $T_\mathcal{B}(\phi^k(\mathcal{B}))$.

Since $f_\mathcal{B} : T_\mathcal{B} \to T_\mathcal{B}$ is an expanding $\mathcal{A}$-relative immersion for $\phi|_\mathcal{B}$, the minimal subforest $T_\mathcal{B}(\phi^k(\mathcal{B}))$ has natural edges whose lengths are all exponential in k. Fix $k \gg 0$ such that all natural edges of $T_\mathcal{B}(\phi^k(\mathcal{B}))$ are longer than $2\mathcal{X} \cdot C \cdot K^{3\mathcal{X}-1}$. By the pigeonhole principle, each natural edge of $T_\mathcal{B}(\phi^k(\mathcal{B}))$ contains a natural edge of $T_*(\phi^k(\mathcal{B}))$ longer than $C \cdot K^{3\mathcal{X}-1}$. Let $\mathbb{G}_k$ be the directed graph of natural edges of the $f_{*,k}$-invariant subforest $T_*(\phi^k(\mathcal{B}))$ where a directed edge $E_i \to E_j$ corresponds to $f_{*,k}$ mapping E_i over E_j. Set $\mathcal{S}_0$ to be those natural edges with length at most C and $\mathcal{S}$ to be those natural edges with directed path from $\mathcal{S}_0$ in $\mathbb{G}_k$ and their $\phi^k(\mathcal{C})$-translates; these lower stratum natural edges will be the *short* natural edges of $T_*(\phi^k(\mathcal{C}))$. Since $f_{*,k}$ is K-Lipschitz and the shortest path between any two natural edges in $\mathbb{G}_k$ has $3\mathcal{X}$ natural edges, all the short natural edges have length at most $C \cdot K^{3\mathcal{X}-1}$. So the short natural edges $\mathcal{S}$ form an orbit-closed $f_{*,k}$-invariant lower stratum subforest of $T_*(\phi^k(\mathcal{C}))$ with bounded components as $\mathcal{S}$ does not cover any natural edge of $T_\mathcal{B}(\phi^k(\mathcal{B}))$.

Collapsing the short natural edges of $T_*(\phi^k(\mathcal{B}))$ induces a map $g'_{*,k} : Y'_{*,k} \to Y'_{*,k}$ with the same cancellation constant $C(g'_{*,k}) = C$. Now iteratively collapse pretrivial edges until the induced map $g_{*,k} : Y_{*,k} \to Y_{*,k}$ has none. As the collapses are supported in the lower stratum, the new map $g_{*,k}$ still induces the immersion $f_{\mathcal{C},k}$ upon collapsing the rest of the lower stratum and, as a result, folding in $g_{*,k}$ may only occur between initial segments of natural edges of $Y_{*,k}$ whose $g_{*,k}$-images are a lower stratum natural edge. However, all natural edges in the lower stratum of $Y_{*,k}$ are longer than C by construction and so no folding in $g_{*,k}$ is possible by bounded cancellation, i.e. $g_{*,k}$ is an immersion.

Collapsing a maximal invariant subforest with bounded components and forgetting bivalent vertices if necessary, we may assume $g_{*,k} : Y_{*,k} \to Y_{*,k}$ is a minimal $\phi^k(\mathcal{A})$-relative immersion for $\phi|_{\phi^k(\mathcal{C})}$. By the injectivity of ϕ, we can view $g_{*,k}$ as a minimal $\mathcal{A}$-relative immersion for $\phi|_\mathcal{C}$. It remains to show that every edge of $Y_{*,k}$ expands under $g_{*,k}$-iteration.

For a contradiction, suppose there is an edge of $Y_{*,k}$ whose $g_{*,k}$-iterates have uniformly bounded length. Since $g_{*,k}$ is minimal, the non-expanding edges in the graph of groups $\mathcal{C}\backslash Y_{*,k}$ contain a setwise-fixed subgraph $\mathcal{F}$ that carries a loxodromic element. The subgraph $\mathcal{F}$ is a free splitting of a $[\phi]$-invariant free factor system. Recall that the $[\phi]$-elliptic free factor system $\mathcal{A}$ decomposes as a union of the maximal $[\phi]$-fixed free factor system and components that eventually get mapped into this fixed system (Proposition 2.4). By construction, the nontrivial point stabilizers of $Y_{*,k}$ are conjugates of $\mathcal{A}$-components. So any vertex of the fixed graph $\mathcal{F}$ is labelled by either the trivial group or a component of the maximal $[\phi]$-fixed free factor system. Thus $\mathcal{F}$ is a free splitting of a $[\phi]$-fixed free factor

system that carries some loxodromic element. However, Propositions 2.3 and 2.4 imply that all $[\phi]$-fixed free factor systems are elliptic — a contradiction. □

We are now ready to inductively construct expanding relative immersions.

Theorem 4.5. *If $\phi : F \to F$ is injective but not surjective, then there is an expanding $\mathcal{A}$-relative immersion for ϕ, where $\mathcal{A}$ is the $[\phi]$-elliptic free factor system.*

Proof. Suppose $\phi : F \to F$ is injective but not surjective. Then the $[\phi]$-elliptic free factor system $\mathcal{A}$ is proper by Proposition 2.4. The naive approach is to assume there exists a chain $\mathcal{A} = \mathcal{B}_0 \prec \cdots \prec \mathcal{B}_n = \{F\}$ in the poset of $[\phi]$-invariant free factor system such that the restrictions $\phi|_{\mathcal{B}_{m+1}}$ are irreducible relative to $\mathcal{B}_m$ for all $m \geq 1$. This assumption is typical when working with automorphisms. For each restriction $\phi|_{\mathcal{B}_{m+1}}$, if the minimal stretch factor $\lambda([\phi], \mathcal{B}_{m+1}, \mathcal{B}_m) = 1$, then there is automatically a $\mathcal{B}_m$-relative immersion for $\phi|_{\mathcal{B}_{m+1}}$; and if $\lambda([\phi], \mathcal{B}_{m+1}, \mathcal{B}_m) > 1$, then there is an expanding $\mathcal{B}_m$-relative immersion for $\phi|_{\mathcal{B}_{m+1}}$ by Proposition 4.1. In either case, there is a $\mathcal{B}_m$-relative immersion for $\phi|_{\mathcal{B}_{m+1}}$. By Proposition 4.2, $\lambda([\phi], \mathcal{B}_1, \mathcal{B}_0) > 1$ and there is an expanding $\mathcal{B}_0$-relative immersion for $\phi|_{\mathcal{B}_1}$. By inductively patching these immersions together using Proposition 4.4, we get an expanding $\mathcal{B}_0$-relative immersion for ϕ and we are done. Unfortunately, since ϕ is not surjective, it could be that no chain $\mathcal{A} = \mathcal{B}_0 \prec \cdots \prec \mathcal{B}_n = \{F\}$ satisfies the naive assumption we made at the start. Recall that our definition of $\phi|_{\mathcal{B}_{m+1}}$ being irreducible relative to $\mathcal{B}_m$ presupposed $\phi^{-1} \cdot \mathcal{B}_m = \mathcal{B}_m$. Fortunately, this is a minor complication that can be fixed. The proof follows the approach described above closely but uses a chain with slightly weaker conditions on it.

We first construct a chain $\mathcal{A} = \mathcal{B}_0 \prec \cdots \prec \mathcal{B}_n = \{F\}$ in the poset of $[\phi]$-invariant free factor system that we will induct on. Let $\mathcal{A} \prec \mathcal{B}_1$ be a chain of $[\phi]$-invariant free factor systems such that $\phi|_{\mathcal{B}_1}$ is irreducible relative to $\mathcal{A}$. Suppose $\mathcal{B}_m$ has been constructed for some $m \geq 1$ and let $\mathcal{C} \succeq \mathcal{B}_m$ be the maximal free factor system in the chain $\mathcal{B}_m \preceq \phi^{-1} \cdot \mathcal{B}_m \preceq \phi^{-2} \cdot \mathcal{B}_m \preceq \cdots$ of $[\phi]$-invariant free factor systems. If $\phi^{-1} \cdot \mathcal{B}_m = \mathcal{B}_m = \mathcal{C}$, then let $\mathcal{B}_m \prec \mathcal{B}_{m+1}$ be a chain of $[\phi]$-invariant free factor systems such that $\phi|_{\mathcal{B}_{m+1}}$ is irreducible relative to $\mathcal{B}_m$. If $\mathcal{B}_m \prec \mathcal{C}$, then let $\mathcal{B}_m \prec \cdots \prec \mathcal{B}_{m+k} = \mathcal{C}$ be the chain of $[\phi]$-invariant free factor systems given by $\mathcal{B}_{m+i} = \phi^{-i} \cdot \mathcal{B}_m$ for $1 \leq i \leq k$.

We proceed by inducting on the resulting chain between $\mathcal{A}$ and $\{F\}$. For the base case, $\phi|_{\mathcal{B}_1}$ is irreducible relative to $\mathcal{A}$; therefore, there is an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}_1}$ by Proposition 4.2. For our induction hypothesis, suppose that there is an expanding $\mathcal{A}$-relative immersion $f_{\mathcal{B}_m} : T_{\mathcal{B}_m} \to T_{\mathcal{B}_m}$ for $\phi|_{\mathcal{B}_m}$ for some $m \geq 1$. By our construction of the chain, either $\phi|_{\mathcal{B}_{m+1}}$ is irreducible relative to $\mathcal{B}$ or $\phi(\mathcal{B}_{m+1})$ is carried by $\mathcal{B}_m$. We deal with these two cases separately.

Case 1. Suppose $\phi|_{\mathcal{B}_{m+1}}$ is irreducible relative to $\mathcal{B}_m$. By Proposition 3.7, there is an irreducible $\mathcal{B}_m$-relative representative $f_* : T_* \to T_*$ for $\phi|_{\mathcal{B}_{m+1}}$ with minimal stretch factor. If $\lambda(f_*) = 1$, then the simplicial embedding $f_{\mathcal{B}_{m+1}} = f_*$ is a $\mathcal{B}_m$-relative immersion for

32

$\phi|_{\mathcal{B}_{m+1}}$. If $\lambda(f_*) > 1$, then there is an expanding $\mathcal{B}$-relative immersion $f_{\mathcal{B}_{m+1}}$ for $\phi|_{\mathcal{B}_{m+1}}$ by Proposition 4.1. In either case, we get a $\mathcal{B}_m$-relative immersion $f_{\mathcal{B}_{m+1}} : T_{\mathcal{B}_{m+1}} \to T_{\mathcal{B}_{m+1}}$ for $\phi|_{\mathcal{B}_{m+1}}$ defined on a $(\mathcal{B}_{m+1}, \mathcal{B}_m)$-forest $T_{\mathcal{B}_{m+1}}$. Thus, there is an expanding $\mathcal{A}$-relative immersion for $\phi|_{\mathcal{B}_{m+1}}$ by the induction hypothesis and Proposition 4.4.

Case 2. Now suppose $\phi(\mathcal{B}_{m+1})$ is carried by $\mathcal{B}_m$. Let $T_{\mathcal{B}_m}(\phi(\mathcal{B}_{m+1})) \subset T_{\mathcal{B}_m}$ be the minimal subforest for $\phi(\mathcal{B}_{m+1})$ and $T_{\mathcal{B}_{m+1}}$ be the same forest after forgetting bivalent vertices. By injectivity of ϕ, we may consider $T_{\mathcal{B}_{m+1}}$ as a $(\mathcal{B}_{m+1}, \mathcal{A})$-forest that comes with a natural $\phi|_{\mathcal{B}_{m+1}}$-equivariant immersion $g : T_{\mathcal{B}_{m+1}} \to T_{\mathcal{B}_m}$. Since $f_{\mathcal{B}_m} : T_{\mathcal{B}_m} \to T_{\mathcal{B}_m}$ is an immersion, we can identify a subdivision of $T_{\mathcal{B}_m}$ with the minimal subforest $T_{\mathcal{B}_{m+1}}(\mathcal{B}_m) \subset T_{\mathcal{B}_{m+1}}$ of $\mathcal{B}_m$. Composing g with the subdivision and inclusion $T_{\mathcal{B}_{m+1}}(\mathcal{B}_m) \subset T_{\mathcal{B}_{m+1}}$ gives an $\mathcal{A}$-relative immersion $f_{\mathcal{B}_{m+1}} : T_{\mathcal{B}_{m+1}} \to T_{\mathcal{B}_{m+1}}$ for $\phi|_{\mathcal{B}_{m+1}}$, which is expanding since its image lies in $T_{\mathcal{B}_{m+1}}(\mathcal{B}_m)$ and its restriction to $T_{\mathcal{B}_{m+1}}(\mathcal{B}_m)$ is the expanding $\mathcal{A}$-relative immersion $f_{\mathcal{B}_m}$ after forgetting bivalent vertices. $\qquad\square$

It will follow from bounded cancellation that expanding relative immersion from the theorem is canonical.

Proposition 4.6. *Suppose $\phi : F \to F$ is injective and there is an expanding $\mathcal{A}$-relative immersion for ϕ, where $\mathcal{A}$ is a $[\phi]$-invariant free factor system. Then there is a unique expanding $\mathcal{A}$-relative immersion for ϕ.*

Proof. Let $\phi : F \to F$ be injective, $\mathcal{A}$ a $[\phi]$-invariant free factor system. Suppose $f : T \to T$ and $f' : T' \to T'$ are expanding $\mathcal{A}$-relative immersions for ϕ. The goal is to show that T and T' are equivariantly isomorphic as simplicial trees. Let $g : T \to T'$ be an equivariant tree map, i.e. ψ-equivariant tree map with $\psi : F \to F$ being the identity automorphism; such a map always exists between $(F, \mathcal{A})$-trees. By taking restrictions to the $\phi^k(F)$-minimal subtrees $T(\phi^k(F))$ and $T'(\phi^k(F))$ and applying deformation retractions, we get an equivariant tree map $g_k : T(\phi^k(F)) \to T'(\phi^k(F))$ with cancellation constant $C(g)$. By Corollary 3.3, we assume g_k is an equivariant natural tree map with cancellation constant $2C(g)$.

Recall that f' is an expanding $\mathcal{A}$-relative immersion for ϕ, so all the natural edges of $T'(\phi^k(F))$ are longer than $2C(g)$ for large enough k. Fix $k \gg 0$, then no folding occurs in g_k by bounded cancellation. So g_k is a subdivision composed with forest collapse and $T'(\phi^k(F))$ has at most the same number of orbits of branch points as $T(\phi^k(F))$. By the same argument, $T(\phi^k(F))$ is an equivariant subdivision and forest collapse of $T'(\phi^k(F))$ and hence has at most the same number of orbits of branch points as $T'(\phi^k(F))$. This implies the two minimal subtrees have the same number of orbits of branch points and g_k is an equivariant simplicial isomorphism after forgetting bivalent vertices. Since f and f' are immersions, $T(\phi^k(F))$ and $T'(\phi^k(F))$ are ϕ^k-equivariant subdivisions of T and T' respectively. Therefore, T and T' are equivariantly isomorphic by injectivity of ϕ. So f and f' are homotopic immersions and hence agree up to *isotopy.* $\qquad\square$

Corollary 4.7. *If $\phi : F \to F$ is injective but not surjective, then there is a unique expanding $\mathcal{A}$-relative immersion for ϕ, where $\mathcal{A}$ is the $[\phi]$-elliptic free factor system.*

This gives us a complete characterization of when an injective endomorphism is induced by an expanding graph immersion.

