The structure and density of k-product-free sets in the free semigroup and group

Freddie Illingworth[†] Lukas Michel[‡] Alex Scott[‡]

March 28, 2025

Abstract

The free semigroup $\mathcal{F}$ on a finite alphabet $\mathcal{A}$ is the set of all finite words with letters from $\mathcal{A}$ equipped with the operation of concatenation. A subset S of $\mathcal{F}$ is k-product-free if no element of S can be obtained by concatenating k words from S, and strongly k-product-free if no element of S is a (non-trivial) concatenation of at most k words from S.

We prove that a k-product-free subset of $\mathcal{F}$ has upper Banach density at most $1/\rho(k)$, where $\rho(k) = \min\{\ell \colon \ell \nmid k-1\}$. We also determine the structure of the extremal k-product-free subsets for all $k \notin \{3, 5, 7, 13\}$; a special case of this proves a conjecture of Leader, Letzter, Narayanan, and Walters. We further determine the structure of all strongly k-product-free sets with maximum density. Finally, we prove that k-product-free subsets of the free group have upper Banach density at most $1/\rho(k)$, which confirms a conjecture of Ortega, Rué, and Serra.

1 Introduction

A subset S of a (semi)group G is said to be *product-free* if $x \cdot y \notin S$ for all $x, y \in S$. Two very natural questions present themselves.

Density: How dense can the largest product-free subset of G be?

Structure: What is the structure of the densest product-free subsets of G?

These problems have been extensively studied over the last fifty years. In the finite abelian case, this culminated in a solution to the density problem by Green and Ruzsa [GR05] and the structure problem by Balasubramian, Prakash, and Ramana [BPR16]. The finite non-abelian case was first investigated by Babai and Sós [BS85]. This case behaves very differently with the possibility of a largest product-free subsets having vanishing density as shown by the seminal work of Gowers [Gow08] on quasirandom groups. Recent breakthroughs include the alternating group where Eberhard [Ebe16] solved the density problem (up to logarithmic factors) and Keevash,

2020 MSC: 20M05 (free semigroups), 20E05 (free non-abelian groups), 05D05 (extremal set theory).

[†]Department of Mathematics, University College London, United Kingdom (f.illingworth@ucl.ac.uk).
The research was done while FI was at Oxford being supported by EPSRC grant EP/V007327/1.

[‡]Mathematical Institute, University of Oxford, United Kingdom ({michel,scott}@maths.ox.ac.uk). Research of AS supported by EPSRC grant EP/V007327/1.

Lifshitz, and Minzer [KLM24] solved the structure problem. We refer the reader to [Ked09, TV17] for surveys of the area.

In this paper, we are interested in the infinite non-abelian setting. In particular, we will work with the *free semigroup* $\mathcal{F}$ and the *free group* $\mathcal{G}$[1] on a finite alphabet $\mathcal{A}$. To define density on $\mathcal{F}$ and $\mathcal{G}$, it is natural to use the measure that, for each n, assigns total weight 1 to the set of all (reduced) words of length n, and gives equal weight to all such words. For a set of words S, we write $\bar{d}(S)$ for its upper asymptotic density and $d^*(S)$ for its upper Banach density under this measure (see Section 2 for details). Since $\bar{d}(S) \leqslant d^*(S)$ for every set S, any result proved for upper Banach density also applies to upper asymptotic density.

The density problem for free semigroups on finite alphabets was solved by Leader, Letzter, Narayanan, and Walters [LLNW20].

Theorem 1.1 ([LLNW20])**.** *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subset \mathcal{F}$ is product-free, then $d^*(S) \leqslant 1/2$.*

There is a simple class of examples of large product-free subsets of $\mathcal{F}$ that show that the constant $1/2$ is best possible here. For a non-empty subset $\Gamma \subset \mathcal{A}$, the *odd-occurrence set* $\mathcal{O}_\Gamma \subset \mathcal{F}$ generated by Γ is the set of all words in which the total number of occurrences of letters from Γ is odd (note that if $\Gamma = \mathcal{A}$, then $\mathcal{O}_\Gamma$ consists of all words of odd length). It is easy to see that these sets are product-free with density $1/2$. Leader, Letzter, Narayanan, and Walters conjectured that these are the only examples.

Conjecture 1.2 ([LLNW20])**.** *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subset \mathcal{F}$ is product-free and $d^*(S) = 1/2$, then $S \subset \mathcal{O}_\Gamma$ for some nonempty subset $\Gamma \subset \mathcal{A}$.*

We confirm Conjecture 1.2 and in fact prove a more general result, Theorem 1.4, but this requires some setup. For $k \geqslant 3$, there are two natural ways to extend the notion of being product-free. Calkin and Erdős [CE96] and Łuczak and Schoen [ŁS97] defined a subset S of a (semi)group to be *k-product-free* ($k \geqslant 2$) if $x_1 \cdot \dots \cdot x_k \notin S$ for all $x_1, \dots, x_k \in S$ and to be *strongly k-product-free* if it is ℓ-product-free for every $\ell = 2, \dots, k$.

We first discuss *strongly* k-product-free sets. Ortega, Rué, and Serra extended Theorem 1.1 to all k, solving the density problem for strongly k-product-free sets in free semigroups and free groups.

Theorem 1.3 ([ORS24])**.** *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{H}$ be the free semigroup or the free group on alphabet $\mathcal{A}$. If $S \subset \mathcal{H}$ is strongly k-product-free, then $\bar{d}(S) \leqslant 1/k$.*

Our first main theorem solves the structure problem for strongly k-product-free sets in free semigroups. Note first that the odd-occurrence set $\mathcal{O}_\Gamma$ can also be described as follows: label each letter in Γ with a 1 and every other letter with a 0 and let the sum of a word be the sum of the labels of its letter; then $\mathcal{O}_\Gamma$ is the set of words with odd sum. The natural generalisation of this to $k \geqslant 3$ provides strongly k-product-free subsets of $\mathcal{F}$ with density $1/k$. We prove that these are the only examples.

Theorem 1.4. *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subset \mathcal{F}$ is strongly k-product-free and $d^*(S) = 1/k$, then the following holds. It is possible*

[1] The free semigroup on alphabet $\mathcal{A}$ is the set of all finite words with letters from $\mathcal{A}$ equipped with the operation of concatenation. The free group on alphabet $\mathcal{A}$ is the group generated by $\mathcal{A}$ whose only relations are that aa^{-1} and $a^{-1}a$ are the empty word for every $a \in \mathcal{A}$.

to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/k\mathbb{Z}$ such that S is a subset of the strongly k-product-free set

$$T := \{w \in \mathcal{F} \colon \text{the sum of the labels of letters in } w \text{ is } 1 \bmod k\}.$$

The set T defined in Theorem 1.4 is strongly k-product-free and has density $1/k$ provided the labelling is non-degenerate (otherwise T is empty; see Remark 3.1 for details). The special case of Theorem 1.4 with $k = 2$ confirms Conjecture 1.2.

We now turn to k-product-free sets in free semigroups. Since this notion is much weaker than being strongly k-product-free, the maximum density of k-product-free sets should be larger than $1/k$. In the special case $|\mathcal{A}| = 1$, the free semigroup $\mathcal{F}$ on alphabet $\mathcal{A}$ is isomorphic to the non-negative integers under addition. In this case, the term 'sum-free' is used in place of 'product-free'. Calkin and Erdős [CE96] conjectured that a k-sum-free subset of the non-negative integers has density at most $1/\rho(k)$ where $\rho(k)$ is

$$\rho(k) := \min\{\ell \in \mathbb{Z}^+ \colon \ell \nmid k - 1\}.$$

Note that the integers which are 1 mod $\rho(k)$ form a k-product-free set since the sum of k such integers is congruent to k mod $\rho(k)$ and $\rho(k)$ was chosen so that $k \not\equiv 1 \bmod \rho(k)$. In particular, $1/\rho(k)$ would be best possible. Łuczak and Schoen [ŁS97] confirmed this conjecture and also solved the structure problem for the non-negative integers. We extend their results by solving the density problem (for all k) and the structure problem (for all but four values of k) for k-product-free subsets of the free semigroup.

For the density problem we prove the following.

Theorem 1.5. *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subset \mathcal{F}$ is k-product-free, then $d^*(S) \leqslant 1/\rho(k)$.*

For the structure problem, we show that the structure of the extremal k-product-free sets is very similar to that of strongly k-product-free sets except everything is modulo $\rho(k)$ (see Section 6 for further discussion of the cases when k is 3, 5, 7, or 13).

Theorem 1.6. *Let $k \geqslant 2$ be an integer with $k \notin \{3, 5, 7, 13\}$ and $\rho = \rho(k)$. Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subset \mathcal{F}$ is k-product-free and $d^*(S) = 1/\rho$, then the following holds. It is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/\rho\mathbb{Z}$ such that S is a subset of the k-product-free set*

$$T := \{w \in \mathcal{F} \colon \text{the sum of the labels of letters in } w \text{ is } 1 \bmod \rho\}.$$

Finally, we consider k-product-free sets in the free group. Theorem 1.3 solves the density problem for strongly k-product-free sets. Ortega, Rué, and Serra [ORS24] made a conjecture corresponding to Calkin and Erdős's for k-product-free sets.

Conjecture 1.7 ([ORS24])**.** *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subset \mathcal{G}$ is k-product-free, then $\bar{d}(S) \leqslant 1/\rho(k)$.*

We prove a strengthening of this conjecture, with upper Banach density in place of upper asymptotic density. See Section 6 for a discussion and conjecture on what the structure of the extremal sets might be.

Theorem 1.8. *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subset \mathcal{G}$ is k-product-free, then $d^*(S) \leqslant 1/\rho(k)$.*

The rest of the paper is structured as follows. In Section 2 we provide the formal definitions of density. Sections 2.1 and 2.2 introduce the notions of upper asymptotic and upper Banach density and prove some important technical lemmas. In Section 2.3, we then state our main density result, Theorem 2.6, from which Theorem 1.5 follows and prove the key technical results for our structural proofs, Lemmas 2.9 and 2.10. We also deduce a result, Corollary 2.7, which shows that the density bounds of Theorems 1.3 and 1.5 hold locally down all subtrees. Before proving Theorem 2.6 we obtain our structural results in Section 3 whose proofs are simpler and already contain some of the key ideas. The proof of Theorem 1.4 is given in Section 3.1 and the proof of Theorem 1.6 in Section 3.2. Section 4 is devoted to our main density result. In Section 4.1 we build the machinery that we use to prove Theorem 2.6 in Section 4.2. In Section 5 we adapt our arguments to the free group. We finish, in Section 6, with some open problems.

2 Density

Throughout this paper $\mathcal{F}$ will be the free semigroup on a finite alphabet $\mathcal{A}$. That is, $\mathcal{F}$ is the set of all finite words whose letters are in $\mathcal{A}$ equipped with the associative operation of concatenation and whose identity is the empty word.

In this section, we will formally define density for subsets of $\mathcal{F}$. We start in Section 2.1 by introducing the upper asymptotic and upper Banach density. While we want to prove upper bounds for these densities, they are difficult to use directly in our proofs for various technical reasons including that they are not additive.

Instead, in Section 2.2, we will define an adapted notion of density that is closely related but much easier to handle. We use this directly to prove some basic properties about the density and to define a notion of density for subtrees of $\mathcal{F}$.

With these notions of density in hand, we state our main density result, Theorem 2.6, in Section 2.3. While we defer the proof of this result to Section 4 due to its complexity, we will immediately deduce some properties about the density of a produce-free set in subtrees of $\mathcal{F}$ which will be crucial for proving our structural results in Section 3.

2.1 Upper asymptotic and upper Banach density

To motivate and provide intuition for the notation we view things from the perspective of a randomly generated word. Let $\mathbf{W} = \alpha_1\alpha_2\cdots$ be a random infinite word where each α_i is an independent uniformly random letter in $\mathcal{A}$. Taking $\mathbf{W}_n = \alpha_1\alpha_2\cdots\alpha_n$, we may view $(\mathbf{W}_n)$ as a random walk on the infinite $|\mathcal{A}|$-ary tree. We say $\mathbf{W}$ *hits* a set $B \subset \mathcal{F}$ if the random walks hits B (equivalently if $\mathbf{W}$ has a prefix in B) and $\mathbf{W}$ *avoids* B otherwise. We equip $\mathcal{F}$ with a measure μ satisfying, for every word $w \in \mathcal{F}$,

$$\mu(w) = \mathbb{P}(\mathbf{W} \text{ hits } w) = |\mathcal{A}|^{-|w|}.$$

Note that, for $B \subset \mathcal{F}$, $\mu(B) = \sum_{w\in B}\mu(w)$ is the expected number of times that $\mathbf{W}$ hits B. This has a useful corollary. A set $C \subset \mathcal{F}$ is *prefix-free* if there are no distinct words $a, b \in C$ where a is a prefix of b. $\mathbf{W}$ can hit a prefix-free set at most once.