Corollary 4.8. *Let $\phi : F \to F$ be an injective endomorphism. Then the following conditions are equivalent:*

1. *$[\phi]$ is induced by a unique expanding graph immersion;*

2. *$[\phi]$ is induced by an expanding graph immersion;*

3. *the trivial conjugacy class is the only conjugacy class with an infinite $[\phi]$-tail;*

4. *the trivial system is the only $[\phi]$-fixed free factor system.*

Proof. By Theorem 2.1, we have (3) $\iff$ (4). By Proposition 4.6, we get (1) $\iff$ (2). We leave the implication (2) $\implies$ (3) as an exercise. It remains to show (3) $\implies$ (2). Suppose the trivial conjugacy class is the only conjugacy class with an infinite $[\phi]$-tail. So ϕ is not surjective and the $[\phi]$-elliptic free factor system is trivial. By Theorem 4.5, there is an expanding ϕ-equivariant immersion $f : T \to T$ defined on a free F-tree T, i.e. $[\phi]$ is induced by an expanding immersion of the marked graph $F \backslash T$. $\qquad \square$

Interlude

Let us summarize the main results of the first part of the paper.

Summary. *If $\phi : F \to F$ is injective but not surjective, then there is:*

1. *a unique maximal proper $[\phi]$-fixed free factor system $\mathcal{A}$;*   (Proposition 2.3)

2. *a unique maximal proper $[\phi]$-invariant free factor system $\mathcal{A}^* \succeq \mathcal{A}$ such that $\phi^k(\mathcal{A}^*)$ is carried by $\mathcal{A}$ for some $k \geq 0$; after replacing the components of $\mathcal{A}$ with conjugates if necessary, we can assume $\mathcal{A} \subset \mathcal{A}^*$;*   (Proposition 2.4)

3. *a unique expanding $\mathcal{A}^*$-relative immersion for ϕ.*   (Corollary 4.7)

We call the free factor system $\mathcal{A}^*$ the $[\phi]$-*elliptic* free factor system. By *expanding immersion*, we mean every edge of the tree grows under iteration.

As a corollary of this summary, we have the following characterization: $[\phi]$ is induced by an expanding graph immersion if and only if the trivial system is the only $[\phi]$-fixed free factor system (Corollary 4.8). We will now contextualize these results and, especially, this corollary. In our previous work [24], we determined exactly when the mapping torus of an expanding graph immersion has a word-hyperbolic fundamental group.

Theorem ([24, Theorem 6.3]). *Let $\phi : F \to F$ be induced by an expanding graph immersion. $F*_\phi$ is word-hyperbolic if and only if it has no $BS(1,d)$ subgroups for $d \geq 2$.*

By Reynolds' result (Corollary 4.3), we knew this theorem applied to all irreducible nonsurjective endomorphisms but, at the time, there was no known algebraic characterization of the general class of endomorphisms induced by expanding graph immersions. It is clear that nontrivial $[\phi]$-periodic subgroups are an obstruction to ϕ being induced by an expanding graph immersion. Corollary 4.8 means nontrivial fixed free factor systems are the only (essential) obstruction. So the above theorem can be restated as:

Theorem. *Let $\phi : F \to F$ be an injective endomorphism with no nontrivial fixed free factor system. $F*_\phi$ is word-hyperbolic if and only if it has no $BS(1,d)$ subgroups for $d \geq 2$.*

The interesting thing about the restatement is that it is purely algebraic, i.e. there is no mention of topological maps. Since fixed free factor systems correspond to *free-by-cyclic* subgroups $F' \rtimes \mathbb{Z}$ of $F*_\phi$, the restatement also suggests how to extend our previous work to all injective nonsurjective endomorphisms.

In the next sections, we use expanding relative immersions to relativize our previous work in [24]. For instance, the sequence of lemmas/propositions in the next section is essentially identical to the sequence in [24, Section 3]. In Section 6, we extend our previous theorem to all injective endomorphisms. However, this will *not* constitute an alternate proof of Brinkmann's theorem [8]: $F \rtimes \mathbb{Z}$ is word-hyperbolic if and only if it has no $\mathbb{Z}^2$ subgroups. We use Brinkmann's theorem as a black box in our generalization.

Before diving into the details, we give an illustration of the considerations in the second part of this paper. To prove that $BS(1,d)$ subgroups are the only obstruction to word-hyperbolicity, the main tool we use is the Bestvina-Feighn combination theorem [3]: as long as the group in question satisfies the *annuli flaring condition*, it will be word-hyperbolic. See Section 7 for the exact statement. For non-examples, we now explain why/how $BS(1,d)$ groups fail the annuli flaring condition.

Consider two Baumslag-Solitar groups $BS(1,d)$ with $d = 1, 2$. Note that $BS(1,1) = \mathbb{Z}^2$. Fix the presentation $BS(1,d) = \langle a, t \mid t^{-1}at = a^d \rangle$ and notice that the group is a mapping torus of an endomorphism ψ of the cyclic group $\langle a \rangle$. So $BS(1,d)$ acts on a simplicial tree $\mathcal{T}$ with point and edge stabilizers conjugate to $\langle a \rangle$ and the quotient graph $BS(1,d)\backslash\mathcal{T}$ has one vertex and edge. The edge has a natural orientation inherited from the nontrivial loop corresponding to t and we can lift the orientation to an equivariant orientation of $\mathcal{T}$. Under this orientation, $\mathcal{T}$ is like a rooted tree: every vertex has exactly d incoming edges and one outgoing edge. Choose the vertex $\star$ on the axis of t whose stabilizer is $\langle a \rangle$ and define the family tree of $\star$-descendants $\mathcal{T}_{\langle a \rangle}$: union of directed edge-paths that terminate on $\star$.

A *class of annuli* of length n in $BS(1,d)$ is an ordered pair (g, p_g) where $g \in BS(1,d)$ is nontrivial and p_g is a geodesic of length n in $\mathcal{T}$ fixed by g. Since g is elliptic, then (up to conjugacy) we may assume $g \in \langle a \rangle$, $\star \in p_g$, and $p_g \subset \mathcal{T}_{\langle a \rangle}$. The orientation on $\mathcal{T}$ implies
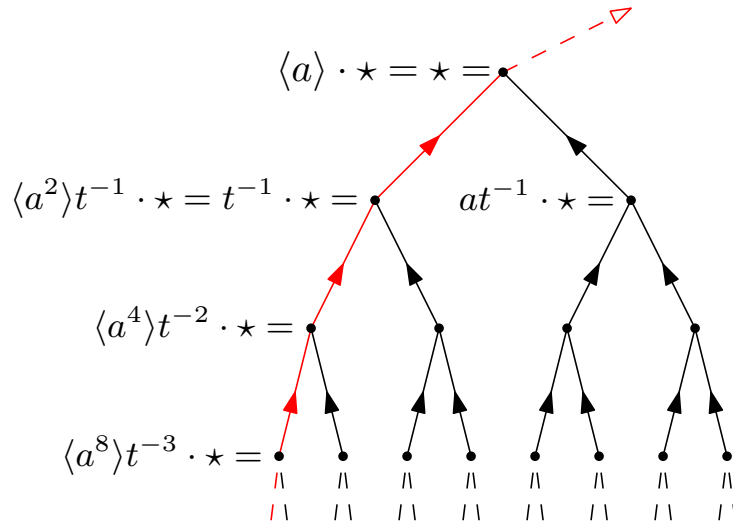
Figure 6: Portion of the trees $\mathcal{T}_{\langle a \rangle} \subset \mathcal{T}$ when $d = 2$. The line on the left is the axis of t.

there are two types of annuli: 1) unidirectional — those whose geodesics are directed paths terminating on $\star$; 2) bidirectional — those whose geodesics are a wedge of two directed paths terminating on $\star$. When $d = 1$, $\mathcal{T}$ is a line and all annuli are unidirectional. Loosely speaking, flaring means expanding exponentially towards some end. It should be intuitive that annuli will not flare because the associated endomorphism ψ is the identity map. When $d = 2$, $\mathcal{T}$ is a regular trivalent tree and annuli can be unidirectional or bidirectional. Since the associated endomorphism ψ is represented by a degree 2 map of the circle, the unidirectional annuli will flare. However, due to the branching in $\mathcal{T}$, it is possible to wedge two flaring unidirectional annuli along their expanding ends to get a bidirectional annulus that shrinks exponentially towards both its ends.

The objective is to rule out both of these behaviors in the setting of a mapping torus $F*_\phi$ that has no $BS(1, d)$ subgroups with $d \geq 1$ and thus conclude that the mapping torus satisfies the annuli flaring condition. Here is an outline of the second part of the paper:

1. In Section 5, we prove that if $F*_\phi$ has no $BS(1, d)$ subgroups with $d \geq 2$, then there is a uniform bound on one of the two directed pieces of any bidirectional annulus. Thus long bidirectional annuli are unidirectional up to bounded error. However, since annuli are not defined yet in this section, everything is done in terms of *pullbacks*, which are special classes of bidirectional annuli.

2. In Section 6, we define annuli and give the correspondence with pullbacks. We then use expanding relative immersions and Brinkmann's theorem to prove the section's main result: if $F*_\phi$ has no $\mathbb{Z}^2$ subgroups, then long unidirectional annuli flare. Thus, together with the previous section, $F*_\phi$ satisfies the annuli flaring condition if it has

36

no $BS(1, d)$ subgroups with $d \geq 1$.

3. Finally, in Section 7, we extend the result to HNN extensions of free groups over free factors $F*_A$ by showing that there is a canonical finite collection of mapping tori $F'*_{\phi'} \leq F*_A$ that carry all long annuli of $F*_A$. Intuitively speaking, the action of $F*_A$ on its Bass-Serre tree is *acylindrical* relative to $\{F'_i*_{\phi'_i}\}$. So $F*_A$ satisfies the annuli flaring condition if and only if the subgroups $\{F'_i*_{\phi'_i}\}$ satisfy the condition.

Geometry of ascending HNN extensions

5 Iterated pullbacks

In our previous work [24, Sections 3], topological pullbacks for a graph immersion $f : \Gamma \to \Gamma$ were used to give sufficient conditions for $\pi_1(f) = \phi : F \to F$ to have an invariant nonfixed *cyclic subgroup system*. The goal of this section is to drop the immersion hypothesis and give sufficient conditions that apply to all injective endomorphisms of F.

Suppose immersions $f_1 : \Gamma_1 \to \Gamma$ and $f_2 : \Gamma_2 \to \Gamma$ induce inclusions of free groups $H_1, H_2 \leq F$ respectively. Then components of the core *(topological) pullback* of (f_1, f_2), also known as *fibered products*, are in one-to-one correspondence with nontrivial intersections $H_1 \cap gH_2g^{-1}$ as g ranges over (H_1, H_2)-double coset representatives of $H_1\backslash F/H_2$. For a graph immersion $f : \Gamma \to \Gamma$ that induces an endomorphism $\phi : F \to F$, we get a one-to-one correspondence between components of the core pullback of (f^k, f^k) and nontrivial conjugacy classes $[\phi^k(F) \cap g\phi^k(F)g^{-1}]$ as $[[g]]$ ranges over $\phi^k(F)$-double cosets for all $k \geq 1$. We will not define pullbacks of topological maps since we will be working (semi-)algebraically in this section.

Given subgroups $H_1, H_2 \leq F$, we define the (algebraic) pullback of (H_1, H_2), denoted by $H_1 \wedge H_2$, to be the set of all nontrivial components $[H_1 \cap gH_2g^{-1}]$ as $[[g]]$ ranges over the (H_1, H_2)-double cosets in F. Note that pullbacks are technically *not* subgroup systems since they are sets of conjugacy classes of subgroups. When H_1 and H_2 are finitely generated, then their pullback is a finite set. For an injective endomorphism $\phi : F \to F$ and $k \geq 1$, define the iterated (algebraic) pullbacks of ϕ to be $\Lambda_k(\phi) = \phi^k(F) \wedge \phi^k(F)$. Furthermore, the iterated pullback depends only on the outer class $[\phi]$ and will be denoted later by $\Lambda_k[\phi]$. In this section, we will simply write Λ_k for brevity.

The (algebraic) mapping torus of an injective endomorphism ϕ is the ascending HNN extension of F with the presentation $F*_\phi = \langle F, t \mid t^{-1}xt = \phi(x), \forall x \in F \rangle$. We now give a third description of Λ_k that will actually be the working description. The mapping torus $F*_\phi$ is the fundamental group of a circle of groups with one vertex group F and edge group F. The edge monomorphisms for this circle of groups are the identity map $id_F : F \to F$ and endomorphism $\phi : F \to F$ and the corresponding Bass-Serre tree $\mathcal{T}$ has one orbit of edges and vertices. The tree also comes with a natural orientation where each vertex has

exactly one outgoing oriented edge and the <u>stable letter</u> $t \in F *_\phi$ acts on its directed axis with positive translation.

By construction, there is a vertex $\star$ of $\mathcal{T}$ whose stabilizer in $F *_\phi$ is exactly F. Let $\mathcal{T}_F$ be the full subtree of $\mathcal{T}$ rooted at $\star$. Colloquially, $\mathcal{T}_F$ is the family tree of $\star$-descendants. We will use the action of $F *_\phi$ on $\mathcal{T}$ to index the vertices of $\mathcal{T}_F$ by cosets in F. The root $\star$ is indexed gF, where $gF = F$ is the coset of F in F. The kth-generation vertices $gt^{-k} \cdot \star$ are indexed by the cosets $g\phi^k(F)$ in F. The intersection $g_1 \phi^k(F) g_1^{-1} \cap g_2 \phi^k(F) g_2^{-1}$ corresponds to the simultaneous stabilizer of the ordered pair of vertices indexed by $g_1 \phi^k(F)$ and $g_2 \phi^k(F)$; the intersection is conjugate in F to the stabilizer of the vertices indexed by $\phi^k(F)$ and $g\phi^k(F)$, where $g = g_1^{-1} g_2$. The intersection is also conjugate (in F) to the stabilizers of the vertices indexed by $\phi^k(F)$ and $hg\phi^k(F)$, where $h \in \phi^k(F)$. Thus, F-orbits of kth-generation ordered pairs of vertices are indexed by $\phi^k(F)$-double cosets in F and conjugacy classes of their nontrivial stabilizers are the components of Λ_k.

There is a function $\phi|_{\Lambda_k} : \Lambda_k \to \Lambda_{k+1}$ induced by ϕ, defined algebraically as

$$[\phi^k(F) \cap g\phi^k(F)g^{-1}] \quad \mapsto \quad [\phi^{k+1}(F) \cap \phi(g)\phi^{k+1}(F)\phi(g)^{-1}].$$

Graphically, the function maps (the conjugacy class of) the stabilizer of $(g_1 t^{-k}, g_2 t^{-k}) \cdot \star$ to (the conjugacy class of) the stabilizer of $t^{-1} \cdot (g_1 t^{-k}, g_2 t^{-k}) \cdot \star$.

Lemma 5.1. $\phi|_{\Lambda_k} : \Lambda_k \to \Lambda_{k+1}$ *as a function on the set of components is injective.*

Graphical proof. If $t^{-1} \cdot (g_1 t^{-k}, g_2 t^{-k}) \cdot \star$ and $t^{-1} \cdot (g_1' t^{-k}, g_2' t^{-k}) \cdot \star$ are in the same F-orbit, then they are in fact in the same $\phi(F)$-orbit (descendants of $t^{-1} \cdot \star$). So $(g_1 t^{-k}, g_2 t^{-k}) \cdot \star$ and $(g_1' t^{-k}, g_2' t^{-k}) \cdot \star$ are in the same F-orbit by the action of t. $\square$

One can give an algebraic proof of the lemma that replaces the action of t with the injectivity of ϕ. The lemma implies there is a chain of injections $\Lambda_0 \to \Lambda_1 \to \Lambda_2 \to \cdots$. Furthermore, the restriction to each component is an isomorphism since it is a conjugation by t in $F *_\phi$. We will be mainly interested in the set complement $\hat{\Lambda}_k = \Lambda_k \setminus \phi(\Lambda_{k-1})$. Equivalently, $\hat{\Lambda}_k = \{ [\phi^k(F) \cap g\phi^k(F)g^{-1}] \in \Lambda_k \; : \; g \notin \phi(F) \}$. Graphically, $\hat{\Lambda}_k$ is the set of components that stabilizes F-orbits of kth-generation pairs of vertices in $\mathcal{T}_F$ whose only common ancestor is $\star$. We might say iterated pullbacks *stabilize* if $\hat{\Lambda}_k$ is empty for some k. The image $\phi(F)$ is *malnormal* in F if and only if $\hat{\Lambda}_1 = \emptyset$; so iterated pullback stability is a sort of generalization of malnormality for $\phi^k(F)$ with respect to ϕ.

Lemma 5.2. *Suppose* $\phi : F \to F$ *is injective and* $k \geq 1$. *If* $\hat{\Lambda}_k$ *is empty, then so is* $\hat{\Lambda}_{k+1}$; *and if* $\hat{\Lambda}_k$ *has only cyclic components, then* $\hat{\Lambda}_{k+1}$ *is empty or has only cyclic components.*

Proof. There is an "inclusion" of components, $\hat{\Lambda}_{k+1} \preccurlyeq \hat{\Lambda}_k$, since a stabilizer of a $(k+1)^{\text{th}}$-generation pair of vertices is contained in a stabilizer of a kth-generation pair of vertices. So $\hat{\Lambda}_k = \emptyset$ implies $\hat{\Lambda}_{k+1} = \emptyset$. Suppose $\hat{\Lambda}_k$ has cyclic components, then $\hat{\Lambda}_{k+1}$ is empty or has cyclic components as the subgroups of a cyclic group are trivial or cyclic. $\square$

The <u>reduced rank</u> of a nontrivial finite rank free group H is $\mathrm{rr}(H) = \mathrm{rank}(H) - 1$ and the reduced rank of a pullback $H_1 \wedge H_2$, where H_1, H_2 are finitely generated free groups, is the sum of the reduced ranks of components in $H_1 \wedge H_2$. The latter is denoted by $\mathrm{rr}(H_1 \wedge H_2)$. In the literature, the reduced rank of the trivial group/pullback is defined to be 0 but we shall not need this fact.

Since the restriction $\phi|_{\Lambda_k}$ gives natural isomorphisms of the components, the chain of injections produces a nondecreasing sequence of positive integers

$$\mathrm{rank}(F) - 1 = \mathrm{rr}(F) \leq \mathrm{rr}(\Lambda_1) \leq \mathrm{rr}(\Lambda_2) \leq \cdots.$$

Observe that $\mathrm{rr}(\Lambda_i) = \mathrm{rr}(\Lambda_{i+1})$ if and only if $\hat{\Lambda}_{i+1}$ is empty or has cyclic components only. By Lemma 5.2, the sequence becomes constant once two consecutive entries are equal. Walter Neumann used topological pullbacks of Stallings graphs to bound the reduced ranks of algebraic pullbacks and improve Hanna Neumann's bound [26].

Theorem 5.3 ([27, Proposition 2.1]). *If $H_1, H_2 \leq F$ are nontrivial finitely generated subgroups, then $\mathrm{rr}(H_1 \wedge H_2) \leq 2 \, \mathrm{rr}(H_1) \, \mathrm{rr}(H_2)$.*

Remark. Although this bound is weaker than the *Strengthened Hanna Neumann Conjecture*, now Friedman-Mineyev's Theorem [12, 20], it is sufficient for our purposes. W. Neumann's proof is an elementary Euler characteristic computation for the topological pullback.

The uniform bound given by W. Neumann's theorem allows us to conclude that the sequence of reduced ranks given above is eventually constant, and hence, past a certain point we are adding only cyclic components or nothing at all to the iterated pullbacks.

Lemma 5.4. *If $\phi : F \to F$ is injective, then either $\hat{\Lambda}_k$ has cyclic components for all $k \geq 2 \, \mathrm{rr}(F)^2$ or $\hat{\Lambda}_k = \emptyset$ for some k.*

Proof. Theorem 5.3 gives us a uniform bound on the reduced ranks of the iterated pullbacks $\mathrm{rr}(\Lambda_k) \leq 2 \, \mathrm{rr}(\phi^k(F))^2 = 2 \, \mathrm{rr}(F)^2$ for all $k \geq 1$. By Lemma 5.2 and the uniform bound on the nondecreasing sequence of reduced ranks, $\mathrm{rr}(\Lambda_k)$ are all equal for $k \geq k_0 = 2 \, \mathrm{rr}(F)^2$. Therefore, $\hat{\Lambda}_{k_0}$ is empty or has only cyclic components. The lemma follows by inductively applying Lemma 5.2. $\square$

We say $\phi : F \to F$ has an <u>invariant cyclic subgroup system with index $d \geq 1$</u> if there is an integer $k \geq 1$, element $x \in F$, and nontrivial cyclic subgroup $\langle c \rangle \leq F$ such that $\phi^k(\langle c \rangle) \leq x \langle c \rangle x^{-1}$ and has index d. Note that this is actually a property of the outer endomorphism $[\phi]$. We can now give the main result of this section:

Proposition 5.5. *Let $\phi : F \to F$ be injective. If $\hat{\Lambda}_k$ is not empty for all $k \geq 1$, then $[\phi]$ has an invariant cyclic subgroup system with index $d \geq 2$.*

Proof. Let $k_0 = 2\operatorname{rr}(F)^2$ and $\phi : F \to F$ be an injective endomorphism such that $\hat{\Lambda}_k$ is not empty for all $k \geq 1$. By Lemma 5.4, $\hat{\Lambda}_k$ has cyclic components for $k \geq k_0$.

So far nothing in the section has used relative immersions but our main motivation for constructing them was this proposition. Note that $\hat{\Lambda}_1 \neq \emptyset$ automatically implies ϕ is not surjective. By Theorem 4.5, there is an expanding $\mathcal{A}^*$-relative immersion $f : T \to T$ for ϕ, where $\mathcal{A}^*$ is the $[\phi]$-elliptic free factor system and T is an $(F, \mathcal{A}^*)$-tree. Recall that the equality $\phi^{-1} \cdot \mathcal{A}^* = \mathcal{A}^*$ is part of our definition of $\mathcal{A}^*$-relative representatives. An element of F is *elliptic* if it fixes a point on T, i.e. its conjugacy class is carried by $\mathcal{A}^*$. By construction of the $[\phi]$-elliptic free factor system (Proposition 2.4), there is an $m \geq 1$ such that $[\phi^k(g)]$ is carried by the maximal $[\phi]$-fixed free factor system $\mathcal{A}$ for all $k \geq m$ and elliptic $g \in F$. As $\mathcal{A}$ is a $[\phi]$-fixed free factor system, the components of $\phi^k(\mathcal{A})$ are free factors of F for all k.

Suppose $[\phi^k(F) \cap g\phi^k(F)g^{-1}] \in \hat{\Lambda}_k$ for some $k \geq k_1 = \max(k_0, m+1)$. As $k \geq k_0$, this component is cyclic and we may assume it has a representative $\phi^k(F) \cap g\phi^k(F)g^{-1}$ generated by a nontrivial element $\phi^k(x) \in \phi^k(F)$. In particular, there exists $\phi^k(y) \in \phi^k(F)$ and $g \in F \setminus \phi(F)$ such that $\phi^k(x) = g\phi^k(y)g^{-1}$. We first show that x is *loxodromic*, i.e. not elliptic. Suppose $x \in F$ is elliptic, i.e. $\mathcal{A}^*$ carries $[x]$. Then some free factor $\phi^k(A) \in \phi^k(\mathcal{A})$ carries $[\phi^k(x)] = [\phi^k(y)]$ as $k \geq m+1$. So y must be elliptic too since $\phi^{-1} \cdot \mathcal{A}^* = \mathcal{A}^*$ and, as noted in the previous paragraph, $\phi^m(A)$ carries $\phi^m(x)$ and $\phi^m(y)$. Thus, up to conjugation in F, we may assume $\phi^m(x), \phi^m(y) \in \phi^m(A)$ and, up to conjugation in $\phi^{k-m}(F)$, we may assume $\phi^k(x), \phi^k(y) \in \phi^k(A)$. *Malnormality* of the free factor $\phi^k(A)$ implies $g \in \phi^{k-m}(F)$. But $k - m \geq 1$ leads to the contradiction $g \in \phi(F)$. Therefore, x is loxodromic. The integer $k \geq k_1$ and component $[\langle \phi^k(x) \rangle] \in \hat{\Lambda}_k$ were arbitrary, so all components of $\hat{\Lambda}_k$ are loxodromic for $k \geq k_1$. Recall from the proof of Lemma 5.2 that there is an infinite descending chain:

$$\hat{\Lambda}_1 \succcurlyeq \hat{\Lambda}_2 \succcurlyeq \hat{\Lambda}_3 \succcurlyeq \cdots$$

Since $\hat{\Lambda}_{k_1}$ has finitely many components and the components are all cyclic, there is a cyclic component in $\hat{\Lambda}_{k_1}$ which "carries" some component of $\hat{\Lambda}_k$ for all $k \geq k_1$. Suppose this component has a representative generated by $c = \phi^{k_1}(x) \in \phi^{k_1}(F)$. Then for all $k \geq k_1$, there is a cyclic component of $\hat{\Lambda}_k$ with a representative generated by $\phi^k(x_k) \in \phi^k(F)$ such that $\langle c \rangle$ carries $\langle \phi^k(x_k) \rangle$. If we let $\alpha \subset T$ be the axis for element c, then the previous sentence implies there are sequences of element $(x_k)_{k \geq k_1}$ and $(t_k)_{k \geq k_1}$ such that the (unoriented) axes of $\phi^k(x_k)$ are all translates $t_k \cdot \alpha$ of the (unoriented) axis α.

For any $k \geq k_1$, replace x_k with its inverse if necessary so that the action of $\phi^k(x_k)$ on its axis is coherent (respects orientation via t_k) with the action of c on α. By passing to a strictly increasing subsequence $(k_i)_{i \geq 1}$, we may assume there is an oriented edge e of T such that the axes α_{k_i} of $(x_{k_i})_{i \geq 1}$ all contain a translate of e. We now pass to the graph of groups $\Gamma = F \backslash T$ in order to avoid mentions of translates and orbits. The axis α will project to an immersed loop $\bar{\alpha}$ in Γ representing c and axes α_{k_i} project to immersed loops $\bar{\alpha}_{k_i}$ that represent x_{k_i} and whose $\bar{f}^{k_i}$-image is a power $\bar{\alpha}^{d_i}$ up to rotation/cyclic homotopy,

where $d_i \geq 1$ and $\bar{f} : \Gamma \to \Gamma$ is the expanding immersion induced by $f : T \to T$. The edge e projects to an edge $\bar{e}$ that is contained in all the loops $\bar{\alpha}_{k_i}$ for $i \geq 1$

The proof now mimics the proof of [24, Proposition 3.11]. Since f is an immersion, it maps axes in T onto axes and $\bar{f}$ maps immersed loops in Γ to immersed loops. So $\bar{f}^{k_i}(\bar{e})$ is a subpath of the immersed loop $\bar{f}^{k_i}(\bar{\alpha}_{k_i}) \simeq \bar{\alpha}^{d_i}$ for all i, and since f is expanding, $\bar{f}^{k_i}(\bar{e})$ contains arbitrarily long powers of $\bar{\alpha}$ as $i \to \infty$. Set N to be the number of subpaths of $\bar{\alpha}$ (up to rotation) that are also loops. Choose $n \gg 0$ such that $\bar{f}^{k_n}(\bar{e})$ contains the loop $\bar{\alpha}^{N+1}$ as a subpath. Then $\bar{f}^{k_{n+1}}(\bar{e})$ is a subpath of $\bar{\alpha}^{d_{n+1}}$ that contains the loop $\bar{f}^{k_{n+1}-k_n}(\bar{\alpha}^{N+1})$ as a subpath. In fact, for all positive integers $j \leq N+1$, the loop $\bar{f}^{k_{n+1}-k_n}(\bar{\alpha}^j)$ is a subpath of $\bar{\alpha}^{d_{n+1}}$. Thus, there is a sequence of loops $(\epsilon_j)_{j=1}^{N+1}$ that are subpaths of $\bar{\alpha}$ and strictly increasing positive integers $(s_j)_{j=1}^{N+1}$ such that $\bar{f}^{k_{n+1}-k_n}(\bar{\alpha}^j) \cdot \epsilon_j$ is $\bar{\alpha}^{s_j}$ up to rotation. By definition of N and pigeonhole principle, some $\epsilon_t = \epsilon_{t'} = \epsilon$ for some $t < t'$ and $\bar{f}^{k_{n+1}-k_n}(\bar{\alpha}^{t'-t})$ is $\bar{\alpha}^{s_{t'}-s_t}$ up to rotation. Lifting back to the $(F, \mathcal{A})$-tree T, we find that $f^{k_{n+1}-k_n}$ maps the axis α to a translate of itself. So $\phi^{k_{n+1}-k_n}(c)$ is conjugate to a nontrivial power c^d; finally, $d \geq 2$ since f is expanding. $\qquad\square$

Remark. A careful examination of the proof reveals that it can be made "more effective" with the pigeonhole principle. For any expanding relative immersion f, we can construct a specific number $k = k(f)$ for which $\hat{\Lambda}_k \neq \emptyset$ implies ϕ has an invariant cyclic subgroup system with index $d \geq 2$. Thus, one would not have to check infinitely many iterated pullbacks of ϕ to know that an invariant cyclic subgroup system with index $d \geq 2$ exists. Of course, for this to be useful, we need to know whether finding the $[\phi]$-elliptic free factor system and expanding relative immersion for ϕ can be made effective. Finally, we remark that the converse of Proposition 5.5 holds but we omit the proof as it is not needed for the rest of the paper.