Observation 2.1. *If $C \subset \mathcal{F}$ is prefix-free, then $\mu(C) \leqslant 1$.*

For a positive integer n and a set $B \subset \mathcal{F}$ the *length n layer of B* is

$$B(n) := \{w \in B \colon |w| = n\},$$

while, for an interval $I \subset \mathbb{Z}^+$,

$$B(I) := \{w \in B \colon |w| \in I\}.$$

Note that the measure μ is defined so that $\mu(\mathcal{F}(n)) = 1$. The density of B on layer n is $|B(n)|/|\mathcal{F}(n)| = \mu(B(n))$, which is the probability that $\mathbf{W}_n$ is in B. The *density of B on interval I* is

$$d^I(B) := \frac{\mu(B(I))}{\mu(\mathcal{F}(I))} = |I|^{-1} \sum_{n \in I} \mu(B(n)).$$

With these definitions in place, we may give standard notions of density. The *upper asymptotic density* of B is

$$\bar{d}(B) := \limsup_{m \to \infty} d^{\{1,2,\ldots,m\}}(B) = \limsup_{m \to \infty} \sum_{n=1}^{m} \mu(B(n))/m.$$

The *upper Banach density* of B is

$$d^*(B) := \limsup_{I \to \infty} d^I(B) = \limsup_{I \to \infty} |I|^{-1} \sum_{n \in I} \mu(B(n)),$$

where I is an interval and the notation $I \to \infty$ denotes that both $|I|$ and $\min I$ tend to infinity[2]. In the following we will exclusively use upper Banach density, but since $\bar{d}(B) \leqslant d^*(B)$ for any set B, our results apply equally well to upper asymptotic density. We remark that these densities always lie in the interval $[0, 1]$.

It should be noted that limit superiors are only subadditive (and not additive). In particular, for disjoint sets $A, B \subset \mathcal{F}$ we have $d^*(A \cup B) \leqslant d^*(A) + d^*(B)$ and equality may not hold. For example, the sets

$$\begin{aligned} A &= \bigcup_{n \in \mathbb{Z}^+} \mathcal{F}(\{(2n-1)! + 1, (2n-1)! + 2, \ldots, (2n)!\}), \\ B &= \bigcup_{n \in \mathbb{Z}^+} \mathcal{F}(\{(2n)! + 1, (2n)! + 2, \ldots, (2n+1)!\}) \end{aligned}$$

are disjoint and both have density 1.

Despite this, in the semigroup of non-negative integers $\mathcal{F} = \mathbb{Z}^+$, $d^*(B)$ satisfies some useful properties. For example, it holds that $|d^I(x+B) - d^I(B)| \leqslant x/|I|$. This implies that $d^*(x+B) = d^*(B)$. Even more importantly, if $x_1, \ldots, x_n \in \mathbb{Z}^+$ are such that $x_1 + B, \ldots, x_n + B$ are disjoint, then $d^I(x_1 + B) + \cdots + d^I(x_n + B) \leqslant 1$, implying that

$$n \cdot d^I(B) \leqslant \sum_{i=1}^{n} \left(d^I(x_i + B) + \frac{x_i}{|I|} \right) \leqslant 1 + \sum_{i=1}^{n} \frac{x_i}{|I|}$$

and so $n \cdot d^*(B) \leqslant 1$. Not only can this provide upper bounds on the density of B, but if we knew that $d^*(B) > 1/n$, we could conclude that the sets $x_1 + B, \ldots, x_n + B$ cannot all

[2] The condition $\min I \to \infty$ is often omitted from the definition. However, some simple analysis shows that, whether or not this condition is included, the resulting density is the same.

be disjoint and thereby deduce some structural information about B. Such arguments were used by Łuczak and Schoen [Łuc95, ŁS97] for their results about sum-free subsets of the non-negative integers and also of right cancellative semigroups [ŁS01].

In the free semigroup $\mathcal{F}$ where $|\mathcal{A}| > 1$, these arguments no longer work. For example, if $w \in \mathcal{F}$, it is easy to see that $d^*(wB) = |\mathcal{A}|^{-|w|} \cdot d^*(B)$ where $wB := \{wb \colon b \in B\}$. Also, the fact that $w_1 B, \dots, w_n B$ are disjoint gives no general upper bound on the density of B. Even if we consider nested sets $B, wB, \dots, w^n B$, taking $B := \mathcal{F} \setminus (w\mathcal{F})$ provides an example where these sets are pairwise disjoint, but $d^*(B) = 1 - |\mathcal{A}|^{-|w|}$ which can be arbitrarily close to 1.

We address these issues in Sections 2.2 and 2.3. By modifying the density that we consider, we can ensure that the density is additive. Importantly, the density of the set $S \subset \mathcal{F}$ whose upper Banach density we want to bound will not change. Moreover, in certain situations, we prove that n disjoint nested copies of B imply that the density of B is at most $1/n$. This will be crucial for proving our structural results.

2.2 Diagonalisation and relative density

Let $S \subset \mathcal{F}$ be a fixed set whose upper Banach density we wish to bound (for example, S might be k-product-free). There is a sequence of intervals (I_j) such that $I_j \to \infty$ and

$$d^{I_j}(S) \to d^*(S), \quad \text{as } j \to \infty.$$

Let $B \subset \mathcal{F}$ be another set. The sequence $(d^{I_j}(B))$ is bounded (all terms are in $[0,1]$) and so, by the Bolzano-Weierstrass theorem, has a convergent subsequence. In particular, by passing to a subsequence of (I_j), we may assume that $d^{I_j}(S) \to d^*(S)$ and $(d^{I_j}(B))$ converges to some limit that we will call $d^{I_\infty}(B)$. Given a countable collection of subsets of $\mathcal{F}$, we may, by a diagonalisation argument, assume there is a subsequence (I_j) such that $d^{I_j}(B) \to d^{I_\infty}(B)$ for every B in the collection where $d^{I_\infty}(S) = d^*(S)$. Throughout this paper we will only ever consider countably many sets and so we may assume that this convergence occurs for all sets we consider. These limits, unlike the corresponding upper Banach densities, are additive. Indeed, if sets A and B are disjoint, then $d^{I_j}(A \cup B) = d^{I_j}(A) + d^{I_j}(B)$ and so $d^{I_\infty}(A \cup B) = d^{I_\infty}(A) + d^{I_\infty}(B)$. It should be noted that while $d^{I_\infty}(S) = d^*(S)$, we only have $d^{I_\infty}(B) \leqslant d^*(B)$ for the other sets.

For our proofs we will need not only to bound the density of a product-free set S but also to bound the density of S on subtrees. We now begin to define this.

The product AB of two sets $A, B \subset \mathcal{F}$ is

$$AB := \{ab \colon a \in A, b \in B\}$$

and the set B^k is the product of k copies of B. Note that B is k-product-free exactly if $B \cap B^k = \varnothing$. A particularly important example of a product is $w\mathcal{F}$ for a word $w \in \mathcal{F}$: this is exactly the subtree of $\mathcal{F}$ consisting of all words starting with w. Similarly $B\mathcal{F}$ is exactly the set of words that have a prefix in B.

For a finite set $B \subset \mathcal{F}$ we write $\min B$ and $\max B$ for the length of the shortest and longest words in B, respectively. Note that if B is finite, then for $n \geqslant \max B$ the random infinite word $\mathbf{W}$ hits $(B\mathcal{F})(n)$ if and only if it hits B.

Observation 2.2. *If $n \geqslant |w|$, then $\mu((w\mathcal{F})(n)) = \mu(w)$. If $C \subset \mathcal{F}$ is prefix-free and finite, then $\mu((C\mathcal{F})(n)) = \mu(C)$ for $n \geqslant \max C$.*

Definition 2.3 (relative density)**.** Let $w \in \mathcal{F}$ and $B \subset \mathcal{F}$. For $n \geqslant |w|$, the relative density of B in $w\mathcal{F}$ on layer n is

$$\frac{|B(n) \cap w\mathcal{F}|}{|\mathcal{F}(n) \cap w\mathcal{F}|} = \frac{\mu(B(n) \cap w\mathcal{F})}{\mu(\mathcal{F}(n) \cap w\mathcal{F})} = \frac{\mu(B(n) \cap w\mathcal{F})}{\mu(w)}$$

which is the probability that $\mathbf{W}_n$ is in B conditioned on the event that $\mathbf{W}$ hits w. If $n < |w|$, then we will take the relative density to be 0 by convention.

Furthermore, if I is an interval with $\min I \geqslant |w|$, then the *relative density of B in $w\mathcal{F}$ on interval I* is

$$d^I_{w\mathcal{F}}(B) := \frac{\mu(B(I) \cap w\mathcal{F})}{\mu(\mathcal{F}(I) \cap w\mathcal{F})} = |I|^{-1}\mu(w)^{-1} \sum_{n \in I} \mu(B(n) \cap w\mathcal{F}) = \mu(w)^{-1} \cdot d^I(B \cap w\mathcal{F}).$$

If $\min I < |w|$, then we will take the relative density to be 0 by convention.

Note that if w is the empty word then this relative density is just $d^I(B)$.

Consider the sequence of intervals (I_j) given above where $d^{I_j}(B) \to d^{I_\infty}(B)$ for every set B in a countable collection. For each word $w \in \mathcal{F}$ and each set in the collection, the sequence $(d^{I_j}_{w\mathcal{F}}(B))$ is bounded (all terms are in $[0,1]$) and so, by the Bolzano-Weierstrass theorem, has a convergent subsequence. Since $\mathcal{F}$ is countable (it consists of only finite words) we may, via a diagonalisation argument, pass to a subsequence (I_j) such that, for every $w \in \mathcal{F}$ and every B in the countable collection, $(d^{I_j}_{w\mathcal{F}}(B))$ converges to some limit $d^{I_\infty}_{w\mathcal{F}}(B)$. In conclusion, we may assume throughout the paper that for any set B we encounter and for all $w \in \mathcal{F}$ we have

$$d^{I_j}_{w\mathcal{F}}(B) \to d^{I_\infty}_{w\mathcal{F}}(B),$$

where $d^{I_\infty}(B) \leqslant d^*(B)$ and $d^{I_\infty}(S) = d^*(S)$ for one fixed set S. Note that when w is the empty word, $d^{I_\infty}_{w\mathcal{F}}$ is the same as d^{I_∞}. As before, these limits are additive. They satisfy the useful property that we may strip away prefixes.

Lemma 2.4. *If $w, v \in \mathcal{F}$, then $d^{I_\infty}_{wv\mathcal{F}}(wB) = d^{I_\infty}_{v\mathcal{F}}(B)$.*

Proof. Let I be any interval with $\min I > |wv|$. Now

$$d^I_{wv\mathcal{F}}(wB) = |I|^{-1}\mu(wv)^{-1} \sum_{n \in I} \mu((wB)(n) \cap wv\mathcal{F}).$$

Removing the leading w from each word in $(wB)(n) \cap wv\mathcal{F}$ shows that $\mu((wB)(n) \cap wv\mathcal{F}) = \mu(w) \cdot \mu(B(n - |w|) \cap v\mathcal{F})$. Also $\mu(wv) = \mu(w)\mu(v)$ and so

$$d^I_{wv\mathcal{F}}(wB) = |I|^{-1}\mu(v)^{-1} \sum_{n \in I - |w|} \mu(B(n) \cap v\mathcal{F}),$$

where $I - |w|$ is the interval obtained by subtracting $|w|$ from each element of I. Thus

$$|d^I_{wv\mathcal{F}}(wB) - d^I_{v\mathcal{F}}(B)| = |I|^{-1}\mu(v)^{-1} \cdot \left| \sum_{n \in I - |w|} \mu(B(n) \cap v\mathcal{F}) - \sum_{n \in I} \mu(B(n) \cap v\mathcal{F}) \right|$$

But, for each integer n, $\mu(B(n) \cap v\mathcal{F}) \in [0,1]$ and so

$$|d^I_{wv\mathcal{F}}(wB) - d^I_{v\mathcal{F}}(B)| \leqslant |I|^{-1}\mu(v)^{-1} \cdot |w|$$

Setting $I = I_j$ and taking j to infinity gives the required result. □

We are now ready to make an important definition that captures the densest that a set B can be down a subtree.

Definition 2.5 (sup density)**.** For a set B in the countable collection, the *sup density of B* is

$$d^{I_\infty}_{\sup}(B) := \sup_{w \in \mathcal{F}} d^{I_\infty}_{w\mathcal{F}}(B).$$

Of course, the sup density satisfies $d^{I_\infty}_{\sup}(B) \geqslant d^{I_\infty}(B)$ (note that the empty word is in $\mathcal{F}$) and so $d^{I_\infty}_{\sup}(S) \geqslant d^*(S)$.

2.3 Key technical density results

We will prove the following strengthening of Theorems 1.3 and 1.5 in Section 4. Note that while it would not be strictly necessary to state this result for an arbitrary sequence of intervals (I_j) (we could just use the sequence of intervals described in Section 2.2 for the proofs of all our main theorems), the extra flexibility allows us to prove Corollary 2.7: the bounds $1/k$ and $1/\rho(k)$ for the density of (strongly) k-product-free sets hold not only globally but also locally down each subtree $w\mathcal{F}$.

Theorem 2.6. *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. Let (I_j) be a sequence of intervals, with $I_j \to \infty$, on which all relevant densities converge.*

(a) *If $S \subset \mathcal{F}$ is strongly k-product-free, then $d^{I_\infty}_{\sup}(S) \leqslant 1/k$.*

(b) *If $S \subset \mathcal{F}$ is k-product-free, then $d^{I_\infty}_{\sup}(S) \leqslant 1/\rho(k)$.*

Note that Theorems 1.3 and 1.5 follow immediately from this: taking (I_j) as the sequence described in Section 2.2 we have $d^*(S) = d^{I_\infty}(S) \leqslant d^{I_\infty}_{\sup}(S)$ and so Theorem 2.6 gives the required upper bounds.

We can also directly prove that the bounds for the upper Banach density hold not only globally but also locally down all subtrees. We will not use this result elsewhere.

Corollary 2.7. *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. Let w be any word of $\mathcal{F}$.*

(a) *If $S \subset \mathcal{F}$ is strongly k-product-free, then $\limsup_{I\to\infty} d^I_{w\mathcal{F}}(S) \leqslant 1/k$.*

(b) *If $S \subset \mathcal{F}$ is k-product-free, then $\limsup_{I\to\infty} d^I_{w\mathcal{F}}(S) \leqslant 1/\rho(k)$.*

Proof. Let $S \subset \mathcal{F}$ be (strongly) k-product-free. By definition of the limit superior, there is a sequence of intervals (I_j) such that $I_j \to \infty$ and

$$d^{I_j}_{w\mathcal{F}}(S) \to \limsup_{I\to\infty} d^I_{w\mathcal{F}}(S), \quad \text{as } j \to \infty.$$

Start with this sequence of intervals and apply the diagonalisation argument outlined in Section 2.2, repeatedly passing to subsequences on which all the densities we need to consider converge. We chose the sequence (I_j) so that $d^{I_\infty}_{w\mathcal{F}}(S) = \limsup_{I\to\infty} d^I_{w\mathcal{F}}(S)$. Therefore,

$$\limsup_{I\to\infty} d^I_{w\mathcal{F}}(S) = d^{I_\infty}_{w\mathcal{F}}(S) \leqslant d^{I_\infty}_{\sup}(S).$$

Theorem 2.6 then gives the required upper bounds. □

Theorem 2.6 is also needed for our structural results, Theorems 1.4 and 1.6. For example, if $S \subset \mathcal{F}$ is strongly k-product-free with $d^*(S) = 1/k$, then applying (a) gives $1/k = d^*(S) = d^{I_\infty}(S) \leqslant d^{I_\infty}_{\sup}(S) \leqslant 1/k$ and so $d^{I_\infty}(S) = d^{I_\infty}_{\sup}(S)$. This suggests that S is uniformly distributed down subtrees which is made precise by the following lemma.

Lemma 2.8. *If $d^{I_\infty}(B) = d^{I_\infty}_{\sup}(B)$, then $d^{I_\infty}_{w\mathcal{F}}(B) = d^{I_\infty}(B)$ for every word $w \in \mathcal{F}$.*

Proof. Let ℓ be a non-negative integer and let I be an interval with $\min I > \ell$. Every word of length greater than ℓ is in exactly one $w\mathcal{F}$ (where $w \in \mathcal{F}(\ell)$). Hence,

$$d^I(B) = \sum_{w\in\mathcal{F}(\ell)} d^I(B \cap w\mathcal{F}) = \sum_{w\in\mathcal{F}(\ell)} \mu(w) \cdot d^I_{w\mathcal{F}}(B).$$

Setting $I = I_j$ and taking j to infinity gives

$$d^{I_\infty}(B) = \sum_{w\in\mathcal{F}(\ell)} \mu(w) \cdot d^{I_\infty}_{w\mathcal{F}}(B).$$

Now $\sum_{w\in\mathcal{F}(\ell)} \mu(w) = \mu(\mathcal{F}(\ell)) = 1$ and every $w \in \mathcal{F}(\ell)$ satisfies $d^{I_\infty}_{w\mathcal{F}}(B) \leqslant d^{I_\infty}_{\sup}(B) = d^{I_\infty}(B)$. Hence we must have $d^{I_\infty}_{w\mathcal{F}}(B) = d^{I_\infty}(B)$ for every $w \in \mathcal{F}(\ell)$. The integer ℓ was arbitrary and so we have the required result. □

The next two lemmas are the key technical results for our structural proofs. We remark that for the non-negative integers (that is, when $|\mathcal{A}| = 1$) they are much more obvious.

Lemma 2.9. *Let $B \subset \mathcal{F}$ be such that $d^{I_\infty}(B) = d^{I_\infty}_{\sup}(B) > 1/n$. Then, for any $w_1, \dots, w_n \in \mathcal{F}$, the sets*

$$w_1B, \quad w_1w_2B, \quad \dots, \quad w_1w_2\cdots w_{n-1}B, \quad w_1w_2\cdots w_nB$$

cannot be pairwise disjoint.

Proof. Assume that these sets are pairwise disjoint. Then, for any word $w \in \mathcal{F}$,

$$d^{I_\infty}_{w\mathcal{F}}(w_1B) + \cdots + d^{I_\infty}_{w\mathcal{F}}(w_1\cdots w_nB) = d^{I_\infty}_{w\mathcal{F}}((w_1B) \cup \cdots \cup (w_1\cdots w_nB)) \leqslant 1.$$

Choose $w = w_1\cdots w_n$. Applying Lemma 2.4 to each term gives

$$d^{I_\infty}_{w_2\cdots w_n\mathcal{F}}(B) + \cdots + d^{I_\infty}_{w_n\mathcal{F}}(B) + d^{I_\infty}_{\mathcal{F}}(B) \leqslant 1.$$

By Lemma 2.8, each term is $d^{I_\infty}(B)$ which contradicts $d^{I_\infty}(B) > 1/n$, as required. □

Lemma 2.10. *Let $B \subset \mathcal{F}$ be such that $d^{I_\infty}(B) = d^{I_\infty}_{\sup}(B) > 2/(2n-1)$. Then, for any $w_1, \dots, w_n, v_1, \dots, v_n \in \mathcal{F}$ and $C \subset B$, either the sets*

$$w_1B, \quad w_1w_2B, \quad \dots, \quad w_1\cdots w_{n-1}B, \quad w_1\cdots w_nC$$

or the sets

$$v_1B, \quad v_1v_2B, \quad \ldots, \quad v_1\cdots v_{n-1}B, \quad v_1\cdots v_n(B\setminus C)$$

are not pairwise disjoint.

Proof. Assume that both collections of sets are pairwise disjoint. Then, as in the proof of Lemma 2.9,

$$d^{I_\infty}_{w_2\cdots w_n\mathcal{F}}(B) + \cdots + d^{I_\infty}_{w_n\mathcal{F}}(B) + d^{I_\infty}_{\mathcal{F}}(C) \leqslant 1$$

and

$$d^{I_\infty}_{v_2\cdots v_n\mathcal{F}}(B) + \cdots + d^{I_\infty}_{v_n\mathcal{F}}(B) + d^{I_\infty}_{\mathcal{F}}(B\setminus C) \leqslant 1.$$

Note that $d^{I_\infty}_{\mathcal{F}}(B\setminus C) = d^{I_\infty}_{\mathcal{F}}(B) - d^{I_\infty}_{\mathcal{F}}(C)$. Applying this and adding the two inequalities, we get

$$d^{I_\infty}_{w_2\cdots w_n\mathcal{F}}(B) + \cdots + d^{I_\infty}_{w_n\mathcal{F}}(B) + d^{I_\infty}_{v_2\cdots v_n\mathcal{F}}(B) + \cdots + d^{I_\infty}_{\mathcal{F}}(B) \leqslant 2.$$

However, by Lemma 2.8, each term is $d^{I_\infty}(B)$ which contradicts $d^{I_\infty}(B) > 2/(2n-1)$, as required. □

3 Structure of product-free sets in the free semigroup

After formally defining density, we now turn to the structure of product-free sets. In this section, we determine the structure of (strongly) k-product-free sets in free semigroups and prove Theorems 1.4 and 1.6 assuming Theorem 2.6.

We start in Section 3.1 by considering strongly k-product-free sets $S \subset \mathcal{F}$ with density $d^*(S) = 1/k$. Recall that in Theorem 1.4 we want to show that we can label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/k\mathbb{Z}$ such that S is a subset of

$$T := \{a \in \mathcal{F} \colon \text{the sum of the labels of letters in } a \text{ is } 1 \bmod k\}.$$

We say that the label of word $a \in \mathcal{F}$ is the sum of the labels of letters in a. Our strategy will be to show that a is labelled with $\ell \in \mathbb{Z}/k\mathbb{Z}$ if and only if $S \cap aS^{k-\ell+1} \neq \varnothing$. To prove this, we will repeatedly construct nested copies of S and apply Lemmas 2.9 and 2.10. For example, given any $x \in S$, we can consider the sets

$$S, \quad aS, \quad axS, \quad ax^2S, \quad \ldots, \quad ax^{k-1}S.$$

By Lemma 2.9 these cannot all be disjoint, but since S is strongly k-product-free we know that the sets $aS, axS, \ldots, ax^{k-1}S$ are disjoint. Therefore, $S \cap ax^rS \neq \varnothing$ for some $r \in \mathbb{Z}/k\mathbb{Z}$ and so $S \cap aS^{r+1} \neq \varnothing$. This way we can identify the label of a. More complicated versions of such arguments eventually allow us to show that $S \subset T$.

To prove our structural result for k-product-free sets $S \subset \mathcal{F}$ with $k \notin \{3, 5, 7, 13\}$, we will show in Section 3.2 that if $d^*(S) = 1/\rho(k)$, then S is strongly $\rho(k)$-product-free. This uses arguments very similar to the strongly k-product-free case. Then, we can simply apply our structural result for strongly $\rho(k)$-product-free sets to obtain Theorem 1.6.

3.1 Structure of strongly k-product-free sets

Let $S \subset \mathcal{F}$ be a strongly k-product-free set with $d^*(S) = 1/k$. Note, by Theorem 2.6, that $d^{I_\infty}(S) = 1/k = d^{I_\infty}_{\sup}(S)$ and so we may and will frequently apply Lemmas 2.9 and 2.10 with $n = k + 1$.

To prove Theorem 1.4, we need to show that there exists a labelling of the letters of $\mathcal{A}$ such that S is a subset of the strongly k-product-free set

$$T := \{a \in \mathcal{F} \colon \text{the label of } a \text{ is } 1 \bmod k\}.$$

Remark 3.1. If some prime divides k and every label given to letters in $\mathcal{A}$, then T will be empty. If there is no such prime, then T will be non-empty by Bezout's lemma. If T is non-empty, then $d^*(T) = 1/k$. Indeed, let $\alpha_1\alpha_2\cdots$ be an infinite random word where the α_i are independent uniformly random letters from $\mathcal{A}$ and let X_n be the sum of the labels of $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then (X_n) is a Markov chain on $\mathbb{Z}/k\mathbb{Z}$ that is irreducible (since $T \neq \varnothing$). The uniform distribution π on $\mathbb{Z}/k\mathbb{Z}$ is stationary for this chain. Let d be the period of (X_n): by the Markov convergence theorem, for each fixed $r \in \{1, \ldots, k-1\}$, the subsequence (X_{nd+r}) converges to π in distribution, and so the averages $|I|^{-1}\sum_{n\in I} X_n$ over long intervals converge to π in distribution. In particular, $d^*(T) = \pi(1) = 1/k$.