In the next section, we will actually be using the contrapositive of the proposition: if ϕ is injective and has no invariant cyclic subgroup system with index $d \geq 2$, then iterated pullbacks of ϕ stabilize. In this case, we can control the types of annuli in the mapping torus $F*_\phi$, which allows us to prove the main theorem: $F*_\phi$ is word-hyperbolic if ϕ also has no fixed cyclic subgroup system, i.e. no invariant cyclic subgroup system with index 1.

6   Hyperbolic endomorphisms

We are finally ready to put all the major pieces together. The first piece involves understanding the relationship between *annuli* in the mapping torus $F*_\phi$ and iterated pullbacks of ϕ. The second piece involves building on Brinkmann's theorem to show that *atoroidal* injective endomorphisms are *hyperbolic*. In our previous work [24], we used these two pieces to give sufficient conditions for the mapping torus to be word-hyperbolic.

Theorem 6.1 ([24, Theorem 6.4])**.** *If $f : \Gamma \to \Gamma$ is a based-hyperbolic graph map and all strictly bidirectional annuli in its mapping torus M_f are shorter than some integer, then $\pi_1(M_f)$ is word-hyperbolic.*

We will define the new terms in the theorem as we go. For the sake of completeness, a sketch of the proof can be found in the next section (Theorem 7.7).

Suppose $\phi : F \to F$ is an injective endomorphism and $f : \Gamma \to \Gamma$ is its topological representative. Recall that the (topological) mapping torus of f is the quotient space $M_f = (\Gamma \times [0,1]) / \sim_f$ with the identification $(x,1) \sim_f (f(x),0)$ for all $x \in \Gamma$ and the algebraic mapping torus $F*_\phi$ is isomorphic to the fundamental group $\pi_1(M_f)$. The edge-space of M_f will be the cross-section in M_f represented by $\Gamma \times \{\frac{1}{2}\}$.

Strictly bidirectional annuli of the mapping torus M_f with length $2L$ can be thought of as the iterated pullbacks $\hat{\Lambda}_L$ of ϕ but it does take some work to give the correspondence. Fix a basepoint $\nu \in S^1$. For integers $L_1 < L_2$, an (topological) annulus in M_f of length $L = L_2 - L_1$ is a homotopy of loops $h : S^1 \times [L_1, L_2] \to M_f$ satisfying the following conditions:

1. it is transverse to the edge-space of M_f;

2. the h-preimage of the edge-space is $S^1 \times ([-L_1, L_2] \cap \mathbb{Z})$;

3. for integers $i \in [L_1, L_2]$, the rings of the annulus $h_i = h(\cdot, i) : S^1 \to M_f$ are locally injective every where except possibly at the basepoint ν;

4. and for the trace of the basepoint $h^\nu = h(\nu, \cdot) : [L_1, L_2] \to M_f$, no subpath between consecutive integer coordinates $[i, i+1]$ is homotopic rel. endpoints into the edge-space.

This is the definition used in [24]. In light of our last description of $\hat{\Lambda}_k$ in the previous section, we give an alternative definition. Let $\mathcal{T}$ be the Bass-Serre tree for $F*_\phi$ and $\star$ be the point whose stabilizer is F. An (algebraic) annulus $([\alpha], p_\alpha)$ in $F*_\phi$ of length $L \geq 2$ is a choice a nontrivial conjugacy class $\overline{[\alpha]}$ in $F*_\phi$ and an orbit of a geodesic path in $p_\alpha \subset \mathcal{T}$ of length L fixed by α. Since elements of $[\alpha]$ act on $\mathcal{T}$ with fixed points, we can always choose a representative $\alpha \in F$. Technically, we have defined a *conjugacy class* of algebraic annuli but the distinction will not be relevant for us.

Lemma 6.2. *Let $\phi : F \to F$ be injective and $f : \Gamma \to \Gamma$ be a topological representative for $[\phi]$. There is a one-to-one correspondence between (homotopy classes of) annuli in M_f of length $L \geq 1$ and (conjugacy classes of) annuli in $F*_\phi$ of length $L + 1 \geq 2$.*

Proof. Given a topological annulus h in M_f of length $L \geq 1$, then the generator of the image $\pi_1(h) : \mathbb{Z} \to \pi_1(M_f)$ determines a conjugacy class $[\alpha]$ in $\pi_1(M_f) \cong F*_\phi$. Condition (3) ensures $\pi_1(h)$ is injective and α is nontrivial. Let $\tilde{h} : \mathbb{R} \times [L_1, L_2] \to \tilde{M}_f$ be the lift of the annulus to the universal cover of M_f. Collapsing the $\tilde{\Gamma}$-direction of $\tilde{M}_f$ produces the Bass-Serre tree $\mathcal{T}$ and Condition (2) ensures the induced map $\bar{h} : \mathbb{R} \times ([L_1, L_2] \cap \mathbb{Z}) \to \mathcal{T}$ is constant on the first factor and its image is a collection of edge-midpoint; each ring h_i determines a conjugacy class in the stabilizer of the corresponding edge-midpoint. By

Conditions (1) and (4), the midpoints extend to a geodesic edge-path p_α in $\mathcal{T}$ of length $L + 1$ fixed by α.

The other direction works in a similar fashion. For any conjugacy class $[\alpha]$ in $F*_\phi$ and two consecutive edge-midpoints in $\mathcal{T}_F$ fixed by $\alpha \in F$, we can construct an annulus of length 1 as follows. Fix a basepoint in the edge-space and assume $\star$ is the vertex between the midpoints. If the midpoints are increasing/decreasing, then $\alpha = \phi(x) \in \phi(F)$ without loss of generality. Let σ, ρ be based loop in the edge-space representing $x, \phi(x) \in F$ and τ a based loop in M_f representing $t \in F*_\phi$. Then the based path $\sigma \cdot \tau \cdot \bar\rho \cdot \bar\tau$ is null-homotopic and can be extended to an annulus with ends σ, ρ and trace τ. If the midpoints are at the same height, then $\alpha = \phi(x) \in \phi(F)$ and $\alpha = g\phi(y)g^{-1} \in g\phi(F)g^{-1}$ for some $g \notin \phi(F)$. Let σ, ρ, γ be based loops in the edge-space representing $x, y, g \in F$ respectively and τ a based loop in M_f representing $t \in F*_\phi$. So the based path $\sigma \cdot \tau\gamma\bar\tau \cdot \bar\rho \cdot \tau\bar\gamma\bar\tau$ is null-homotopic and can be extended to an homotopy between σ, ρ satisfying Conditions (1)-(3) and having trace $\tau\gamma\bar\tau$. This trace satisfies Condition (4) because $tgt^{-1} \notin F$ and thus the homotopy is a topological annulus with ends σ, ρ. Given a geodesic path in $\mathcal{T}$ of length $L \geq 2$ fixed by α, we can replace the path with a translate in $\mathcal{T}_F$ without affecting the class $[\alpha]$ and then construct a topological annulus in M_f of length $L - 1$ by concatenating length 1 annuli from the preceding discussion. This concludes the correspondence between topological and algebraic annuli. $\qquad \square$

The natural orientation on $\mathcal{T}$ gives a dichotomy for annuli $([\alpha], p_\alpha)$ in $F*_\phi$:

1. all edges of p_α have the same orientation— we say α is <u>unidirectional</u>.

2. p_α switches from increasing to decreasing exactly once — we say α is <u>bidirectional</u>.

Each vertex of $\mathcal{T}$ has exactly one outgoing edge and hence the geodesic p_α cannot switch from decreasing to increasing because $[F : F] = 1$. For similar reasons, bidirectional annuli do not exist if and only if $\phi(F)$ is malnormal in F. The next proposition generalizes this equivalence of bidirectional annuli (or lack thereof) and malnormality.

An annulus $([\alpha], p_\alpha)$ in $F*_\phi$ is <u>strictly bidirectional</u> if the switch from increasing to decreasing occurs at the midpoint of p_α.

Lemma 6.3. *Let $\phi : F \to F$ be injective. For any integer $L \geq 1$, the mapping torus $F*_\phi$ has a strictly bidirectional annulus of length $2L$ if and only if $\hat\Lambda_L[\phi]$ is not empty.*

Proof. If there is a strictly bidirectional annulus $([\alpha], p_\alpha)$ in $F*_\phi$ of length $2L$, then we may assume the midpoint of p_α is $\star$ after replacing p_α with a translate. Then the stabilizer of p_α contains α and so it is nontrivial. Since the stabilizer of p_α is the stabilizer of its (ordered) endpoints, the conjugacy class of the stabilizer is a component of $\hat\Lambda_L[\phi]$.

Conversely, if $\hat\Lambda_L[\phi]$ is not empty, then some path in $\mathcal{T}_F$ of length $2L$ with midpoint at $\star$ has a nontrivial stabilizer. Choose a nontrivial element α in this stabilizer and $([\alpha], p_\alpha)$ is a strictly bidirectional annulus in $F*_\phi$ of length $2L$. $\qquad \square$

Let $\phi : F \to F$ be injective and $f : \Gamma \to \Gamma$ be a topological representative for ϕ. If $[\phi]$ has no invariant cyclic subgroup system with index $d \geq 2$, then there is an integer $L \geq 1$ for which $\hat{\Lambda}_L[\phi]$ is empty (Proposition 5.5) and all strictly bidirectional annuli in M_f are shorter than $2L$ (Lemma 6.3). This sets up the second hypothesis of Theorem 6.1.

As for the first hypothesis, we begin by defining *(based-) hyperbolicity.* For a real number $\lambda > 1$ and integer $n \geq 1$, we say a graph map $f : \Gamma \to \Gamma$ is (based-) (λ, n)-hyperbolic if all (based) loops $\sigma : S^1 \to \Gamma$ (with the basepoint mapped to a vertex) satisfy the inequality

$$\lambda |f^n(\sigma)| \leq \max(|f^{2n}(\sigma)|, |\sigma|)$$

where $|\cdot|$ is the combinatorial length after tightening; whether tightening respects a basepoint (based homotopy) or not (free homotopy) will be apparent from the context.

When a graph map is (based-) (λ', n)-hyperbolic for some $\lambda' > 1, n \geq 1$, then it is (based-) (λ^k, nk)-hyperbolic for all $k \geq 1$ and $\lambda \in (1, \lambda']$. So the constants can be omitted and when we do need them, we can assume $\lambda > 1$ is any preferred integer. As defined, hyperbolicity is a property of the homotopy class $[f]$; meanwhile, based-hyperbolicity is a property of the map f.

A graph map $f : \Gamma \to \Gamma$ is (topologically) atoroidal if it is π_1-injective and there is no nontrivial loop σ in Γ and integer $k \geq 1$ such that $f^k(\sigma) \simeq \sigma$. This again is a property of the homotopy class $[f]$. Bestvina-Feighn-Handel showed, as a step in [4, Theorem 5.1], that hyperbolic atoroidal homotopy equivalences are based-hyperbolic. Their argument is reproduced here, modified to drop the π_1-surjectivity assumption. This allows us to consider the growth rate of loops without basepoints for the rest of the section

Lemma 6.4. *If the graph map $f : \Gamma \to \Gamma$ is atoroidal and $(3, n)$-hyperbolic, then it is based-$(2, n')$-hyperbolic.*

To avoid context-ambiguity in the proof, we use $\|\cdot\|$ for lengths of free homotopy classes of loops and $|\cdot|$ for lengths of loops rel. basepoints. However, the distinction is not needed after the proof as all loops afterwards will be considered up to free homotopy.

Proof. Suppose $f : \Gamma \to \Gamma$ is atoroidal and $(3, n)$-hyperbolic for some integer $n \geq 1$. Set M to be the maximum length of $f^k(s)$ rel. basepoint over all embedded based loops s in Γ for $k \in \{0, n, 2n\}$.

Suppose $|f^n(\sigma)| \geq 4M$ for some immersed based loop σ and pick an embedded based loop s with same basepoint as σ so that the concatenation $s \cdot \sigma$ is an immersed loop, i.e. $\|s \cdot \sigma\| = |s| + |\sigma|$. As the graph map f is $(3, n)$-hyperbolic, we get

$$3 \|f^n(s \cdot \sigma)\| \leq \max(\|f^{2n}(s \cdot \sigma)\|, \|s \cdot \sigma\|).$$

For a concatenation of based loops $\rho_1 \cdot \rho_2$, we get $| |\rho_1| - |\rho_2| | \leq \|\rho_1 \cdot \rho_2\| \leq |\rho_1| + |\rho_2|$.

Case 1. If $\|f^{2n}(s \cdot \sigma)\| \geq 3 \|f^n(s \cdot \sigma)\|$, then

$$\begin{aligned}
|f^{2n}(\sigma)| \geq \|f^{2n}(s \cdot \sigma)\| - |f^{2n}(s)| &\geq 3 \|f^n(s \cdot \sigma)\| - M \\
&\geq 3 |f^n(\sigma)| - 3 |f^n(s)| - M \\
&\geq 3 |f^n(\sigma)| - 4M \\
&\geq 2 |f^n(\sigma)|.
\end{aligned}$$

Case 2. If $\|s \cdot \sigma\| \geq 3 \|f^n(s \cdot \sigma)\|$, then

$$|\sigma| = \|s \cdot \sigma\| - |s| \geq 3 \|f^n(s \cdot \sigma)\| - M \geq 3 |f^n(\sigma)| - 4M \geq 2 |f^n(\sigma)|$$

Combining both cases: $2 |f^n(\sigma)| \leq \max(|f^{2n}(\sigma)|, |\sigma|)$. If $|f^{nk}(\sigma)| \geq 4M$ for an immersed based loop σ in Γ and $k \geq 1$, then by induction

$$2^k |f^{nk}(\sigma)| \leq \max(|f^{2nk}(\sigma)|, |\sigma|).$$

For any bound B, there are only finitely many immersed based loops σ' in Γ with $|\sigma'| \leq B$. Since f is atoroidal, there is an integer $k \gg 0$ such that $|f^{nk}(\sigma')| \geq 8M$ for every based loop σ' with $|\sigma'| \leq 4M$ and we conclude that f is based-$(2, nk)$-hyperbolic . $\qquad \square$

When p is a subpath of an immersed loop σ and $n \geq 1$, then $[f^n(p)]_\sigma$ is the subpath of $[f^n(p)]$ that survives in $[f^n(\sigma)]$ and $|f^n(p)|_\sigma$ is the length of $[f^n(p)]_\sigma$. Bounded cancellation implies $|f^n(p)| \leq |f^n(p)|_\sigma + 2C(f^n)$. The next lemma is based on Brinkmann's Lemma 4.2 in [8] with a few changes made to account for graph maps that are not π_1-surjective.