To deduce the structure of S, we need to identify the labels for all words $a \in \mathcal{F}$. Clearly, everything in S should be labelled 1 mod k. For any other $a \in \mathcal{F}$, appending a word from S should increase the label by 1. So, if a has label ℓ and we append $i = -\ell \in \mathbb{Z}/k\mathbb{Z}$ words from S to a, we should get the label 0, and appending one more word from S should give the label 1, which might itself be a word from S. On the other hand, for any other $j \in \mathbb{Z}/k\mathbb{Z} \setminus \{i\}$, appending $j + 1$ words from S to a should give a label different from 1 and should therefore never yield a word from S.

Based on this intuition, for $i = 0, 1, \ldots, k-1$ define

$$T_i := \{a \in \mathcal{F} \colon S \cap aS^{i+1} \neq \varnothing\}.$$

Then, everything in T_i should have the label $-i \in \mathbb{Z}/k\mathbb{Z}$. So, we expect that $S \subset T_{k-1}$ and that $T_iT_j \subset T_{i+j}$. This is exactly what we will show and which allows us to deduce the structure of T_{k-1}, which will be the set T from above.

Remark 3.2. Throughout we will view the indices of the T_i as elements of $\mathbb{Z}/k\mathbb{Z}$ and, in particular, all addition of indices is modulo k.

Note that our definition of T_i is slightly arbitrary. Whether we append or prepend words from S to some $a \in \mathcal{F}$, the change in the label of a should always be the same. So, we could also have defined T_i as the set $\{a \in \mathcal{F} \colon S \cap S^{i+1}a \neq \varnothing\}$. Fortunately, the following result tells us that these definitions are equivalent.

Proposition 3.3. *For any positive integer r and any $a \in \mathcal{F}$,*

$$S \cap S^r a \neq \varnothing \Leftrightarrow S \cap S^{r-1}aS \neq \varnothing \Leftrightarrow \cdots \Leftrightarrow S \cap SaS^{r-1} \neq \varnothing \Leftrightarrow S \cap aS^r \neq \varnothing.$$

Proof. We first prove the case $r = 1$. Suppose that $S \cap Sa \neq \varnothing$. Then there is some x such that $x, xa \in S$. Consider the sets $S, xS, x^2S, \ldots, x^{k-1}S, x^{k-1}aS = x^{k-2}(xa)S$. By Lemma 2.9, these cannot all be pairwise disjoint. Since S is strongly k-product-free and

$x \in S$, the sets $S, xS, \ldots, x^{k-1}S$ are pairwise disjoint. Since S is strongly k-product-free and $xa \in S$, the sets $S, xS, \ldots, x^{k-2}S, x^{k-2}(xa)S$ are pairwise disjoint. Thus $x^{k-1}S$ and $x^{k-1}aS$ are not disjoint and so $S \cap aS \neq \varnothing$.

Let $f\colon \mathcal{F} \to \mathcal{F}$ be the *reverse map* that reverses each word of $\mathcal{F}$ (that is, reads them from right to left). The function f is a measure-preserving involution. Let $\overline{S} = f(S)$. Now $\overline{S}$ is a strongly k-product-free subset of $\mathcal{F}$ with $d^*(\overline{S}) = d^*(S) = 1/k$. In particular, the previous paragraph shows that $\overline{S} \cap \overline{S}a \neq \varnothing \Rightarrow \overline{S} \cap a\overline{S} \neq \varnothing$. Now, $\overline{S} \cap \overline{S}a = f(S \cap aS)$ and $\overline{S} \cap a\overline{S} = f(S \cap Sa)$ and so $S \cap aS \neq \varnothing \Rightarrow S \cap Sa \neq \varnothing$ concluding the case $r = 1$.

For the general case it suffices to prove that for all non-negative integers i, j: $S \cap S^{i+1}aS^j \neq \varnothing \Leftrightarrow S \cap S^i a S^{j+1} \neq \varnothing$. Suppose that $S \cap S^{i+1}aS^j \neq \varnothing$. Then there is $x_i \in S^i$ and $x_j \in S^j$ such that $S \cap Sx_i a x_j \neq \varnothing$. Applying the $r = 1$ case to the word $x_i a x_j$ shows that $S \cap x_i a x_j S \neq \varnothing$ and so $S \cap S^i a S^{j+1} \neq \varnothing$. The other direction is analogous. □

If the sets $T_0, T_1, \ldots, T_{k-1}$ are supposed to correctly identify the labels of all words $a \in \mathcal{F}$, then every a should be in exactly one of these sets, and S should satisfy $S \subset T_{k-1}$. This is proved by the following proposition.

Proposition 3.4. *The sets $T_0, T_1, \ldots, T_{k-1}$ partition $\mathcal{F}$ and $S \subset T_{k-1}$.*

Proof. Let $a \in \mathcal{F}$ and $x \in S$. Consider the sets S, aS, axS, ax^2S, $\ldots$, $ax^{k-1}S$. By Lemma 2.9, these cannot all be pairwise disjoint. Since S is strongly k-product-free and $x \in S$, the sets aS, axS, $\ldots$, $ax^{k-1}S$ are pairwise disjoint. Hence there is some $r \in \{0, 1, \ldots, k-1\}$ such that $S \cap ax^rS \neq \varnothing$ and so $S \cap aS^{r+1} \neq \varnothing$. That is, $\bigcup_{r=0}^{k-1} T_r = \mathcal{F}$.

We next show that the T_i are pairwise disjoint (and so partition $\mathcal{F}$). Suppose that $a \in T_i \cap T_j$ where $0 \leqslant i < j \leqslant k-1$. Since $a \in T_i$, Proposition 3.3 implies $S \cap SaS^i \neq \varnothing$ and so there is $x \in S$ and $y \in S^i$ such that $xay \in S$. Let

$$C := \{s \in S\colon ays \in S\} \subset S.$$

Consider the $k+1$ sets

$$S, \quad xS, \quad x^2S, \quad \ldots, \quad x^{k-1}S, \quad x^{k-1}ay(S \setminus C).$$

As S is strongly k-product-free and $x \in S$, the first k of these sets are pairwise disjoint. Similarly, noting that $x^{k-1}ay = x^{k-2}(xay)$ and $xay \in S$, we have that the last set is disjoint from each of the first $k-1$. Finally, the last two sets are disjoint by the definition of C. Hence, all $k+1$ sets are pairwise disjoint.

Since $a \in T_j$ there are $z_1, \ldots, z_{j+1} \in S$ such that $z_1 \cdots z_{j+1}a \in S$. Consider the $k+1$ sets

$$S, \quad z_1S, \quad z_1^2S, \quad \ldots, \quad z_1^{k-j}S, \quad z_1^{k-j}z_2S, \quad \ldots, \quad z_1^{k-j}z_2\cdots z_jS, \quad z_1^{k-j}z_2\cdots z_{j+1}ayC.$$

The first k of these sets are pairwise disjoint as S is strongly k-product-free. Similarly, noting that $z_1z_2\cdots z_{j+1}a \in S$, $y \in S^i$, and $i < j$, the last set is disjoint from each of S, $z_1S, \ldots, z_1^{k-j-1}S$. Now, by the definition of C, $ayC \subset S$. Using this and product-freeness shows that the last set is disjoint from each of $z_1^{k-j}S, z_1^{k-j}z_2S, \ldots, z_1^{k-j}z_2\cdots z_jS$. Hence, all $k+1$ sets are pairwise disjoint. This contradicts Lemma 2.10 and so the T_i do partition $\mathcal{F}$.

It remains to show that $S \subset T_{k-1}$. Since S is strongly k-product-free, for any $x \in S$, the set S is disjoint from each of $xS, xS^2, \ldots, xS^{k-1}$ and so $x \notin T_0 \cup \cdots \cup T_{k-2}$. Since the T_i partition $\mathcal{F}$, we must have $x \in T_{k-1}$, as required. □

Given these two results, we already know that $a \in T_i$ should be labelled by $-i \in \mathbb{Z}/k\mathbb{Z}$. Next, we want to show that the label of a product ab should be the sum of the labels of a and b. We begin by proving that this is true whenever we append a word from S.

Proposition 3.5. *The following hold for all $j \in \mathbb{Z}/k\mathbb{Z}$.*

(a) *If $ax \in T_j$ and $x \in S$, then $a \in T_{j+1}$.*
(b) *$T_{j+1}S \subset T_j$.*

Proof. We first prove (a). Suppose that $0 \leqslant j \leqslant k-2$. We have $S \cap axS^{j+1} \neq \varnothing$ and $x \in S$, so $S \cap aS^{j+2} \neq \varnothing$ and so $a \in T_{j+1}$.

Now suppose that $j = k-1$. Consider the sets $S, aS, axS, ax^2S, \ldots, ax^{k-1}S$. By Lemma 2.9, these cannot all be pairwise disjoint. Since S is strongly k-product-free and $x \in S$, the sets $aS, axS, \ldots, ax^{k-1}S$ are pairwise disjoint. Also, as $ax \in T_{k-1}$ (and so ax is not in $T_0 \cup T_1 \cup \cdots \cup T_{k-2}$ by Proposition 3.4), S is disjoint from each of $axS, ax^2S, \ldots, ax^{k-1}S$. Thus S and aS are not disjoint and so $a \in T_0$, as required.

We now prove (b). Let $a \in T_{j+1}$ and $x \in S$. Suppose that $ax \in T_i$ (such an i exists by Proposition 3.4). By (a), $i+1 = j+1 \bmod k$ and so $i = j \bmod k$, as required. □

It is now an easy consequence that the labels of all T_i are very well-behaved with respect to products.

Proposition 3.6. *For all $i, j \in \mathbb{Z}/k\mathbb{Z}$, $T_iT_j \subset T_{i+j}$.*

Proof. Let $a \in T_i$ and $b \in T_j$. As $b \in T_j$ there are $x_1, x_2, \ldots, x_{j+1} \in S$ such that $bx_1x_2 \cdots x_{j+1} \in S$. By Proposition 3.5(b),

$$abx_1 \cdots x_{j+1} = a(bx_1 \cdots x_{j+1}) \in T_{i-1}.$$

Applying Proposition 3.5(a) $j+1$ times, once to remove each x_ℓ, gives $ab \in T_{i-1+(j+1)} = T_{i+j}$, as required. □

Finally, this allows us to complete the proof of Theorem 1.4.

Proof of Theorem 1.4. For each letter $\alpha \in \mathcal{A}$, there is, by Proposition 3.4, a unique $i \in \mathbb{Z}/k\mathbb{Z}$ such that $\alpha \in T_i$. Label α with i. By Proposition 3.6, for each i,

$$T_i = \{w \in \mathcal{F} : \text{the sum of the labels of letters in } w \text{ is } i \bmod k\}.$$

In particular, by Proposition 3.4,

$$S \subset T_{k-1} = \{w \in \mathcal{F} : \text{the sum of the labels of letters in } w \text{ is } -1 \bmod k\}.$$

Note that T_{k-1} is strongly k-product-free: if w is the concatenation of $\ell \geqslant 2$ words from T_{k-1}, then the sum of the labels of letters in w is $-\ell \bmod k$.

To obtain the result given in the statement of Theorem 1.4 (i.e. with $1 \bmod k$ instead of $-1 \bmod k$) simply multiply the label of each letter by -1. □

3.2 Structure of k-product-free sets

Let $k \geqslant 2$ be an integer with $k \notin \{3, 5, 7, 13\}$, let $\rho = \rho(k)$, and let $S \subset \mathcal{F}$ be k-product-free satisfying $d^*(S) = 1/\rho$. Note, by Theorem 2.6, that $d^{I_\infty}(S) = 1/\rho = d^{I_\infty}_{\text{sup}}(S)$ and so we may and will frequently apply Lemma 2.9 with $n = \rho + 1$.

To prove Theorem 1.6, we will show that, in fact, S is strongly ρ-product-free and so the result follows from Theorem 1.4. To this end we make the following definition.

Definition 3.7. For a set $A \subset \mathbb{Z}^+$, the set $S_A \subset \mathcal{F}$ is

$$S_A := \bigcap_{i \in A} S^i,$$

where we will omit set parentheses so, for example, $S_1 = S$ and $S_{1,3} = S \cap S^3$.