Lemma 6.5. *Let $f : \Gamma \to \Gamma$ be a π_1-injective graph map with an invariant union of roses $R_* \subset \Gamma$ such that the restriction $f|_{R_*} : R_* \to R_*$ is $(4, n)$-hyperbolic. For some constant L_c, if p in R_* is a subpath (edge-path) of an immersed loop σ in Γ and $|f^n(p)|_\sigma \geq L_c$, then*

$$3 |f^n(p)|_\sigma \leq \max \left(|f^{2n}(p)|_\sigma, |p| \right).$$

The number L_c is the <u>critical length</u> of the triple (f, Γ, R_*).

Proof. Let $f : \Gamma \to \Gamma$ be a graph map and $R_* \subset \Gamma$ be an f-invariant union of roses such that the restriction $f|_{R_*}$ is $(4, n)$-hyperbolic. Set M to be the maximum length of $f^k(s)$ rel. basepoint over all petals s in R_* and $k \in \{n, 2n\}$. Choose $L_c = 2C(f^{2n}) + 5M$ where $C(f^{2n})$ is the cancellation constant for f^{2n}. Recall the triangle inequality: $||p_1| - |p_2|| \leq |p_1 \cdot p_2| \leq |p_1| + |p_2|$ for any path decomposition of a loop $p_1 \cdot p_2$. A remark on the context: paths $[p_i]$ are tightened rel. endpoints but the loop $[p_1 \cdot p_2]$ is tightened by free homotopy.

Given a subpath $p \subset R_*$ of some immersed loop σ in Γ, pick a petal s in R_* such that $s \cdot p$ is an immersed loop in R_*. As the restriction to R_* is $(4, n)$-hyperbolic, we get

$$4 |f^n(s \cdot p)| \leq \max(|f^{2n}(s \cdot p)|, |s \cdot p|).$$

If $4\,|f^n(s \cdot p)| \leq |f^{2n}(s \cdot p)|$, then

$$\begin{aligned}
4\,|f^n(p)|_\sigma \leq 4\,|f^n(p)| &\leq 4\,|f^n(s \cdot p)| + 4\,|f^n(s)| \\
&\leq |f^{2n}(s \cdot p)| + 4M \\
&\leq |f^{2n}(p)| + 5M \\
&\leq |f^{2n}(p)|_\sigma + 2C(f^{2n}) + 5M \quad \text{(bounded cancellation)}
\end{aligned}$$

Similarly, if $4\,|f^n(s \cdot p)| \leq |s \cdot p|$, then

$$\begin{aligned}
4\,|f^n(p)|_\sigma &\leq |p| + 1 + 4M \\
&\leq |p| + 5M + 2C(f^{2n}) \quad \text{(since } M \geq 1)
\end{aligned}$$

Since $L_c = 2C(f^{2n}) + 5M$, we have the desired implication:

$$|f^n(p)|_\sigma \geq L_c \implies 3\,|f^n(p)|_\sigma \leq \max\big(\,|f^{2n}(p)|_\sigma, |p|\,\big). \qquad \square$$

An (outer class of an) injective endomorphism $\phi : F \to F$ is (algebraically) atoroidal if it has no invariant cyclic subgroup system with index $d = 1$. If $f : \Gamma \to \Gamma$ is a topological representative for ϕ, then $[f]$ is (topologically) atoroidal if and only if $[\phi]$ is (algebraically) atoroidal. The next proposition is an extension of Brinkmann's result [8, Proposition 7.1] and is the main technical result of the section.

Proposition 6.6. *If $\phi : F \to F$ is injective and atoroidal, then $[\phi]$ has a $(2, n)$-hyperbolic topological representative for some integer $n \geq 1$.*

Proof. Suppose $\phi : F \to F$ is an injective and atoroidal endomorphism. If ϕ is surjective, then the proposition is precisely Brinkmann's result. So we assume ϕ is not surjective. By Theorem 4.5, there is an expanding $\mathcal{A}^*$-relative immersion $g : T \to T$ for ϕ, where $\mathcal{A}^*$ is the $[\phi]$-elliptic free factor system. Fix some $\mathcal{A}^*$-marked roses $(R_{\mathcal{A}^*}, \alpha_{\mathcal{A}^*})$ and set (Γ, α) to be the $(R_{\mathcal{A}^*}, \alpha_{\mathcal{A}^*})$-vertex blow-up of the graph of groups $F \backslash T$, i.e. (Γ, α) is a marked graph with subgraphs $R_{\mathcal{A}^*}$ corresponding to the free factor system $\alpha(\mathcal{A}^*)$. The roses $R_{\mathcal{A}^*}$ form the *lower stratum* of Γ and the remaining edges the *top stratum*.

We outline the proof which follows closely Brinkmann's strategy. Patch together a homotopy equivalence of the lower stratum with the expanding relative immersion to get some topological representative f of $[\phi]$. By Brinkmann's theorem, the restriction of f to the lower stratum is hyperbolic. The expanding relative immersion on the top stratum means loops that are *mostly* top stratum will have uniform exponential growth under forward iteration. Lemma 6.5 implies loops that are *mostly* lower stratum will have uniform exponential growth under forward and/or backward iteration. The heart of the proof lies in quantifying what being mostly top/lower stratum means and showing that all loops are one and/or the other. Of course, there are a few minor technicalities that need addressing;

for instance, the restriction to the lower stratum is almost but not exactly a homotopy equivalence. This concludes the outline.

Recall that the maximal $[\phi]$-fixed free factor system $\mathcal{A}$ is a subset of $\mathcal{A}^*$ and there is an integer $k_0 \geq 0$ such that $\phi^{k_0}(\mathcal{A}^*)$ is carried by $\mathcal{A}$ (Proposition 2.4). So we may find a topological representative $f_{\mathcal{A}^*} : R_{\mathcal{A}^*} \to R_{\mathcal{A}^*}$ for $[\phi|_{\mathcal{A}^*}]$ whose restriction to the periodic components $R_{\mathcal{A}}$, denoted by $f_{\mathcal{A}}$, is a homotopy equivalence. As ϕ is atoroidal, the restriction $f_{\mathcal{A}}$ is $(4, n_0)$-hyperbolic for some $n_0 \geq 1$ (Brinkmann's theorem).

If σ is an immersed loop in $R_{\mathcal{A}^*}$, then $[f^k(\sigma)]$ is a loop in the periodic roses $R_{\mathcal{A}}$ for all $k \geq k_0$. Since the restriction $f_{\mathcal{A}}$ is $(4, n_0)$-hyperbolic and $f_{\mathcal{A}^*}^{n_0 k}(\sigma)$ is a loop in $R_{\mathcal{A}}$ for any loop σ in $R_{\mathcal{A}^*}$ and $k \geq k_0$, we get the inequality

$$4 \cdot |f_{\mathcal{A}}^{n_0}(f_{\mathcal{A}^*}^{n_0 k}(\sigma))| \leq \max(|f_{\mathcal{A}}^{2n_0}(f_{\mathcal{A}^*}^{n_0 k}(\sigma))|, |f_{\mathcal{A}^*}^{n_0 k}(\sigma)|) \text{ for all loops } \sigma \text{ in } R_{\mathcal{A}^*} \text{ and } k \geq k_0.$$

Choose an integer $k_1 \geq 1$ so that $4^{k_1} \geq 4K^{n_0 k_0}$, where $K = K(f_{\mathcal{A}^*})$ is the Lipschitz constant for $f_{\mathcal{A}^*}$. Suppose σ is a loop in $R_{\mathcal{A}^*}$.

If $4 \cdot |f_{\mathcal{A}}^{n_0}(f_{\mathcal{A}^*}^{n_0(k_0+k_1-1)}(\sigma))| \leq |f_{\mathcal{A}}^{2n_0}(f_{\mathcal{A}^*}^{n_0(k_0+k_1-1)}(\sigma))|$, then by induction

$$4^{k_0+k_1} \cdot |f_{\mathcal{A}^*}^{n_0(k_0+k_1)}(\sigma)| \leq |f_{\mathcal{A}^*}^{2n_0(k_0+k_1)}(\sigma)|.$$

If $4 \cdot |f_{\mathcal{A}}^{n_0}(f_{\mathcal{A}^*}^{n_0(k_0+k_1-1)}(\sigma))| \leq |f_{\mathcal{A}^*}^{n_0(k_0+k_1-1)}(\sigma)|$, then by induction and Lipschitz property

$$4^{k_1} \cdot |f_{\mathcal{A}^*}^{n_0(k_0+k_1)}(\sigma)| \leq |f_{\mathcal{A}^*}^{n_0 k_0}(\sigma)| \leq K^{n_0 k_0} \cdot |\sigma|$$

and $4 \cdot |f_{\mathcal{A}^*}^{n_0(k_0+k_1)}(\sigma)| \leq |\sigma|$ by choice of k_1.

Therefore, the lower stratum map $f_{\mathcal{A}^*}$ is $(4, n_1)$-hyperbolic with $n_1 = n_0(k_0 + k_1)$.

Let $f : \Gamma \to \Gamma$ be a topological representative for $[\phi]$ that extends $f_{\mathcal{A}^*}$ and induces the expanding $\mathcal{A}^*$-relative immersion $g : T \to T$ upon collapsing the lower stratum in the universal cover $\tilde{\Gamma}$. For an arbitrary immersed loop σ in Γ that intersects the top stratum, define σ_{top} (σ_{low} resp.) to be the collection of maximal subpaths of σ in the top (lower resp.) stratum. For all $n \geq 1$, define $[f^n(\sigma_{top})]_\sigma$ ($[f^n(\sigma_{low})]_\sigma$ resp.) to be the collection of paths $[f^n(p)]_\sigma$ where p is some path in σ_{top} (σ_{low} resp.). That f induces an immersion g upon collapsing the lower stratum implies that the top stratum is *persistent*: $f(\sigma)_{top}$ survives in $[f(\sigma)]$.

As the relative immersion $g : T \to T$ is expanding, there is an integer $k_2 \geq 1$, such that $g^{k_2}(e)$ has length ≥ 2 for all edges e in T; and as f induces g, for any immersed loop σ in Γ and path p in σ_{top}, we get $2|p|_\sigma \leq |f^{k_2}(p)|_\sigma$. We may replace n_1 and k_2 with a common multiple and assume $n_1 = k_2$. A similar inequality holds in the lower stratum: by the $(4, n_1)$-hyperbolicity of $f|_{R_{\mathcal{A}^*}}$ and Lemma 6.5, there is a critical length $L_c = L_c(f, \Gamma, R_{\mathcal{A}^*})$ such that for any immersed loop σ in Γ that intersects the top stratum and path p in σ_{low},

$$|f^{n_1}(p)|_\sigma \geq L_c \implies 3|f^{n_1}(p)|_\sigma \leq \max(|f^{2n_1}(p)|_\sigma, |p|).$$

Set M to be the maximal length amongst all paths in $f^{n_1}(e)_{low}$ for all top stratum edges e of Γ. For any integer $k \geq 1$, we distinguish two cases:

Case 1. If $|f^{n_1 k}(\sigma)_{low}| \leq (L_c + 6M)\,|f^{n_1 k}(\sigma)_{top}|$, then

$$|f^{n_1 k}(\sigma)| = |f^{n_1 k}(\sigma)_{low}| + |f^{n_1 k}(\sigma)_{top}| \leq (L_c + 6M + 1)\,|f^{n_1 k}(\sigma)_{top}| \text{ and}$$
$$2^k\,|f^{n_1 k}(\sigma)_{top}| \leq |f^{n_1 k}(f^{n_1 k}(\sigma)_{top})|_\sigma \leq |f^{2n_1 k}(\sigma)|.$$

Additionally, if $2^k \geq 2(L_c + 6M + 1)$, then $2\,|f^{n_1 k}(\sigma)| \leq |f^{2n_1 k}(\sigma)|$.

Case 2. Suppose $|f^{n_1 k}(\sigma)_{low}| \geq (L_c + 6M)\,|f^{n_1 k}(\sigma)_{top}|$. Set m to be the number of paths in $f^{n_1 k}(\sigma)_{low}$. Then m is also the number of paths in $f^{n_1 k}(\sigma)_{top}$ and $m \leq |f^{n_1 k}(\sigma)_{top}|$. By the pigeonhole principle, some path ρ in $f^{n_1 k}(\sigma)_{low}$ satisfies $|\rho| \geq L_c + 6M$. As $|\rho| \geq 6M$, we have $3(|\rho| - 2M) \geq 2|\rho|$. Set $\sigma' = f^{n_1(k-1)}(\sigma)$. By definition of M and persistence of $f^{n_1}(\sigma')_{top}$, there must be a path p' in σ'_{low} such that $[f^{n_1}(p')]_{\sigma'}$ is a subpath of ρ, $|f^{n_1}(p')|_{\sigma'} \geq |\rho| - 2M \geq L_c$, and $3\,|f^{n_1}(p')|_{\sigma'} \leq \max(|f^{2n_1}(p')|_{\sigma'}, |p'|)$.

If $|f^{2n_1}(p')|_{\sigma'} \geq 3|f^{n_1}(p')|_{\sigma'}$, then $|f^{2n_1}(p')|_{\sigma'} \geq 2|\rho|$ and $|f^{n_1(k+1)}(p')|_{\sigma'} \geq 3^{k-1} \cdot 2|\rho|$.

If $|p'| \geq 3|f^{n_1}(p')|_{\sigma'}$, then $|p'| \geq 2|\rho|$. By inducting on the same argument used at the start of the case, there must be a path p in σ_{low} such that $|f^{n_1 k}(p)|_\sigma$ is a subpath of ρ and $|p| \geq 2^k\,|\rho|$. In either case, we get $2^k|\rho| \leq \max(|f^{n_1(k+1)}(p')|_{\sigma'}, |p|)$. Define $f^{n_1 k}(\sigma)_{crit}$ to be the set of paths ρ in $f^{n_1 k}(\sigma)_{low}$ with $|\rho| \geq L_c + 6M$. Altogether, we have shown:

$$2^k|f^{n_1 k}(\sigma)_{crit}| \leq \max(|f^{2n_1 k}(\sigma)|, |\sigma|).$$

The following computation is lifted from Brinkmann [8, Proof of Proposition 7.1]. Set $A = |f^{n_1 k}(\sigma)_{crit}|$ to be the total length of paths in $f^{n_1 k}(\sigma)_{crit}$, $B = |f^{n_1 k}(\sigma)_{low}| - A$ to be the total length of the remaining paths in $f^{n_1 k}(\sigma)_{low}$, and $C = |f^{n_1 k}(\sigma)_{top}|$. We now find a positive lower bound of $\frac{A}{A+B+C}$ that is independent of σ and k. We assumed $A + B \geq (L_c + 6M)\,C$ and so $\frac{A}{A+B+C} \geq \frac{A(L_c+6M)}{(A+B)(L_c+6M+1)}$ and we can focus on the factor $\frac{A}{A+B} = 1 - \frac{B}{A+B}$. Recall $m \leq C$, so $A + B \geq (L_c + 6M)\,m$. Since each path p in $f^{n_1 k}(\sigma)_{low}$ but not in $f^{n_1 k}(\sigma)_{crit}$ satisfies $|p| < L_c + 6M$ and there are at most m of them, $B \leq m\,(L_c + 6M - 1)$. Combining the last two inequalities gives the bound

$$1 - \frac{B}{A+B} \geq 1 - \frac{m\,(L_c + 6M - 1)}{m\,(L_c + 6M)} \geq \frac{1}{L_c + 6M}.$$

Altogether, $\frac{A}{A+B+C} \geq \frac{1}{L_c+6M+1}$.