Since S is k-product-free, $S_{1,k} = \varnothing$. It is enough for us to show that $S_{1,2} = S_{1,3} = \cdots = S_{1,\rho} = \varnothing$ as then S is strongly ρ-product-free. Note that the case $k = 2$ is immediate and so we assume that $k \geqslant 3$ from now on.

We need a quick technical lemma about the size of ρ.

Lemma 3.8. *Let $k \geqslant 3$ be an integer with $k \notin \{3, 5, 7, 13\}$ and let $\rho = \rho(k)$. Then*

$$k - 1 \geqslant \max\{(\rho - t)t(t + 1) : t \in \{1, 2, \ldots, \rho - 1\}\}. \tag{1}$$

Proof. By the arithmetic mean-geometric mean inequality, for any $t \in [0, \rho]$,

$$(\rho - t)t(t + 1) = 4(\rho - t)\tfrac{t}{2}\tfrac{t+1}{2} \leqslant 4\left(\tfrac{\rho+1/2}{3}\right)^3 = 4/27 \cdot (\rho + 1/2)^3.$$

On the other hand, Lev [Lev03, Lem. 18] proved that, for all positive integers $k \geqslant 2$,

$$\rho(k) \leqslant 2 \log_2 k + 2.$$

Now, for all $k \geqslant 2400$,

$$k - 1 \geqslant 4/27 \cdot (2 \log_2 k + 5/2)^3,$$

and so (1) holds. Now, if $\rho \geqslant 10$, then the definition of ρ gives $k - 1 \geqslant 5 \times 7 \times 8 \times 9 = 2520$ and so (1) holds. We are left to check the remaining cases.

- If $\rho = 2$, then the right-hand side of (1) is 2. The smallest $k \geqslant 3$ with $\rho = 2$ is 4.
- If $\rho = 3$, then the right-hand side of (1) is 6. The only $k \leqslant 6$ with $\rho = 3$ are 3 and 5.
- If $\rho = 4$, then the right-hand side of (1) is 12. The only $k \leqslant 12$ with $\rho = 4$ is 7.
- If $\rho = 5$, then the right-hand side of (1) is 24. The only $k \leqslant 24$ with $\rho = 5$ is 13.
- ρ is always the power of a prime so there are no k with $\rho = 6$.
- If $\rho = 7$, then the right-hand side of (1) is 60. The smallest k with $\rho = 7$ is 61.
- If $\rho = 8$, then the right-hand side of (1) is 90. The smallest k with $\rho = 8$ is 421.
- If $\rho = 9$, then the right-hand side of (1) is 126. The smallest k with $\rho = 9$ is 841. □

We first show that $S_{1,\rho}$ is empty.

Proposition 3.9. $S_{1,\rho} = \varnothing$.

Proof. Suppose that $S_{1,\rho} \neq \varnothing$ and let $t \in \mathbb{Z}^+$ be maximal with $S_{1,\rho,2\rho-1,\dots,t(\rho-1)+1} \neq \varnothing$ where the indices form an arithmetic progression with common difference $\rho - 1$. Such a t must exist as $k \equiv 1 \bmod \rho - 1$ and $S_{1,k} = \varnothing$. Let $w \in S_{1,\rho,2\rho-1,\dots,t(\rho-1)+1}$.

Taking $t = \rho - 1$ inside the maximum in (1), we have $k - 1 \geqslant \rho(\rho - 1)$. We split into two cases based on the size of k.

First suppose that $k - 1 > 2\rho(\rho - 1)$. Let $\alpha \in \mathbb{Z}^+$ be minimal such that $(\alpha - 1)\rho(\rho - 1) \geqslant k - 1$. Note that $\alpha \geqslant 4$. Write $\rho = 2a + b$ where $a = \lfloor \rho/2 \rfloor$ and $b \in \{0, 1\}$. Consider the following sets

$$\begin{aligned} &S, \quad w^{\rho-1}S, \qquad w^{\alpha(\rho-1)}S, \quad w^{(\alpha+1)(\rho-1)}S, \qquad w^{2\alpha(\rho-1)}S, \quad w^{(2\alpha+1)(\rho-1)}S, \quad \dots, \\ &w^{(a-1)\alpha(\rho-1)}S, \quad w^{((a-1)\alpha+1)(\rho-1)}S, \qquad w^{a\alpha(\rho-1)}S, \quad w^{(a\alpha+b)(\rho-1)}S. \end{aligned}$$

We remark that these sets are formed by starting with S and then alternating between prepending $w^{\rho-1}$ and $w^{(\alpha-1)(\rho-1)}$. Two sets that differ only by a prepending of $w^{\rho-1}$ are called a *pair*: the pairs are the first and second sets; the third and fourth sets; .... The number of sets listed is $2a + b + 1 = \rho + 1$ and so these cannot all be pairwise disjoint by Lemma 2.9.

We first show that sets in different pairs are disjoint. If two such sets meet, then $S \cap w^{\ell(\rho-1)}S \neq \varnothing$ for some integer ℓ satisfying $\alpha - 1 \leqslant \ell \leqslant a\alpha + b$. We will show that, for such an ℓ, $w^{\ell(\rho-1)} \in S^{k-1}$ which contradicts $S_{1,k} = \varnothing$. Since $w \in S_{1,\rho,2\rho-1,\dots,t(\rho-1)+1}$, we have $w^{\ell(\rho-1)} \in S_{\ell(\rho-1),(\ell+1)(\rho-1),\dots,\ell(\rho-1)(t(\rho-1)+1)}$ where the indices form an arithmetic progression with common difference $\rho - 1$. It suffices to show that $k - 1$ is in this arithmetic progression. Since $k - 1$ is a multiple of $\rho - 1$, it is enough to show that $\ell(\rho - 1) \leqslant k - 1 \leqslant \ell(\rho - 1)(t(\rho - 1) + 1)$ for all integers ℓ satisfying $\alpha - 1 \leqslant \ell \leqslant a\alpha + b$. Now,

$$\ell(\rho - 1)(t(\rho - 1) + 1) \geqslant \ell(\rho - 1)\rho \geqslant (\alpha - 1)\rho(\rho - 1) \geqslant k - 1$$

and

$$\begin{aligned} \ell(\rho - 1) &\leqslant (a\alpha + b)(\rho - 1) = (a\alpha + \rho - 2a)(\rho - 1) \\ &= a(\alpha - 2)(\rho - 1) + \rho(\rho - 1) \leqslant \rho/2 \cdot (\alpha - 2)(\rho - 1) + \rho(\rho - 1) \\ &= \alpha/2 \cdot \rho(\rho - 1) \leqslant (\alpha - 2)\rho(\rho - 1) < k - 1, \end{aligned}$$

where we used the minimality of α and the fact that $\alpha \geqslant 4$ in the final and penultimate inequality respectively.

We second show that sets in the same pair are disjoint which gives the contradiction required to conclude the case $k - 1 > 2\rho(\rho - 1)$. If two sets in the same pair are not disjoint, then $S \cap w^{\rho-1}S \neq \varnothing$. But $w^{\rho-1} \in S_{\rho-1,2(\rho-1),\dots,(t(\rho-1)+1)(\rho-1)}$ and so if $S \cap w^{\rho-1}S \neq \varnothing$, then $S_{1,\rho,2\rho-1,\dots,(t(\rho-1)+1)(\rho-1)+1} \neq \varnothing$ which contradicts the maximality of t.

Second suppose that $2\rho(\rho - 1) \geqslant k - 1 \geqslant \rho(\rho - 1)$. Consider the the following $\rho + 1$ sets

$$S, \quad w^{\rho-1}S, \quad w^{2(\rho-1)}S, \quad \dots, \quad w^{\rho(\rho-1)}S.$$

Since t is maximal, consecutive sets are disjoint as in the previous case. If non-consecutive sets are not disjoint, then $S \cap w^{\ell(\rho-1)}S \neq \varnothing$ for some integer ℓ satisfying

$2 \leqslant \ell \leqslant \rho$. As before, $w^{\ell(\rho-1)} \in S_{\ell(\rho-1),(\ell+1)(\rho-1),\ldots,\ell(\rho-1)(t(\rho-1)+1)}$ and so it suffices to show that $\ell(\rho-1) \leqslant k-1 \leqslant \ell(\rho-1)(t(\rho-1)+1)$ for such ℓ. This is the case as

$$\ell(\rho-1)(t(\rho-1)+1) \geqslant \ell(\rho-1)\rho \geqslant 2\rho(\rho-1) \geqslant k-1$$

and

$$\ell(\rho-1) \leqslant \rho(\rho-1) \leqslant k-1.$$

Thus all $\rho+1$ sets are disjoint contradicting Lemma 2.9 and so $S_{1,\rho}$ is indeed empty. □

We now show that $S_{1,2}, \ldots, S_{1,\rho-1}$ are all empty.

Proposition 3.10. *For all* $1 \leqslant d \leqslant \rho-1$, $S_{1,d+1} = \varnothing$.

Proof. We argue via downwards induction on d with the base case $d = \rho-1$ given by Proposition 3.9. Suppose, towards a contradiction, that the result is false and let $1 \leqslant d \leqslant \rho-2$ be largest with $S_{1,d+1} \neq \varnothing$ and let $t \in \mathbb{Z}^+$ be maximal with $S_{1,d+1,2d+1,\ldots,td+1} \neq \varnothing$. Such a t exists as $k \equiv 1 \bmod d$. Let $w \in S_{1,d+1,2d+1,\ldots,td+1}$.

By the definition of ρ, both d and $d+1$ divide $k-1$. Since d and $d+1$ are coprime we may write $k-1 = \alpha d(d+1)$ for some positive integer α. Let $s \in \mathbb{Z}^+$ be largest such that $ds \leqslant \rho-1$. Write $\rho = a(s+1)+b$ where $a = \lfloor \rho/(s+1) \rfloor$ and $b \in \{0,1,\ldots,s\}$. Consider the following $\rho+1$ sets

$$\begin{array}{llllllll} S, & w^d S, & \ldots, & w^{sd}S, & w^{(\alpha+s)d}S, & w^{(\alpha+s)d+d}S, & \ldots, & w^{(\alpha+s)d+sd}S, \\ w^{2(\alpha+s)d}S, & \ldots, & w^{2(\alpha+s)d+sd}S, & \ldots, & w^{a(\alpha+s)d}S, & \ldots, & w^{a(\alpha+s)d+bd}S. & \end{array}$$

We remark that these sets are formed by starting with S, then prepending w^d a total of s times, prepending $w^{\alpha d}$, then prepending w^d a total of s times, prepending $w^{\alpha d}$, and so on. We group up the sets: the 1^{st} through d^{th} sets are in the first group; the $(d+1)^{\text{st}}$ through $(2d)^{\text{th}}$ sets are in the second group; and so on.

We first show that sets in different groups are disjoint. If two such sets meet, then $S \cap w^{\ell d}S \neq \varnothing$ for some integer ℓ satisfying $\alpha \leqslant \ell \leqslant a(\alpha+s)+b$. Now $w^{\ell d} \in S_{\ell d,(\ell+1)d,\ldots,\ell d(td+1)}$ and so, since $k-1$ is a multiple of d, it suffices to show that $\ell d \leqslant k-1 \leqslant \ell d(td+1)$ for all such ℓ. Firstly,

$$\ell d(td+1) \geqslant \alpha d(d+1) = k-1.$$

Now,

$$\begin{aligned} \ell d &\leqslant (a(\alpha+s)+b)d = (a(\alpha+s)+\rho-a(s+1))d \\ &= (a(\alpha-1)+\rho)d \end{aligned}$$

and we wish to show this is at most $k-1 = \alpha d(d+1)$ and so it is enough to show that $a(\alpha-1)+\rho \leqslant \alpha(d+1)$. By the maximality of s, $d(s+1) \geqslant \rho$ and so $d \geqslant \rho/(s+1) \geqslant a$. Hence, it suffices to show that $d(\alpha-1)+\rho \leqslant \alpha(d+1)$, or equivalently $\rho \leqslant \alpha+d$. But, by Lemma 3.8,

$$\begin{aligned} \alpha d(d+1) = k-1 &\geqslant \max\{(\rho-t)t(t+1) \colon t \in \{1,2,\ldots,\rho-1\}\} \\ &\geqslant (\rho-d)d(d+1), \end{aligned}$$

and so we do indeed have $\rho \leqslant \alpha + d$.

Next we show that sets in the same group are disjoint. If two consecutive sets in the same group meet, then $S \cap w^d S \neq \varnothing$. But $w^d \in S_{d,2d,\ldots,(td+1)d}$ and so if $S \cap w^d S \neq \varnothing$, then $S_{1,d+1,2d+1,\ldots,(td+1)d+1} \neq \varnothing$ which contradicts the maximality of t. If two non-consecutive sets in the same group meet, then $S \cap w^{\ell d} S \neq \varnothing$ for some integer ℓ with $2 \leqslant \ell \leqslant s$. But $w^{\ell d} \in S_{\ell d}$ and so $S_{1,\ell d+1} \neq \varnothing$. However, $d < 2d \leqslant \ell d \leqslant ds \leqslant \rho - 1$ and so this contradicts the maximality of d.

Hence, all $\rho + 1$ sets are pairwise disjoint which contradicts Lemma 2.9, as required. □

Propositions 3.9 and 3.10 together show that S is strongly ρ-product-free. Theorem 1.6 then follows from Theorem 1.4.

4 Density of product-free sets in the free semigroup

In this section we will bound the density of a (strongly) k-product-free set $S \subset \mathcal{F}$ and so prove Theorem 2.6. To motivate our approach, assume that S is strongly 3-product-free and that we want to bound $d^{I_\infty}(S)$. We might hope that $d^{I_\infty}(S) = d^{I_\infty}(S^2) = d^{I_\infty}(S^3)$. Because all of these sets are disjoint, this would imply that $d^{I_\infty}(S \cup S^2 \cup S^3) = 3 \cdot d^{I_\infty}(S)$ and so $d^{I_\infty}(S) \leqslant 1/3$, as required.

If $S\mathcal{F}$ covers all of $\mathcal{F}$, such an argument works. Indeed, note that for all $w \in S$ we have $d^{I_\infty}_{w\mathcal{F}}(S^2) \geqslant d^{I_\infty}_{w\mathcal{F}}(wS) = d^{I_\infty}(S)$. If $S\mathcal{F}$ covers all of $\mathcal{F}$, an averaging argument would then show that $d^{I_\infty}(S^2) \geqslant d^{I_\infty}(S)$, and so $d^{I_\infty}(S \cup S^2) \geqslant 2 \cdot d^{I_\infty}(S)$. To include S^3 in the union, we can just repeat the argument. For $w \in S$ we have $d^{I_\infty}_{w\mathcal{F}}(S^2 \cup S^3) \geqslant d^{I_\infty}_{w\mathcal{F}}(w(S \cup S^2)) = d^{I_\infty}(S \cup S^2) \geqslant 2 \cdot d^{I_\infty}(S)$, giving $d^{I_\infty}(S^2 \cup S^3) \geqslant 2 \cdot d^{I_\infty}(S)$ and thus $d^{I_\infty}(S \cup S^2 \cup S^3) \geqslant 3 \cdot d^{I_\infty}(S)$, as required.

If $S\mathcal{F}$ fails to cover some part of $\mathcal{F}$, we simply want to ignore the uncovered part. Within $S\mathcal{F}$, the relative density of S is at least $d^{I_\infty}(S)$ (we define relative density properly after Lemma 4.3), so it would suffice to show that the relative densities of $S \cup S^2$ and $S \cup S^2 \cup S^3$ are at least $2 \cdot d^{I_\infty}(S)$ and $3 \cdot d^{I_\infty}(S)$ respectively. While the relative density of $S \cup S^2$ can be computed as before, this fails for $S \cup S^2 \cup S^3$. We only know that $S \cup S^2$ has a high density within $S\mathcal{F}$, for example $d^{I_\infty}_{v\mathcal{F}}(S \cup S^2) \geqslant 2 \cdot d^{I_\infty}(S)$ for some $v \in S$. This does not suffice to get a lower bound on $d^{I_\infty}_{w\mathcal{F}}(S^2 \cup S^3)$ in the calculation above.

Instead, note that $d^{I_\infty}_{wv\mathcal{F}}(S^2 \cup S^3) \geqslant d^{I_\infty}_{wv\mathcal{F}}(w(S \cup S^2)) = d^{I_\infty}_{v\mathcal{F}}(S \cup S^2) \geqslant 2 \cdot d^{I_\infty}(S)$ which tells us that the relative density of $S^2 \cup S^3$ in $Sv\mathcal{F}$ is at least $2 \cdot d^{I_\infty}(S)$. If we could show that $Sv\mathcal{F}$ covers essentially all of $S\mathcal{F}$, this would imply that the relative density of $S^2 \cup S^3$ in $S\mathcal{F}$ is at least $2 \cdot d^{I_\infty}(S)$ which in turn would suffice to show that $d^{I_\infty}(S) \leqslant 1/3$.

The technical arguments in Section 4.1 are mostly devoted to showing that this is true up to some small error. In Section 4.2 we will then use this machinery to prove Theorem 2.6. While proving an upper bound on the density of a strongly k-product-free set works as outlined above, the proof for k-product-free sets is more involved and requires some ideas similar to those from Section 3.2.

4.1 Steeplechases

In this section we essentially show that $Sv\mathcal{F}$ covers almost all of $S\mathcal{F}$ up to some small error. The main idea is to approximate S by finite prefix-free subsets (C_t) of S such that, for all t, $C_{t+1} \subset C_t\mathcal{F}$ and $C_t\mathcal{F}$ covers almost all of S (here almost all means up to some set of density ε). In particular, for all large t and all $t' > t$, $C_{t'}\mathcal{F}$ covers almost all of $C_t\mathcal{F}$.

Now, $C_t v\mathcal{F}$ covers an $|\mathcal{A}|^{-|v|}$-fraction of $C_t\mathcal{F}$. Since $C_{t+1}\mathcal{F}$ covers almost all of $C_t\mathcal{F}$, it covers almost all of $C_t\mathcal{F} \setminus C_t v\mathcal{F}$. So $C_{t+1}v\mathcal{F}$ will cover almost an $|\mathcal{A}|^{-|v|}$-fraction of $C_t\mathcal{F} \setminus C_t v\mathcal{F}$. Repeating this argument with $C_{t+2}\mathcal{F}$, we get that $C_{t+2}v\mathcal{F}$ will cover almost an $|\mathcal{A}|^{-|v|}$-fraction of $C_t\mathcal{F} \setminus (C_t v\mathcal{F} \cup C_{t+1}v\mathcal{F})$. Iterating this further shows that $\bigcup_{\ell \geqslant t} C_\ell v\mathcal{F}$ covers almost all of $C_t\mathcal{F}$ (now up to some set of density 2ε). This implies that $Sv\mathcal{F}$ covers almost all of $S\mathcal{F}$, which is what we needed.

The existence of the prefix-free subsets (C_t) which approximate S is established in Lemma 4.2. Lemma 4.3 then shows that $\bigcup_{\ell \geqslant t} C_\ell v\mathcal{F}$ covers almost all of $C_t\mathcal{F}$. Finally, Lemma 4.6, our key technical lemma, uses these to give a result that will allow us to carry out the density arguments sketched above.

We start by defining the properties of the sets (C_t) from the preceding discussion.

Definition 4.1 (steeplechase)**.** An infinite sequence (C_t) of subsets of $\mathcal{F}$ is a *steeplechase* if, for each positive integer t,

- C_t is prefix-free and finite,
- every word in C_{t+1} has a proper prefix in C_t (in particular, $C_{t+1}\mathcal{F} \subset C_t\mathcal{F}$).

Steeplechase (C_t) is *spread* if $\max C_t < \min C_{t+1}$ for all t and is *ε-tight* if, for all m, n, $|\mu(C_m) - \mu(C_n)| \leqslant \varepsilon$.

Every steeplechase contains a spread steeplechase. Indeed, note that $\min C_t \geqslant t$, since every word in C_i has a proper prefix in C_{i-1}. Let $\ell_1 = \max C_1$. Then $\min C_{\ell_1+1} > \ell_1 = \max C_1$. Let $\ell_2 = \max C_{\ell_1+1}$. Then $\min C_{\ell_2+1} > \max C_{\ell_1+1}$. Iteratively doing this gives a spread steeplechase $C_1, C_{\ell_1+1}, C_{\ell_2+1}, \ldots$.

Since C_t is prefix-free, $\mu(C_t) \in [0,1]$. Also, for each t, $C_{t+1}\mathcal{F} \subset C_t\mathcal{F}$ and so, by Observation 2.2, the sequence $(\mu(C_t))$ is non-increasing. In particular, this sequence tends to a limit. Hence the sequence is Cauchy: for any $\varepsilon > 0$, there is a K such that, for all $m, n \geqslant K$, $|\mu(C_m) - \mu(C_n)| \leqslant \varepsilon$. Thus, ignoring the first few C_t gives an ε-tight steeplechase.

In particular, given any steeplechase (C_t) we may, by passing to a subsequence, assume that (C_t) is both spread and ε-tight.

The following lemma shows that, for any set $B \subset \mathcal{F}$, there is a steeplechase that approximates B very well.

Lemma 4.2. *Let $\varepsilon > 0$ and $B \subset \mathcal{F}$. There is an ε-tight spread steeplechase (C_t) such that*

- $C_1 \cup C_2 \cup \cdots \subset B$,
- *for all t and all large n (in terms of t), $\mu((B \setminus C_t\mathcal{F})(n)) \leqslant \varepsilon$,*
- *for all t, $\mu(C_t) \geqslant d^{I_\infty}(B) - \varepsilon$.*

Proof. For $x \in B$, let the *headcount* of x be

$$h(x) = |\{b \in B\colon b \text{ is a prefix of } x\}|.$$

For each positive integer t, let $D_t = \{x \in B\colon h(x) = t\}$. Note that each D_t is prefix-free and so $\mu(D_t) \leqslant 1$ for all t. Iteratively do the following procedure for each positive integer t.

1. Let ℓ_t be such that $\mu(D_t(\{\ell_t + 1, \ell_t + 2, \ldots\})) \leqslant \varepsilon/2^t$.
2. Let $C_t = D_t(\{1, 2, \ldots, \ell_t\})$.
3. Remove $(D_t \setminus C_t)\mathcal{F}$ from B (including from all later D_i).

Let B' be the set remaining at the end of this procedure. Note that in each application of step 3, for every word w, either $h(w)$ does not change or w is removed from B. In particular, every word in C_t has a proper prefix in C_{t-1}. Also, by construction, C_t is a finite subset of B. Thus (C_t) is a steeplechase and $C_1 \cup C_2 \cup \cdots \subset B$.

Fix t and let $n > \max\{\ell_1, \ldots, \ell_t\}$. Step 3 removes all words with headcount r and length greater than ℓ_r from B. Thus, any word of length n in B' has headcount greater than t and so is in $C_t\mathcal{F}$. Thus, $B'(n) \subset C_t\mathcal{F}(n)$. Next note that B' is obtained from B by deleting all the $(D_r \setminus C_r)\mathcal{F}$ and so,

$$\mu((B \setminus C_t\mathcal{F})(n)) \leqslant \mu((B \setminus B')(n)) \leqslant \sum_r \mu(((D_r \setminus C_r)\mathcal{F})(n)) \leqslant \sum_r \mu(D_r \setminus C_r) \leqslant \varepsilon.$$

Finally, this implies that $\mu(B(n)) \leqslant \mu(C_t\mathcal{F}(n)) + \varepsilon \leqslant \mu(C_t) + \varepsilon$. Averaging this over $n \in I_j$ and taking $j \to \infty$ gives $d^{I_\infty}(B) \leqslant \mu(C_t) + \varepsilon$.

Hence (C_t) is a steeplechase satisfying all three conditions. As noted above, we may, by passing to a subsequence, assume that (C_t) is spread and ε-tight. Passing to a subsequence does not affect the three conditions. □

We call a steeplechase (C_t) given by Lemma 4.2 an *ε-capturing steeplechase for B*. As promised, we now show for any word w that $\bigcup_{\ell \geqslant 1} C_\ell w\mathcal{F}$ covers almost all of $C_1\mathcal{F}$.