Additionally, if $2^k \geq 2(L_c + 6M + 1)$, then $2|f^{n_1 k}(\sigma)| \leq \max(|f^{2n_1 k}(\sigma)|, |\sigma|)$.

Choose $k \geq 1$ so that $2^k \geq 2(L_c + 6M + 1)$; the two exhaustive cases above imply f is $(2, n_1 k)$-hyperbolic. $\qquad\square$

All the heavy lifting is done and we can prove our main theorem

Theorem 6.7. *Let $\phi : F \to F$ be an injective endomorphism. Then the following statements are equivalent:*

48

1. $F*_\phi$ is word-hyperbolic;

2. $F*_\phi$ contains no $BS(1,d)$ subgroups with $d \geq 1$;

3. $[\phi]$ has no invariant cyclic subgroup system with index $d \geq 1$;

4. all/some topological representatives of $[\phi]$ are based-hyperbolic and all strictly bidirectional annuli in their mapping tori are shorter than some integer.

Proof. This proof is mostly a matter of bookkeeping.

$(1) \Longrightarrow (2)$: $BS(1,d)$ subgroups are well-known obstructions to word-hyperbolicity.

$(2) \Longrightarrow (3)$: If $[\phi]$ has an invariant cyclic subgroup system with index $d \geq 1$, then there is a subgroup of $F*_\phi$ isomorphic to a quotient of $BS(1,d)$; use normal forms to show the subgroup is in fact isomorphic to $BS(1,d)$ (See [18, Lemma 2.3] for details).

$(3) \Longrightarrow (4)$: Suppose $[\phi]$ has no invariant cyclic subgroup system with index $d \geq 1$. Then, by Proposition 6.6, $[\phi]$ has a $(2,n)$-hyperbolic topological representative $f : \Gamma \to \Gamma$ defined on a marked graph (Γ, α). That is, for (any conjugacy class) of $x \in F$, we have

$$2 \, \|\phi^n(x)\|_\alpha \leq \max(\|\phi^{2n}(x)\|_\alpha, \|x\|_\alpha).$$

For any marked graph (Γ', β), choose difference of markings $g : \Gamma \to \Gamma'$ and $h : \Gamma' \to \Gamma$, i.e. graph maps such that $[\pi_1(g) \circ \alpha] = [\beta]$ and $[\pi_1(h) \circ \beta] = [\alpha]$. Then $K = \max(K(g), K(h))$ satisfies $\frac{1}{K}\|x\|_\beta \leq \|x\|_\alpha \leq K\|x\|_\beta$ for all $x \in F$, where $K(g), K(h)$ are the respective Lipschitz constants. Fix an integer $m \geq 1$ such that $2^m > K^2$, then for every $x \in F$,

$$\frac{2^m}{K}\|\phi^{mn}(x)\|_\beta \leq \max(K \, \|\phi^{2mn}(x)\|_\beta, K \, \|x\|_\beta)$$

and $(\Gamma', \beta, [\phi])$ is $(2^m K^{-2}, mn)$-hyperbolic, i.e. all topological representatives of $[\phi]$ on (Γ', β) are hyperbolic. Since $[\phi]$ is atoroidal, all topological representatives of $[\phi]$ on (Γ', β) are based-hyperbolic (Lemma 6.4). As (Γ', β) was arbitrary, all topological representatives of $[\phi]$ are based-hyperbolic.

Let $f' : \Gamma' \to \Gamma'$ be any topological representative for $[\phi]$. Since $[\phi]$ has no invariant cyclic subgroup with index $d \geq 2$, all strictly bidirectional annuli in $M_{f'}$ are shorter than some integer (Proposition 5.5 with Lemmas 6.3 and 6.2).

$(4) \Longrightarrow (1)$ — Theorem 6.1. See Theorem 7.7 below for a sketch of the proof. $\square$

7 HNN extensions over free factors

In the last section, we extend the previous characterization to HNN extensions of free groups defined over free factors. Precisely, let $A \leq F$ be a (nontrivial) free factor and $\phi : A \to F$ be an injective homomorphism. The HNN extension of F over A is given by:

$$F*_A = \langle F, t \mid t^{-1}xt = \phi(x), \forall x \in A \rangle.$$

HNN extensions, just like mapping tori, have associated Bass-Serre trees which we can use to define algebraic annuli as in the previous section. Unlike the Bass-Serre tree of a mapping torus, that of an HNN extension need not behave like a rooted tree: generally, the number of outgoing and incoming edges at a vertex will match the indices of A and $\phi(A)$ in F respectively. However, since free factors are malnormal, these HNN extensions share with mapping tori the dichotomy of annuli: annuli are either unidirectional or bidirectional.

We will say a conjugacy class of elements $[g]$ in F has <u>one forward iterate</u> if it contains an element $a \in A$. By malnormality of A, the conjugacy class of $[\phi(a)]$ depends only on the class $[g]$. We set the forward iterate of such a class $[g]$ to be $\phi \cdot [g] = [\phi(a)]$. Inductively, $[g]$ has $n + 1$ forward iterates if it has n forward iterates and the iterate $\phi^n \cdot [g]$ has one forward iterate; in this case, set $\phi^{n+1} \cdot [g] = \phi \cdot (\phi^n \cdot [g])$.

Similar to our use of infinite tails (backward iteration) to construct a canonical fixed free factor system in Section 2, we shall use forward iteration to construct a canonical invariant free factor system of A. The proof will use a simpler variation of descent.

Proposition 7.1. *If $A \leq F$ is a free factor and $\phi : A \to F$ is injective, then there is a unique maximal $[\phi]$-invariant free factor system of A. Precisely, $\mathcal{F}$ carries every conjugacy class with $2 \cdot \mathrm{rank}(A)$ forward iterates and every $[\phi]$-invariant subgroup system.*

Proof. Let $A \leq F$ be a free factor, $\phi : A \to F$ be injective, and $L(A) - 1 = 2 \cdot \mathrm{rank}(A) - 1$ be the length of the longest chain of free factor systems in A. Any conjugacy class in a $[\phi]$-invariant free factor system of A has infinitely many forward iterates. So if only the trivial conjugacy class has $L(A)$ forward iterates, then the trivial system is the unique $[\phi]$-invariant free factor system. Thus, we may assume some fixed nontrivial conjugacy class $[g]$ has $L(A)$ forward iterates. Since $A \leq F$ is a free factor and $\phi : A \to F$ is injective, we can iterate ϕ^{-1} on the poset of free factor systems of A!

Claim (Descent)**.** *For some $n \geq 0$, let $\mathcal{A}_n \prec \cdots \prec \mathcal{A}_0$ be the chain of nontrivial free factor systems of A such that $\mathcal{A}_i = \phi^{-i} \cdot \{A\}$ and $\mathcal{A}_n$ carries all $[\phi]$-invariant subgroup systems of A. If $\mathcal{A}_n$ is not $[\phi]$-invariant, then $\mathcal{A}_{n+1} = \phi^{-1} \cdot \mathcal{A}_n \prec \mathcal{A}_n$ is a nontrivial free factor system of A that carries all $[\phi]$-invariant subgroup systems.*

Starting with $\{A\}$, the descent will eventually stop at some $n < L(A) - 1$ with a nontrivial free factor system $\mathcal{A}_n = \phi^{-n} \cdot \{A\}$ that carries all conjugacy classes with $n + 1$ forward iterates and all $[\phi]$-invariant subgroup systems of A. Since descent stopped, $\mathcal{A}_n$ must be $[\phi]$-invariant and hence the unique maximal system amongst all $[\phi]$-invariant free factor systems of A. $\qquad\square$

Proof of descent. Let $n \geq 0$ and $\mathcal{A}_n \prec \cdots \prec A_0 = \{A\}$ be the chain of nontrivial free factor systems such that $\mathcal{A}_i = \phi^{-i} \cdot \{A\}$ and $\mathcal{A}_n$ carries all $[\phi]$-invariant subgroup systems of A. Nontriviality of $\mathcal{A}_n$ implies $n < L(A) - 1$. Now suppose that $\mathcal{A}_n$ is not $[\phi]$-invariant.

The existence of a nontrivial conjugacy class $[g]$ with $L(A) \geq n + 2$ forward iterates implies $\mathcal{A}_{n+1} = \phi^{-1} \cdot \mathcal{A}_n$ is a nontrivial free factor system of A. If $n = 0$, then clearly

$\mathcal{A}_{n+1} \preceq \mathcal{A}_n$. If $n > 1$, then $\mathcal{A}_n \prec \mathcal{A}_{n-1}$ by assumption; hence, $\phi^{-1} \cdot \mathcal{A}_n \preceq \phi^{-1} \cdot \mathcal{A}_{n-1}$ and, equivalently, $\mathcal{A}_{n+1} \preceq \mathcal{A}_n$. If $\mathcal{A}_n = \mathcal{A}_{n+1}$, then $\phi(\mathcal{A}_n)$ is carried by $\mathcal{A}_n$; but $\mathcal{A}_n$ is not $[\phi]$-invariant, thus $\mathcal{A}_{n+1} \prec \mathcal{A}_n$. Let $\mathcal{B}$ be any $[\phi]$-invariant subgroup system in A. By assumption, $\mathcal{A}_n$ carries $\mathcal{B}$ and $\phi(\mathcal{B})$; therefore, $\mathcal{A}_{n+1} = \phi^{-1} \cdot \mathcal{A}_n$ carries $\mathcal{B}$. $\qquad\square$

The unique $[\phi]$-invariant free factor system of A given by Proposition 7.1, denoted by $\mathcal{F}$, is the canonical $[\phi]$-invariant free factor system and it captures the long-term dynamics of $[\phi]$. The system $\mathcal{F}$ is proper (in A) if and only if A is not $[\phi]$-invariant; it is trivial exactly when only the trivial conjugacy class has $L(A) = 2 \cdot \operatorname{rank}(A)$ forward iterates. The canonical invariant free factor system allows us to naturally extend definitions and results that required iteration of an injective endomorphism. For instance, we can now say $\phi : A \to F$ has an invariant cyclic subgroup system with index $d \geq 1$ if a restriction $\phi|_{\mathcal{F}} : \mathcal{F} \to \mathcal{F}$ has an invariant cyclic subgroup system with index $d \geq 1$. In particular, $[\phi]$ is atoroidal if $[\phi|_{\mathcal{F}}]$ is atoroidal.

For $k \geq L(A)$, we define the iterated pullbacks $\Lambda_k[\phi]$ of $[\phi]$ to be the iterated pullbacks $\Lambda_k[\phi|_{\mathcal{F}}]$ of the restriction $[\phi|_{\mathcal{F}}]$. Similarly define $\hat{\Lambda}_k$ for $k \geq L(A)$. Since connectedness did not play any role in the proof of Proposition 5.5, we immediately get the following extension when we replace endomorphisms of F with endomorphisms of $\mathcal{F}$.

Proposition 7.2. *If $A \leq F$ is a free factor, $\phi : A \to F$ is injective, and $\hat{\Lambda}_k[\phi]$ is not empty for all $k \geq L(A)$, then ϕ has an invariant cyclic subgroup system with index $d \geq 2$.*

Remark. By the same token, the results in this section do not need A to be "connected." So there is a natural generalization of these results that readily follows if we replace the free factor $A \leq F$ with a free factor system $\mathcal{A}$ of F.

For the rest of the section, we will extend the results of Section 6 to the injective homomorphism $\phi : A \to F$. Fix a marked graph (Γ, α) with no bivalent vertices such that the free factor systems $\mathcal{F}$ and A correspond to nested core subgraphs $\Gamma_{\mathcal{F}} \subset \Gamma_A \subset \Gamma$ respectively, i.e. $\Gamma_{\mathcal{F}}$ and Γ_A have markings $\alpha_{\mathcal{F}} : \mathcal{F} \to \pi_1(\Gamma_{\mathcal{F}})$ and $\alpha_A : A \to \pi_1(\Gamma_A)$ respectively such that the inclusion maps $c_{\mathcal{F}} : \Gamma_{\mathcal{F}} \to \Gamma_A$ and $c_A : \Gamma_A \to \Gamma$ satisfy $[\pi_1(c_{\mathcal{F}}) \circ \alpha_{\mathcal{F}}] = [\alpha_A|_{\mathcal{F}}]$ and $[\pi_1(c_A) \circ \alpha_A] = [\alpha|_A]$. A topological representative for $[\phi]$ will be a graph map $f : (\Gamma_A, \Gamma_{\mathcal{F}}) \to (\Gamma, \Gamma_{\mathcal{F}})$ with no pretrivial edges such that $[\pi_1(f) \circ \alpha_A] = [\alpha \circ \phi]$. Thus, the invariant restriction $f|_{\mathcal{F}} : \Gamma_{\mathcal{F}} \to \Gamma_{\mathcal{F}}$ is a topological representative for the restriction $[\phi|_{\mathcal{F}}]$.

The following is the analogue of the mapping torus. Let f be a topological representative for $\phi : A \to F$. We define the classifying space to be $M_f = (\Gamma \sqcup (\Gamma_A \times [0,1])) / \sim$ where we identify $(x, 0) \sim x$ and $(x, 1) \sim f(x)$ for all $x \in \Gamma_A \subset \Gamma$. The edge-space of M_f will be the cross-section represented by $\Gamma_A \times \{\frac{1}{2}\}$. By construction, $\pi_1(M_f) \cong F*_A$. Topological annuli in M_f and algebraic annuli in $F*_A$ are defined exactly as before and the correspondence between them is the same. Similarly, the correspondence between strictly bidirectional annuli and the iterated pullbacks remains. We then get this natural extension of Lemma 6.3.