Lemma 4.3. *Let (C_t) be an ε-tight spread steeplechase. For every $w \in \mathcal{F}$ there is an N such that the following holds. If $C = C_1 \cup C_2 \cup \cdots \cup C_N$ and $n \geqslant \max C_N + |w|$, then*

$$\mu((C_1\mathcal{F} \setminus Cw\mathcal{F})(n)) \leqslant 2\varepsilon.$$

Proof. Let N sufficiently large in terms of $|w|$ and $|\mathcal{A}|$ and let $n \geqslant \max C_N + |w|$. Now

$$\mu((C_1\mathcal{F} \setminus C_N\mathcal{F})(n)) = \mu((C_1\mathcal{F})(n)) - \mu((C_N\mathcal{F})(n)) = \mu(C_1) - \mu(C_N) \leqslant \varepsilon,$$

since $C_N\mathcal{F} \subset C_1\mathcal{F}$ and (C_t) is ε-tight. Hence, it suffices to prove that

$$\mu((C_N\mathcal{F} \setminus Cw\mathcal{F})(n)) \leqslant \varepsilon.$$

Let X be the following finite prefix-free set

$$X = \{s \in C_N\colon s \text{ has no prefix in } Cw\}.$$

Note that $(C_N\mathcal{F} \setminus Cw\mathcal{F})(n) \subset (X\mathcal{F})(n)$ and so

$$\mu((C_N\mathcal{F} \setminus Cw\mathcal{F})(n)) \leqslant \mu((X\mathcal{F})(n)) = \mu(X).$$

Recall the random infinite word $\mathbf{W} = \alpha_1\alpha_2\cdots$ and corresponding random walk defined in Section 2. Since X is prefix-free, $\mu(X) = \mathbb{P}(\mathbf{W} \text{ hits } X)$ and it suffices to show this probability is at most ε. Let K be the largest integer with $1 + K|w| \leqslant N - |w|$. If $\mathbf{W}$ hits X, then $\mathbf{W}$ hits C_N and so, since (C_t) is a steeplechase, $\mathbf{W}$ hits each of $C_1, C_{1+|w|}, \ldots, C_{1+K|w|}$. Also, $\mathbf{W}$ must avoid each of $C_1 w, C_{1+|w|}w, \ldots, C_{1+K|w|}w$ in order to hit X.

We reveal the letters of $\mathbf{W}$ one-by-one. We wait until $\mathbf{W}$ hits/avoids C_1 (this will certainly be known by the time the length of $\mathbf{W}$ is $\max C_1$). If $\mathbf{W}$ avoids C_1, then $\mathbf{W}$ avoids X. If $\mathbf{W}$ hits C_1, then we reveal the next $|w|$ letters of $\mathbf{W}$ and check if they spell w (this has probability $|\mathcal{A}|^{-|w|}$). If they do, then $\mathbf{W}$ avoids X. If they do not, then we wait until $\mathbf{W}$ hits/avoids $C_{1+|w|}$: note that this has not already happened since (C_t) is spread and so $\min C_{1+|w|} \geqslant \max C_1 + |w|$. If $\mathbf{W}$ avoids $C_{1+|w|}$, then $\mathbf{W}$ avoids X. If $\mathbf{W}$ hits $C_{1+|w|}$, then we reveal the next $|w|$ letters of $\mathbf{W}$ and check if they spell w (this has probability $|A|^{-|w|}$). We continue this procedure with the final check being whether the next $|w|$ letters of $\mathbf{W}$ after it hits $C_{1+K|w|}$ spell w. Note that each check has probability $|\mathcal{A}|^{-|w|}$ and is independent of the previous checks (new letters are involved in each check). If $\mathbf{W}$ hits X, then $\mathbf{W}$ must fail each of these spelling checks and so the probability that $\mathbf{W}$ hits X is at most

$$(1 - |\mathcal{A}|^{-|w|})^{K+1}.$$

By taking N (and so K) sufficiently large in terms of $|w|$ and $|\mathcal{A}|$ we may ensure this is at most ε, as required. □

Before proving our key technical result for our density proofs (Lemma 4.6) we will need to define the relative density of B on $C\mathcal{F}$. If $C \subset \mathcal{F}$ is finite and interval I satisfies $\min I \geqslant \max C$, then the *relative density of B in $C\mathcal{F}$ on interval I* is

$$d^I_{C\mathcal{F}}(B) := \frac{\mu(B(I) \cap C\mathcal{F})}{\mu(\mathcal{F}(I) \cap C\mathcal{F})}.$$

Suppose C is also prefix-free. Then, by Observation 2.2, $\mu(\mathcal{F}(I) \cap C\mathcal{F}) = |I|\mu(C)$. Also $(c\mathcal{F} \colon c \in C)$ partition $C\mathcal{F}$. In particular,

$$\begin{aligned}
d^I_{C\mathcal{F}}(B) &= |I|^{-1}\mu(C)^{-1} \sum_{n \in I} \mu(B(n) \cap C\mathcal{F}) \\
&= \sum_{c \in C} |I|^{-1}\mu(C)^{-1} \sum_{n \in I} \mu(B(n) \cap c\mathcal{F}) \\
&= \sum_{c \in C} \frac{\mu(c)}{\mu(C)} \cdot d^I_{c\mathcal{F}}(B) \\
&= \mu(C)^{-1} \sum_{c \in C} d^I(B \cap c\mathcal{F}) \\
&= \mu(C)^{-1} \cdot d^I(B \cap C\mathcal{F}).
\end{aligned}$$

Recall that, for every set B that we consider in this paper and every word c, the sequence $d^{I_j}_{c\mathcal{F}}(B)$ converges (to $d^{I_\infty}_{c\mathcal{F}}(B)$). Thus the third line of the displayed equation implies that the sequence $d^{I_j}_{C\mathcal{F}}(B)$ converges to a limit $d^{I_\infty}_{C\mathcal{F}}(B)$. Again, these limits are additive.

Observation 4.4. *Let $C \subset \mathcal{F}$ be finite and prefix-free. Then*

$$d^{I_\infty}_{C\mathcal{F}}(B) = \sum_{c \in C} \frac{\mu(c)}{\mu(C)} \cdot d^{I_\infty}_{c\mathcal{F}}(B) = \mu(C)^{-1} \cdot d^{I_\infty}(B \cap C\mathcal{F}).$$

Observation 4.5. *The first equality of Observation 4.4 implies that* $d^{I_\infty}_{C\mathcal{F}}(B) \leqslant d^{I_\infty}_{\sup}(B)$.

Now for the key technical lemma for our density results.

Lemma 4.6. *Let* $\varepsilon > 0$ *and let* $A, B \subset \mathcal{F}$. *If* (C_t) *is an* ε*-capturing steeplechase for* A *with* $\mu(C_1) \geqslant 2\varepsilon + \varepsilon^{1/3}$, *then*

$$d^{I_\infty}_{C_1\mathcal{F}}(AB) \geqslant d^{I_\infty}_{\sup}(B) - 3\varepsilon^{1/3} \geqslant d^{I_\infty}_{C_1\mathcal{F}}(B) - 3\varepsilon^{1/3}.$$

Proof. Let $w \in \mathcal{F}$ be such that

$$d^{I_\infty}_{w\mathcal{F}}(B) \geqslant d^{I_\infty}_{\sup}(B) - \varepsilon^{1/3}.$$

Apply Lemma 4.3 to (C_t) and w to give an N such that letting $C = C_1 \cup \cdots \cup C_N$, if $n \geqslant \max C_N + |w|$, then

$$\mu((C_1\mathcal{F} \setminus Cw\mathcal{F})(n)) \leqslant 2\varepsilon.$$

We may greedily choose $\widetilde{C} \subset C$ (starting with shorter words first) such that $\widetilde{C}w$ is prefix-free and $\widetilde{C}w\mathcal{F} = Cw\mathcal{F}$. Note that

$$\begin{aligned} 2\varepsilon &\geqslant \mu((C_1\mathcal{F} \setminus \widetilde{C}w\mathcal{F})(n)) \geqslant \mu((C_1\mathcal{F})(n)) - \mu((\widetilde{C}w\mathcal{F})(n)) \\ &= \mu(C_1) - \mu(\widetilde{C}w) \geqslant 2\varepsilon + \varepsilon^{1/3} - \mu(\widetilde{C}w) \end{aligned}$$

and so $\mu(\widetilde{C}w) \geqslant \varepsilon^{1/3}$.

Let I be an interval with $\min I \geqslant \max C_N + |w|$ and let $X \subset \mathcal{F}$. Note that every word in $\widetilde{C}w\mathcal{F} = Cw\mathcal{F}$ has a prefix in $C_1 \cup \cdots \cup C_N$. But since (C_t) is a steeplechase, every such word has a prefix in C_1. Thus $\widetilde{C}w\mathcal{F} \subset C_1\mathcal{F}$ and so

$$\begin{aligned} d^{I}_{C_1\mathcal{F}}(X) &= |I|^{-1}\mu(C_1)^{-1} \sum_{n\in I} \mu(X(n) \cap C_1\mathcal{F}) \\ &\geqslant |I|^{-1}\mu(C_1)^{-1} \sum_{n\in I} \mu(X(n) \cap \widetilde{C}w\mathcal{F}) \\ &\geqslant |I|^{-1} \sum_{n\in I} \frac{\mu(X(n) \cap \widetilde{C}w\mathcal{F})}{\mu(\widetilde{C}w) + 2\varepsilon}. \end{aligned}$$

Using the fact that $x/(y + 2\varepsilon) \geqslant x/y - 2\varepsilon x/y^2 \geqslant x/y - 2\varepsilon^{1/3}$ for $\varepsilon > 0$, $x \in [0, 1]$, and $y \geqslant \varepsilon^{1/3}$, we have

$$d^{I}_{C_1\mathcal{F}}(X) \geqslant d^{I}_{\widetilde{C}w\mathcal{F}}(X) - 2\varepsilon^{1/3}.$$

Setting $X = AB$, $I = I_j$, and taking j to infinity gives

$$d^{I_\infty}_{C_1\mathcal{F}}(AB) \geqslant d^{I_\infty}_{\widetilde{C}w\mathcal{F}}(AB) - 2\varepsilon^{1/3}. \tag{2}$$

Now,

$$\begin{aligned} d^{I_\infty}_{\widetilde{C}w\mathcal{F}}(AB) &= \sum_{c\in\widetilde{C}} \frac{\mu(c)}{\mu(\widetilde{C})} \cdot d^{I_\infty}_{cw\mathcal{F}}(AB) \\ &\geqslant \sum_{c\in\widetilde{C}} \frac{\mu(c)}{\mu(\widetilde{C})} \cdot d^{I_\infty}_{cw\mathcal{F}}(cB) \\ &= \sum_{c\in\widetilde{C}} \frac{\mu(c)}{\mu(\widetilde{C})} \cdot d^{I_\infty}_{w\mathcal{F}}(B) \\ &= d^{I_\infty}_{w\mathcal{F}}(B) \geqslant d^{I_\infty}_{\sup}(B) - \varepsilon^{1/3}, \end{aligned}$$

where the first equality used Observation 4.4, the first inequality used the fact that $c \in \widetilde{C} \subset C \subset A$, the second equality used Lemma 2.4, and the second inequality is due to the choice of w. Combining this with (2) gives the first inequality claimed, while Observation 4.5 gives the second. □

4.2 Density of (strongly) k-product-free sets

In this section we prove Theorem 2.6, making use of the machinery developed in the previous section. Part (a) has a simple iterating proof which uses that a strongly k-product-free S is disjoint from each of $S^2, S^3, \ldots, S^k$.

Proof of Theorem 2.6(a). Let $S \subset \mathcal{F}$ be strongly k-product-free and let w be any word of $\mathcal{F}$. It suffices to show that $d^{I_\infty}_{w\mathcal{F}}(S) \leqslant 1/k$. Now $d^{I_\infty}_{w\mathcal{F}}(S) = d^{I_\infty}_{w\mathcal{F}}(S \cap w\mathcal{F})$ and so we may restrict S to $S \cap w\mathcal{F}$ (it remains strongly k-product-free). Thus we may and will assume that $S \subset w\mathcal{F}$. If $d^{I_\infty}_{w\mathcal{F}}(S) = 0$, then we are done. Otherwise, by the definition of $d^{I}_{w\mathcal{F}}$ and the fact that $S \subset w\mathcal{F}$, $d^{I_\infty}(S) = \mu(w) \cdot d^{I_\infty}_{w\mathcal{F}}(S) > 0$.