Lemma 7.3. *Let $A \leq F$ be a free factor and $\phi : A \to F$ be injective. For any integer $L \geq L(A)$, the HNN extension $F*_A$ has a strictly bidirectional annulus of length $2L$ if and only if $\hat{\Lambda}_L[\phi]$ is not empty.*

The definition of (based-) hyperbolicity extends naturally using the invariant restriction. For $\lambda > 1$ and $n \geq L(A)$, a topological representative $f : (\Gamma_A, \Gamma_{\mathcal{F}}) \to (\Gamma, \Gamma_{\mathcal{F}})$ for $[\phi]$ is (based-) (λ, n)-hyperbolic if $f|_{\mathcal{F}} : \Gamma_{\mathcal{F}} \to \Gamma_{\mathcal{F}}$ is (based-) (λ, n)-hyperbolic and $[f]$ is atoroidal if $[f|_{\mathcal{F}}]$ is atoroidal. When $\mathcal{F}$ is trivial, any topological representative for $[\phi]$ is vacuously (based-) hyperbolic and atoroidal. Again, since connectedness played no role in the proofs of Lemma 6.5 and Proposition 6.6, we get this extension:

Proposition 7.4. *Let $A \leq F$ be a free factor. If $\phi : A \to F$ is injective and atoroidal, then $[\phi]$ has a $(2, n)$-hyperbolic topological representative for some integer $n \geq L(A)$.*

We are now ready to state and almost prove the main result of the section.

Theorem 7.5. *Let $A \leq F$ be a free factor and $\phi : A \to F$ be an injective homomorphism. Then the following are equivalent:*

1. *$F*_A$ is word-hyperbolic;*

2. *$F*_A$ contains no $BS(1, d)$ subgroups with $d \geq 1$;*

3. *$[\phi]$ has no invariant cyclic subgroup system with index $d \geq 1$;*

4. *all/some topological representative of $[\phi]$ are based-hyperbolic and all strictly bidirectional annuli in their classifying spaces are shorter than some integer.*

The proof follows the same steps as Theorem 6.7 and the only missing ingredient is an extension of Theorem 6.1 (Theorem 7.7 below) that proves the implication (4) $\implies$ (1).

Remark. Every HNN extension of the form $F*_A$ where F is a finite rank free group and $A \leq F$ is a free factor can be written as a mapping torus $\mathbb{F}*_\psi$ of an injective endomorphism $\psi : \mathbb{F} \to \mathbb{F}$ where $\mathbb{F}$ is a free group with possibly infinite rank. Conversely, every finitely generated subgroup of a mapping torus $\mathbb{F}*_\psi$ where $\mathbb{F}$ has possibly infinite rank can be written as an HNN extension $F*_A$ where F has finite rank and $A \leq F$ is a free factor [10]. This argument is Feighn-Handel's result that $F*_A$ and $\mathbb{F}*_\psi$ are *coherent*. As a corollary, we get the following amusing statement:

Corollary 7.6. *Let $\psi : \mathbb{F} \to \mathbb{F}$ be an injective endomorphism of an infinite rank free group. $\mathbb{F}*_\psi$ is locally word-hyperbolic if and only if it contains no $BS(1, d)$ subgroups with $d \geq 1$.*

By *locally word-hyperbolic*, we mean every finitely generated subgroup is word-hyperbolic.

The following extension will complete our proof of Theorem 7.5.

Theorem 7.7. *Let $A \leq F$ be a free factor and $\phi : A \to F$ be injective. If a topological representative f of $[\phi]$ is based-hyperbolic and all strictly bidirectional annuli in its classifying space M_f are shorter than some integer, then $F *_A$ is word-hyperbolic.*

Roughly speaking, the extension follows from the fact that annuli of M_f longer than $L(A)$ are annuli of the mapping torus of $f|_{\mathcal{F}}$ up to some controlled error. Due to the close similarity with the proof of Theorem 6.1 given in [24], we will only sketch a proof of the extension.

We start by defining *annuli flaring*; let $h : S^1 \times [-M, M] \to M_f$ be a topological annulus in M_f of length $2M$. The girth of h is the combinatorial length $|h_0|$ of the middle ring $h_0 = h(\cdot, 0) : S^1 \to M_f$ in the edge-space. Let $\lambda > 1$ be a real number; the annulus h is $\underline{\lambda\text{-hyperbolic}}$ if

$$\lambda |h_0| \leq \max(|h_{-L}|, |h_L|).$$

For integers i between $-M$ and $M - 1$ (inclusive), define $\tau_i : (i, i + 1) \to \Gamma$ by projecting the trace $h^\nu|_{(i,i+1)}$ to the vertex-space Γ as follows:

$$\tau_i(t) = \begin{cases} x & \text{if } h^\nu(t) = (x, s) \in \Gamma_A \times [0, 1] \text{ and } s < \frac{1}{2} \\ f(x) & \text{if } h^\nu(t) = (x, s) \in \Gamma_A \times [0, 1] \text{ and } s > \frac{1}{2} \\ x & \text{if } h^\nu(t) = x \in \Gamma \end{cases}$$

We say h is $\underline{\rho\text{-thin}}$ if $|\tau_i| + 1 \leq \rho$ for all i, where $|\cdot|$ is the length after tightening the path rel. endpoints. We say M_f satisfies the annuli flaring condition if there are $\lambda > 1, M \geq L(A)$, and a function $H : \mathbb{R} \to \mathbb{R}$ such that any ρ-thin annulus of length $2M$ with girth at least $H(\rho)$ is λ-hyperbolic. Bestvina-Feighn's combination theorem [3] states that $\pi_1(M_f)$ is word-hyperbolic if M_f satisfies the annuli flaring condition.

Sketch proof of Theorem 7.7. Suppose $A \leq F$ is a free factor, $\phi : A \to F$ is injective, and $\mathcal{F}$ is the canonical $[\phi]$-invariant free factor system of A. Let $f : (\Gamma_A, \Gamma_{\mathcal{F}}) \to (\Gamma, \Gamma_{\mathcal{F}})$ be a based-hyperbolic topological representative for $[\phi]$ and $K = K(f) > 1$ be some Lipschitz constant for f. If $\mathcal{F}$ is trivial, then all unidirectional annuli in M_f are shorter than $L(A)$. So M_f vacuously satisfies the annuli flaring condition with $M = L(A)$ by the annulus dichotomy (malnormality of A). Thus, we may assume $\mathcal{F}$ is not trivial. The assumptions and tool introduced in the next two paragraphs are needed in case A is not $[\phi]$-invariant.

Let $\mathcal{A}_n \prec \cdots \prec \mathcal{A}_0 \prec \mathcal{A}_{-1}$ be the chain where $\mathcal{A}_{-1} = \{F\}, \mathcal{A}_0 = \{A\}, \mathcal{A}_n = \mathcal{F}$, and generally $\mathcal{A}_i = \phi^{-1} \cdot \mathcal{A}_{i-1}$ for $0 \leq i \leq n$. By invariance of $\mathcal{A}_n$, we can extend the chain by setting $\mathcal{A}_i = \mathcal{A}_n$ for $i > n$. For $0 \leq i < n$, $\mathcal{A}_{i-1}$ carries $\phi(\mathcal{A}_i)$ by definition. We assume Γ has a filtration of nonempty core subgraphs $\cdots \subset \Gamma_1 \subset \Gamma_0 \subset \Gamma_{-1}$ where $\Gamma_{-1} = \Gamma_{\mathcal{F}} \vee R$ is a wedge between $\Gamma_{\mathcal{F}}$ and a rose R and each Γ_i is a $\mathcal{A}_i$-marked graph. If $n > 0$, we will also assume the topological representative $f : (\Gamma_0, \Gamma_n) \to (\Gamma_{-1}, \Gamma_n)$ is of the form $(\Gamma_0, \ldots, \Gamma_{n-1}, \Gamma_n) \to (\Gamma_{-1}, \ldots, \Gamma_{n-2}, \Gamma_n)$. Note that since $\Gamma_n = \Gamma_{\mathcal{F}}$ was f-invariant to begin with, changes can made to the given representative f to allow for these

additional assumptions without changing the invariant restriction $f|_{\mathcal{F}}$. In particular, f is still based-hyperbolic after the changes.

Using this filtration, we introduce a tool needed to iterate based loops not in $\Gamma_{\mathcal{F}}$: for any integer $i \geq 0$ and any immersed based loop $s : S^1 \to \Gamma_0$ that is freely homotopic into $\Gamma_{\mathcal{F}}$, there is a "closest point" projection (with respect to Hausdorff distance of lifts to the universal cover) of s to an immersed based loop in Γ_0 denoted by $\lfloor s \rfloor_i : S^1 \to \Gamma_0$ freely homotopic to s such that $f^i \circ \lfloor s \rfloor_i$ is a based loop in Γ_0. Recall that based loops must send basepoints to vertices. Since we assumed the "complement" of Γ_n is a rose, the projections $\lfloor s \rfloor_i$ are immersed based loops in Γ_i. So $\lfloor s \rfloor_i = \lfloor s \rfloor_n$ for $i \geq n$. Implicit in this is the fact that, for all $i \geq 0$, any conjugacy class carried by $\mathcal{A}_i$ but not $\mathcal{A}_{i+1}$ is mapped by ϕ to a conjugacy class carried by $\mathcal{A}_{i-1}$ but not $\mathcal{A}_i$. We are ready to start proving the theorem.

Suppose f is based-$(2, m)$-hyperbolic for some $m \geq L(A)$ and strictly bidirectional annuli in M_f are shorter than $2L \geq 2L(A)$. Fix $r \geq 1$ such that $2^{L+r} \geq 3 \cdot K^{Lm}$. Now set $M = (2L + r)m$ and

$$H(\rho) = 2^M \cdot 4\rho \cdot \frac{K^{2M} - 1}{K - 1}.$$

Suppose h is an arbitrary ρ-thin annulus of length $2M$ with girth $|h_0| \geq H(\rho)$. We need to show that h is 2-hyperbolic. Without loss of generality, assume the truncation $h|_{[-M,0]}$ is unidirectional and *increasing*.

For $-M \leq k \leq 0$, consider the rings of $h_k : S^1 \to M_f$ in the edge-space $\Gamma_A \times \{\frac{1}{2}\}$. This can be viewed as an immersed based loop $h_k : S^1 \to \Gamma_0$ by forgetting the $\{\frac{1}{2}\}$-factor. Since $M \geq L(A)$ and $h_{-M} \simeq_{M_f} h_0$ (free homotopy in M_f) is not trivial, the ring h_{-M} is freely homotopic (in Γ) into $\Gamma_n = \Gamma_{\mathcal{F}}$. Since Γ_n is f-invariant and $h_{-M} \simeq_{M_f} h_k$ for all $k \leq 0$, the rings h_k are freely homotopic into Γ_n. In particular, these rings all have projections to subgraphs Γ_i for all i.

For the first step, project h_{-1} to Γ_1 so that $f \circ \lfloor h_{-1} \rfloor_1 : S^1 \to \Gamma_0$ is a based loop freely homotopic to h_0 in Γ. The ρ-thin assumption (and "closest point" projection) means: $|f \circ \lfloor h_{-1} \rfloor_1| \geq |h_0| - 2\rho$. Similarly, we project h_{-2} to Γ_2 so that $f \circ \lfloor h_{-2} \rfloor_2 : S^1 \to \Gamma_1$ is a based loop freely homotopic to $\lfloor h_{-1} \rfloor_1$ in Γ. Then $|f \circ \lfloor h_{-2} \rfloor_2| \geq |\lfloor h_{-1} \rfloor_1| - 2\rho$ by ρ-thinness and, by the K-Lipschitz property of f,

$$|f^2 \circ \lfloor h_{-2} \rfloor_2| \geq |f \circ \lfloor h_{-1} \rfloor_1| - 2\rho \cdot K \geq |h_0| - 2\rho \cdot (1 + K).$$

By induction, the f-invariance of Γ_n, and the fact $M > L(A) > n$, we get this inequality:

$$|f^M \circ \lfloor h_{-M} \rfloor_n| \geq |h_0| - 2\rho \cdot (1 + \cdots + K^{M-1}) = |h_0| - 2\rho \cdot \frac{K^M - 1}{K - 1}.$$

Since f is based-$(2, m)$-hyperbolic, we know that

$$2|f^M \circ \lfloor h_{-M} \rfloor_n| \leq \max(|f^{M+m} \circ \lfloor h_{-M} \rfloor_n|, |f^{M-m} \circ \lfloor h_{-M} \rfloor_n|).$$

54

There are three cases to consider of which we will only prove one.

Case 1: Suppose $2|f^M \circ \lfloor h_{-M} \rfloor_n| \leq |f^{M-m} \circ \lfloor h_{-M} \rfloor_n|$.

Then by induction on based-$(2, m)$-hyperbolicity, we get:

$$
\begin{aligned}
|h_{-M}| \geq |\lfloor h_{-M} \rfloor_n| &\geq 2^{2L+r} |f^M \circ \lfloor h_{-M} \rfloor_n| \\
&\geq 2^{2L+r} |h_0| - 2^{2L+r} \cdot 2\rho \cdot \frac{K^M - 1}{K - 1} \\
&\geq 3|h_0| - H(\rho) \\
&\geq 2|h_0| \qquad\qquad\qquad\qquad\qquad \text{as } |h_0| \geq H(\rho).
\end{aligned}
$$

So h is 2-hyperbolic in this case. The proof so far has been nearly identical to the proof of Theorem 6.1 in [24]. But here, we needed to be careful when applying iterates of f to the rings h_k since A may not be $[\phi]$-invariant. This is why we introduced the projections $\lfloor \cdot \rfloor_i$.

There are two more cases remaining.

Case 2: Suppose $2|f^M \circ \lfloor h_{-M} \rfloor_n| \leq |f^{M+m} \circ \lfloor h_{-M} \rfloor_n|$ and h is unidirectional.

Case 3: Suppose $2|f^M \circ \lfloor h_{-M} \rfloor_n| \leq |f^{M+m} \circ \lfloor h_{-M} \rfloor_n|$ and h is bidirectional.

Case 3 is where the bound on strictly bidirectional annuli is needed. We leave the details of these cases to the reader. Alternatively, one could read the proof of Theorem 6.1, compare how Case 1 was handled in the two proofs, and adjust the proofs of the remaining cases accordingly. We have covered all the cases and M_f satisfies the annuli flaring condition. By the combination theorem, $F *_A \cong \pi_1(M_f)$ is word-hyperbolic. $\qquad\square$

Epilogue

We would like to conclude this paper with a discussion of a few questions: the first question is a natural generalization of Section 7; the rest are problems from Ilya Kapovich's paper [18, Section 6] that could be answered using expanding relative immersions.

Problem 1. Suppose $A \leq F$ is a vertex group of a cyclic splitting of F and $\phi : A \to F$ is injective. Is $F *_A$ word-hyperbolic if it contains no $BS(1, d)$ subgroups for $d \geq 1$? Can it have $BS(m, d)$ subgroups with $m, d > 1$?

A *cyclic splitting* of F is an edge of groups decomposition of F with a nontrivial cyclic edge group. Vertex groups of cyclic splittings are generalizations of free factors and it would be interesting to see if the ideas in Section 7 can be adapted to this case. The first obstacle is finding the appropriate generalization of Proposition 7.1 since ϕ^{-1}-iteration need not be as well-behaved. Furthermore, the vertex groups A are not always malnormal which means annuli in $F *_A$ could exhibit more complicated behavior.

For the remaining problems, assume $\phi : F \to F$ is injective.

Problem 2 ([18, Problem 6.4])**.** What kind of isoperimetric functions can $F *_\phi$ have?

The automorphism case of this problem was answered by Bridson-Groves [7]: they showed that free-by-cyclic groups have quadratic isoperimetric functions. Implicit in the second part of the paper is the idea that $F*_\phi$ is hyperbolic relative to a canonical finite collection of free-by-cyclic groups when it has no $BS(1, d)$ subgroups for $d \geq 2$. Furthermore, when $A \leq F$ is a free factor and $\psi : A \to F$ is injective, then $F*_A$ is hyperbolic relative to a canonical finite collection of ascending HNN extensions. This would imply that $F*_A$ has a quadratic isoperimetric function when it has no $BS(1, d)$ subgroups for $d \geq 2$. We intend to complete this direction in future work by employing a combination theorem for relatively hyperbolic groups [13, 23].

When $F*_\phi$ is word-hyperbolic, Mahan Mj proved there is a continuous extension to the Gromov boundary of the inclusion map $F \leq F*_\phi$ [22]; this map is known as the *Cannon-Thurston map*.

Problem 3 ([18, Problem 6.7]). If $F*_\phi$ is word-hyperbolic, is the Cannon-Thurston map (uniformly) finite-to-one? Is there a corresponding ending laminations theorem?

For the first of these questions, it seems that expanding relative immersions reduce the problem to the automorphism case, which has been answered [14, 21]. As for the ending laminations, what is missing is the appropriate formulation of the theorem that replaces the usual short exact sequence with a graph of groups decomposition.

Problem 4 ([18, Problem 6.8]). Let $\partial\phi : \partial F \to \partial F$ be the extension of ϕ to the boundary. Can we classify injective endomorphisms φ by the dynamics of boundary extensions $\partial\varphi$?

Levitt-Lustig covered the surjective case when they proved that *most* automorphisms of non-elementary word-hyperbolic groups have boundary extensions with *north-south dynamics* [19]: two fixed points — a repellor and an attractor. Expanding relative immersions should imply that *most* injective nonsurjective endomorphisms of F have *sink dynamics*: a single fixed point and it is an attractor.

There are of course more unresolved questions about injective endomorphisms and their expanding relative immersions. These are left for the reader to ask and answer.

A   Relative train tracks

The objective in this appendix is to sketch the proof that irreducible relative representatives with minimal stretch factor are train tracks. Bestvina-Handel's construction of train tracks for irreducible automorphisms [6] translates verbatim to the non-free forest setting.

Let $\phi : F \to F$ be an injective endomorphism, $\mathcal{A} \prec \mathcal{B}$ be $[\phi]$-invariant free factor systems such that $\phi^{-1} \cdot \mathcal{A} = \mathcal{A}$, and T_* be a $(\mathcal{B}, \mathcal{A})$-forest possibly with bivalent vertices. Recall, an $\mathcal{A}$-*relative weak representative* for the restriction $\phi|_{\mathcal{B}}$ is a $\phi|_{\mathcal{B}}$-equivariant graph map $f_* : T_* \to T_*$; and an $\mathcal{A}$-*relative representative* is an $\mathcal{A}$-relative weak representative f_*

with no pretrivial edges and whose underlying forest T_* has no bivalent vertices. Finally, we say a relative weak representative is *minimal* if it has no orbit-closed invariant subforests with bounded components.

An $\mathcal{A}$-relative weak representative f_* is *irreducible* if the *transition matrix* $A(f_*)$ is irreducible and, in this case, the *stretch factor* $\lambda(f_*)$ is the Perron-Frobenius eigenvalue of $A(f_*)$. An $\underline{\mathcal{A}\text{-relative train track}}$ for $\phi|_{\mathcal{B}}$ is an $\mathcal{A}$-relative representative f_* for $\phi|_{\mathcal{B}}$ that additionally satisfies the property: the edge-paths $f_*^n(e)$ are immersed for all edges e in T_* and integers $n \geq 1$. We have set the stage for the Bestvina-Handel's result.

Theorem A.1 ([6, Theorem 1.7]). *Let $\phi|_{\mathcal{B}}$ be irreducible relative to $\mathcal{A}$ and $f_* : T_* \to T_*$ be an irreducible $\mathcal{A}$-relative representative for $\phi|_{\mathcal{B}}$. If f_* has minimal stretch factor, then it is an irreducible $\mathcal{A}$-relative train track.*

The minimality is amongst irreducible representatives rather than irreducible weak representatives. The argument relies on understanding how various moves on an irreducible $\mathcal{A}$-relative representative f_* affect $\lambda(f_*)$ and invoking minimality of $\lambda(f_*)$ to conclude that no such moves are possible. Note that although the moves are described locally, they must be performed equivariantly if we want the resultant forests to be $(\mathcal{B}, \mathcal{A})$-forests. The proofs of these moves/lemmas are omitted since they are the same as the proofs in [6].

Remark. The reviewer pointed out to us that this outline is also the strategy carried out by Dicks-Ventura [9]. Recently, Bestvina [2] and Francaviglia-Martino [11] gave an alternative approach to proving this theorem using the *Lipschitz metric* on *relative outer space*.

The first move is *subdivision*, which occurs at an interior point of an edge that is in the preimage of vertices under the weak representative.

Lemma A.2 ([6, Lemma 1.10]). *If $f_* : T_* \to T_*$ is an irreducible $\mathcal{A}$-relative weak representative for $\phi|_{\mathcal{B}}$ and $f_*' : T_*' \to T_*'$ is induced by a subdivision, then f_*' is an irreducible $\mathcal{A}$-relative weak representative and $\lambda(f_*') = \lambda(f_*)$.*

The next move is *bivalent homotopy*, which decreases the number of edges.

Lemma A.3 ([6, Lemma 1.13]). *If $f_* : T_* \to T_*$ is an irreducible $\mathcal{A}$-relative weak representative for $\phi|_{\mathcal{B}}$ and $f_*'' : T_*'' \to T_*''$ is an irreducible $\mathcal{A}$-relative weak representative induced by a bivalent homotopy followed by collapse of a maximal invariant subforest with bounded components, then $\lambda(f_*'') \leq \lambda(f_*)$.*

The last move we need is *folding*, which occurs between a pair of oriented edges originating from the same vertex that have the same image under the weak representative.

Lemma A.4 ([6, Lemma 1.15]). *Suppose $f_* : T_* \to T_*$ is an irreducible $\mathcal{A}$-relative weak representative for $\phi|_{\mathcal{B}}$ and $f_*' : T_*' \to T_*'$ is induced by a fold. If f_*' is an $\mathcal{A}$-relative weak representative, then it is irreducible and $\lambda(f_*') = \lambda(f_*)$. Otherwise, if $f_*'' : T_*'' \to T_*''$ is an irreducible $\mathcal{A}$-relative weak representative induced by a homotopy of f_*' that makes the final map locally injective on the interior of edges, followed by collapse of a maximal invariant subforest with bounded components, then $\lambda(f_*'') < \lambda(f_*)$.*

Sketch proof of Theorem A.1. Let $\phi|_\mathcal{B}$ be irreducible relative to $\mathcal{A}$ and $f_* : T_* \to T_*$ be an irreducible $\mathcal{A}$-relative representative. If $\lambda(f_*) = 1$, then f_* is a simplicial embedding with minimal stretch factor and we are done. So we may assume $\lambda(f_*) > 1$.

Suppose for the contrapositive that f_* is not an $\mathcal{A}$-relative train track, then the edge-path $f_*^n(e)$ is not immersed for some edge e in T_* and integer $n \geq 1$. Let n be the smallest such integer and assume $\star$ is an interior point of an edge e at which f_*^n fails to be locally injective. Appropriately subdivide T_* so that a neighborhood U of $\star$ and its iterates $f_*^k(U)$ $(1 \leq k \leq n)$ satisfy nice properties: 1) U is an interval whose boundary consists of distinct vertices; 2) f_*^k is locally injective on U for $1 \leq k < n$; 3) f_*^n folds U at $\star$ to an edge; and 4) $\star \notin f_*^k(U)$ for $1 \leq k \leq n$. Then iteratively fold $f_*^{n-1}(U)$, ..., $f_*^2(U)$, and $f_*(U)$. By minimality of n, all the folds except the last one induce an irreducible $\mathcal{A}$-relative weak representative. By the first case of Lemma A.4, this irreducible $\mathcal{A}$-relative weak representative has the same stretch factor as f_*. By construction, the last fold induces a map f_*' that fails to be an $\mathcal{A}$-relative weak representative as it fails to be locally injective at $\star$. Apply a tightening homotopy on f_*' to make it locally injective at $\star$, then collapse a maximal invariant subforest with bounded components to get $f_*'' : T_*'' \to T_*''$, a minimal $\mathcal{A}$-relative weak representative for $\phi|_\mathcal{B}$. By Lemma 3.6, the map f_*'' is irreducible. By the second case of Lemma A.4, the stretch factor is strictly smaller: $\lambda(f_*'') < \lambda(f_*)$. Sequentially apply bivalent homotopies and collapse maximal invariant subforests with bounded components to get an irreducible $\mathcal{A}$-relative representative f_*'''. The stretch factor satisfies $\lambda(f_*''') \leq \lambda(f_*'') < \lambda(f_*)$ by Lemma A.3 and so $\lambda(f_*)$ was not minimal. $\qquad\square$

References

[1] Mladen Bestvina. Questions in geometric group theory. *preprint*, 2004.

[2] Mladen Bestvina. A Bers-like proof of the existence of train tracks for free group automorphisms. *Fund. Math.*, 214(1):1–12, 2011.

[3] Mladen Bestvina and Mark Feighn. A combination theorem for negatively curved groups. *J. Differential Geom.*, 35(1):85–101, 1992.

[4] Mladen Bestvina, Mark Feighn, and Michael Handel. Laminations, trees, and irreducible automorphisms of free groups. *Geom. Funct. Anal.*, 7(2):215–244, 1997.

[5] Mladen Bestvina, Mark Feighn, and Michael Handel. The Tits alternative for $\mathrm{Out}(F_n)$. I. Dynamics of exponentially-growing automorphisms. *Ann. of Math. (2)*, 151(2):517–623, 2000.

[6] Mladen Bestvina and Michael Handel. Train tracks and automorphisms of free groups. *Ann. of Math. (2)*, 135(1):1–51, 1992.

[7] Martin R. Bridson and Daniel Groves. The quadratic isoperimetric inequality for mapping tori of free group automorphisms. *Mem. Amer. Math. Soc.*, 203(955):xii+152, 2010.

[8] Peter Brinkmann. Hyperbolic automorphisms of free groups. *Geom. Funct. Anal.*, 10(5):1071–1089, 2000.

[9] Warren Dicks and Enric Ventura. *The group fixed by a family of injective endomorphisms of a free group*, volume 195 of *Contemporary Mathematics*. American Mathematical Society, Providence, RI, 1996.

[10] Mark Feighn and Michael Handel. Mapping tori of free group automorphisms are coherent. *Ann. of Math. (2)*, 149(3):1061–1077, 1999.

[11] Stefano Francaviglia and Armando Martino. Stretching factors, metrics and train tracks for free products. *Illinois J. Math.*, 59(4):859–899, 2015.

[12] Joel Friedman. Sheaves on graphs, their homological invariants, and a proof of the Hanna Neumann conjecture: with an appendix by Warren Dicks. *Mem. Amer. Math. Soc.*, 233(1100):xii+106, 2015.

[13] François Gautero. Geodesics in trees of hyperbolic and relatively hyperbolic spaces. *Proc. Edinb. Math. Soc. (2)*, 59(3):701–740, 2016.

[14] Pritam Ghosh. Limits of conjugacy classes under iterates of hyperbolic elements of $\mathrm{Out}(\mathbb{F})$. *Groups Geom. Dyn.*, 14(1):177–211, 2020.

[15] Misha Gromov. Hyperbolic groups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987.

[16] Mark F. Hagen and Daniel T. Wise. Cubulating hyperbolic free-by-cyclic groups: the general case. *Geom. Funct. Anal.*, 25(1):134–179, 2015.

[17] Mark F. Hagen and Daniel T. Wise. Cubulating hyperbolic free-by-cyclic groups: the irreducible case. *Duke Math. J.*, 165(9):1753–1813, 2016.

[18] Ilya Kapovich. Mapping tori of endomorphisms of free groups. *Comm. Algebra*, 28(6):2895–2917, 2000.

[19] Gilbert Levitt and Martin Lustig. Most automorphisms of a hyperbolic group have very simple dynamics. *Ann. Sci. École Norm. Sup. (4)*, 33(4):507–517, 2000.

[20] Igor Mineyev. Submultiplicativity and the Hanna Neumann conjecture. *Ann. of Math. (2)*, 175(1):393–414, 2012.

[21] Mahan Mitra. Ending laminations for hyperbolic group extensions. *Geom. Funct. Anal.*, 7(2):379–402, 1997.

[22] Mahan Mitra. Cannon-Thurston maps for trees of hyperbolic metric spaces. *J. Differential Geom.*, 48(1):135–164, 1998.

[23] Mahan Mj and Lawrence Reeves. A combination theorem for strong relative hyperbolicity. *Geom. Topol.*, 12(3):1777–1798, 2008.

[24] Jean Pierre Mutanguha. Hyperbolic immersions of free groups. *Groups Geom. Dyn.*, 14(4):1253–1275, 2020.

[25] Jean Pierre Mutanguha. Irreducible nonsurjective endomorphisms of F_n are hyperbolic. *Bull. Lond. Math. Soc.*, 52(5):960–976, 2020.

[26] Hanna Neumann. On the intersection of finitely generated free groups. Addendum. *Publ. Math. Debrecen*, 5:128, 1957.

[27] Walter D. Neumann. On intersections of finitely generated subgroups of free groups. In *Groups—Canberra 1989*, volume 1456 of *Lecture Notes in Math.*, pages 161–170. Springer, Berlin, 1990.

[28] Patrick R. Reynolds. *Dynamics of irreducible endomorphisms of F_N*. PhD thesis, University of Illinois at Urbana-Champaign, 2011.

[29] Jean-Pierre Serre. *Arbres, amalgames,* SL$_2$. Société Mathématique de France, Paris, 1977. Avec un sommaire anglais, Rédigé avec la collaboration de Hyman Bass, Astérisque, No. 46.

[30] John R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.

[31] William P. Thurston. Three-dimensional manifolds, Kleinian groups and hyperbolic geometry. *Bull. Amer. Math. Soc. (N.S.)*, 6(3):357–381, 1982.

[32] William P. Thurston. Hyperbolic structures on 3-manifolds. I. Deformation of acylindrical manifolds. *Ann. of Math. (2)*, 124(2):203–246, 1986.