Let $\varepsilon > 0$ be sufficiently small compared to $d^{I_\infty}(S)$ and let (C_t) be an ε-capturing steeplechase for S, as given by Lemma 4.2. Then $\mu(C_1) \geqslant d^{I_\infty}(S) - \varepsilon \geqslant 2\varepsilon + \varepsilon^{1/3}$ and so we may use Lemma 4.6 applied to S and some $B \subset \mathcal{F}$. We will show by induction that, for $1 \leqslant i \leqslant k$,

$$d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^i) \geqslant i \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3(i-1) \cdot \varepsilon^{1/3}.$$

For $i = 1$ this holds with equality. Suppose it holds for some $1 \leqslant i < k$. Lemma 4.6 implies that

$$\begin{aligned} d^{I_\infty}_{C_1\mathcal{F}}(S^2 \cup S^3 \cup \cdots \cup S^{i+1}) &\geqslant d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^i) - 3\varepsilon^{1/3} \\ &\geqslant i \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3i \cdot \varepsilon^{1/3}. \end{aligned}$$

Since $i < k$ and S is strongly k-product-free, S is disjoint from $S^2 \cup S^3 \cup \cdots \cup S^{i+1}$ and so

$$d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^{i+1}) \geqslant (i+1) \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3i \cdot \varepsilon^{1/3},$$

which completes the induction. Taking $i = k$ we obtain

$$1 \geqslant d^{I_\infty}_{C_1\mathcal{F}}(S \cup S^2 \cup \cdots \cup S^k) \geqslant k \cdot d^{I_\infty}_{C_1\mathcal{F}}(S) - 3(k-1) \cdot \varepsilon^{1/3}.$$

Note that $C_1 \subset S \subset w\mathcal{F}$ is prefix-free so $\mu(C_1) \leqslant \mu(w)$. Since (C_t) is ε-capturing for S, Observation 4.4 and the second bullet point of Lemma 4.2 imply that

$$\begin{aligned} d^{I_\infty}_{C_1\mathcal{F}}(S) = \mu(C_1)^{-1} \cdot d^{I_\infty}(S \cap C_1\mathcal{F}) &\geqslant \mu(w)^{-1} \cdot d^{I_\infty}(S \cap C_1\mathcal{F}) \\ &\geqslant \mu(w)^{-1} \cdot (d^{I_\infty}(S) - \varepsilon) = d^{I_\infty}_{w\mathcal{F}}(S) - \varepsilon \cdot \mu(w)^{-1}, \end{aligned}$$

where the final inequality used the fact that $d^{I_\infty}(S) = \mu(w) \cdot d^{I_\infty}_{w\mathcal{F}}(S)$ as noted at the beginning of this proof. Putting everything together

$$1 \geqslant k \cdot d^{I_\infty}_{w\mathcal{F}}(S) - k\varepsilon \cdot \mu(w)^{-1} - 3(k-1) \cdot \varepsilon^{1/3}.$$

Since ε is arbitrarily small, we obtain $d^{I_\infty}_{w\mathcal{F}}(S) \leqslant 1/k$, as required. □

The argument for part (b) (k-product-free sets) is more involved. It is not necessary to keep track of the error term depending on ε (as we eventually take ε to zero). We introduce some notation to simplify the argument. Write $x \lesssim y$ to mean that $x \leqslant y + f(\varepsilon)$ where the error term $f(\varepsilon)$ goes to zero as ε goes to zero (in all cases $f(\varepsilon)$ will be a polynomial in $\varepsilon^{1/3}$).

To improve clarity and motivate the proof we first sketch a proof of that if S is 3-product-free, then $d^{I_\infty}(S) \leqslant 1/3$.

Proof that $d^{I_\infty}(S) \leqslant 1/3$ for 3-product-free S. Let $S \subset \mathcal{F}$ be 3-product-free and let $\varepsilon > 0$ be sufficiently small. Let (C_t) be an ε-capturing steeplechase for $S_{1,2} = S \cap S^2$ (recall Definition 3.7), as given by Lemma 4.2. Since S is 3-product-free, $S_{1,2}$ is strongly 3-product-free.

We claim that $d^{I_\infty}(S \cap C_1\mathcal{F}) \lesssim 1/3 \cdot \mu(C_1)$. If $\mu(C_1) < 2\varepsilon + \varepsilon^{1/3}$, then this is immediate. Otherwise $\mu(C_1) \geqslant 2\varepsilon + \varepsilon^{1/3}$ and so, by Lemma 4.6, $d^{I_\infty}_{C_1\mathcal{F}}(S_{1,2}S) \gtrsim d^{I_\infty}_{C_1\mathcal{F}}(S)$. Since S is 3-product-free, S and $S_{1,2}S$ are disjoint and so

$$d^{I_\infty}_{C_1\mathcal{F}}(S \cup S_{1,2}S) \gtrsim 2 \cdot d^{I_\infty}_{C_1\mathcal{F}}(S).$$

Then, applying Lemma 4.6 again,

$$d^{I_\infty}_{C_1\mathcal{F}}(S_{1,2}S \cup S_{1,2}^2S) = d^{I_\infty}_{C_1\mathcal{F}}(S_{1,2}(S \cup S_{1,2}S)) \gtrsim d^{I_\infty}_{C_1\mathcal{F}}(S \cup S_{1,2}S) \gtrsim 2 \cdot d^{I_\infty}_{C_1\mathcal{F}}(S).$$

Since S is 3-product-free, S is disjoint from $S_{1,2}S \cup S_{1,2}^2S$ and so

$$1 \geqslant d^{I_\infty}_{C_1\mathcal{F}}(S \cup S_{1,2}S \cup S_{1,2}^2S) \gtrsim 3 \cdot d^{I_\infty}_{C_1\mathcal{F}}(S).$$

Observation 4.4 then gives $d^{I_\infty}(S \cap C_1\mathcal{F}) \lesssim 1/3 \cdot \mu(C_1)$, as claimed.

We have bounded the density of S on the part of $\mathcal{F}$ where $S_{1,2}$ is dense. We now bound the density of S on the rest. Let $S' = S \setminus (S_{1,2} \cup C_1\mathcal{F})$ and (D_t) be an ε-capturing steeplechase for S', as given by Lemma 4.2. By passing to a subsequence we may and will assume that $\min D_1 > \max C_1$. Since S is 3-product-free and $S' \cap S^2 = \varnothing$, S' is strongly 3-product-free.

We claim that $d^{I_\infty}(S' \cap D_1\mathcal{F}) \lesssim 1/3 \cdot \mu(D_1)$. If $\mu(D_1) < 2\varepsilon + \varepsilon^{1/3}$, then this is immediate. Otherwise, by Lemma 4.6, $d^{I_\infty}_{D_1\mathcal{F}}(S'S) \gtrsim d^{I_\infty}_{D_1\mathcal{F}}(S) \geqslant d^{I_\infty}_{D_1\mathcal{F}}(S')$. Now S' and $S'S$ are disjoint since $S' \cap S^2 = \varnothing$. Thus

$$d^{I_\infty}_{D_1\mathcal{F}}(S' \cup S'S) \gtrsim 2 \cdot d^{I_\infty}_{D_1\mathcal{F}}(S').$$

Then, applying Lemma 4.6 again,

$$d^{I_\infty}_{D_1\mathcal{F}}((S')^2 \cup (S')^2S) = d^{I_\infty}_{D_1\mathcal{F}}(S'(S' \cup S'S)) \gtrsim 2 \cdot d^{I_\infty}_{D_1\mathcal{F}}(S').$$

Since S is 3-product-free and $S' \cap S^2 = \varnothing$, S' is disjoint from $(S')^2 \cup (S')^2S$ and so

$$1 \geqslant d^{I_\infty}_{D_1\mathcal{F}}(S' \cup (S')^2 \cup (S')^2S) \gtrsim 3 \cdot d^{I_\infty}_{D_1\mathcal{F}}(S').$$

Observation 4.4 then gives $d^{I_\infty}(S' \cap D_1\mathcal{F}) \lesssim 1/3 \cdot \mu(D_1)$, as claimed.

By the definition of S' and since $\min D_1 > \max C_1$, it follows that $C_1\mathcal{F}$ and $D_1\mathcal{F}$ are disjoint (see proof of Proposition 4.7 for more details). In particular, C_1 and D_1 are disjoint and their union is prefix-free. Hence $\mu(C_1) + \mu(D_1) \leqslant 1$. Thus,

$$d^{I_\infty}(S \cap C_1\mathcal{F}) + d^{I_\infty}(S' \cap D_1\mathcal{F}) \lesssim 1/3 \cdot (\mu(C_1) + \mu(D_1)) \leqslant 1/3.$$

Since (C_t) and (D_t) are ε-capturing, it follows (see the proof of Claim 4.7.2 below) that very little of S lies outside $(S \cap C_1\mathcal{F}) \cup (S' \cap D_1\mathcal{F})$. In particular, $d^{I_\infty}(S) \lesssim 1/3$. Since ε can be arbitrarily small, we have $d^{I_\infty}(S) \leqslant 1/3$. □

For general k the argument is a more involved version of the above. We first consider some S_{A_1}, take some ε-capturing steeplechase, (C_t), for S_{A_1} and show that the density of S relative to $C_1\mathcal{F}$ is at most $1/\rho$. We then repeat this step for some $S_{A_2}, S_{A_3}, \ldots$. In future steps we may use the fact that we have dealt with previous S_{A_i}. Proposition 4.7 says that if we have chosen a suitable sequence $A_1, A_2, \ldots$, then we obtain the required bound on $d^{I_\infty}_{\sup}(S)$, and Proposition 4.8 shows that for each k there is a suitable sequence of A_i. These combine to complete the proof of Theorem 2.6(b).

Note in the statement below that dA_ℓ is the sumset

$$dA_\ell := \{a_1 + \cdots + a_d \colon a_1, \ldots, a_d \in A_\ell\}.$$

Proposition 4.7. *Let $k \geqslant 2$ be an integer and $A_1, \ldots, A_m \subset \mathbb{N}$ be a sequence of sets with $A_m = \{1\}$. Suppose that for all $\ell \in [m]$ there exist positive integers $d_1, \ldots, d_{\rho-1}$ (which may depend on ℓ) such that, for all $1 \leqslant i \leqslant j \leqslant \rho - 1$, either*

- *$k \in \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$ or*
- *$A_{\ell'} \subset \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$ for some $1 \leqslant \ell' < \ell$.*

If $S \subset \mathcal{F}$ is k-product-free, then $d^{I_\infty}_{\sup}(S) \leqslant 1/\rho(k)$.

Proof. Let $\rho = \rho(k)$ and let w be any word of $\mathcal{F}$. It suffices to show that $d^{I_\infty}_{w\mathcal{F}}(S) \leqslant 1/\rho$. Now $d^{I_\infty}_{w\mathcal{F}}(S) = d^{I_\infty}_{w\mathcal{F}}(S \cap w\mathcal{F})$ and so we may restrict S to $S \cap w\mathcal{F}$ (it remains k-product-free). Thus we may and will assume that $S \subset w\mathcal{F}$.

Let $\varepsilon > 0$ be sufficiently small. We define the following sets and steeplechases. Take $S^{(1)} := S$, $R^{(1)} := S_{A_1} \cap S^{(1)}$, and let $(C_t^{(1)})$ be an ε-capturing steeplechase for $R^{(1)}$, as given by Lemma 4.2. For $\ell = 2, 3, \ldots, m$, do the following:

- Set $S^{(\ell)} := S \setminus (S_{A_1} \cup \cdots \cup S_{A_{\ell-1}} \cup (C_1^{(1)} \cup \cdots \cup C_1^{(\ell-1)})\mathcal{F})$ and $R^{(\ell)} := S_{A_\ell} \cap S^{(\ell)}$.
- Take $(C_t^{(\ell)})$ to be an ε-capturing steeplechase for $R^{(\ell)}$. By passing to a subsequence of the steeplechase we may and will assume that $\min C_1^{(\ell)} > \max C_1^{(\ell-1)}$.

Claim 4.7.1. *For each $\ell \in [m]$, $\mu(C_1^{(\ell)}) \gtrsim \rho \cdot d^{I_\infty}(S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})$.*

Proof. For clarity, we will write R for $R^{(\ell)}$ in this proof. Firstly, if $\mu(C_1^{(\ell)}) < 2\varepsilon + \varepsilon^{1/3}$, then

$$\rho \cdot d^{I_\infty}(S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F}) \leqslant \rho \cdot d^{I_\infty}(C_1^{(\ell)}\mathcal{F}) = \rho \cdot \mu(C_1^{(\ell)}) \lesssim \mu(C_1^{(\ell)}).$$

Hence, we may assume and will assume that $\mu(C_1^{(\ell)}) \geqslant 2\varepsilon + \varepsilon^{1/3}$. We will prove by downwards induction on $1 \leqslant i \leqslant \rho$ that

$$d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S \cup R^{d_i}S \cup R^{d_i+d_{i+1}}S \cup \cdots \cup R^{d_i+\cdots+d_{\rho-1}}S) \gtrsim (\rho - i + 1) \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}).$$

For $i = \rho$ this holds since $S^{(\ell)} \subset S$. Suppose that it holds for some $1 < i \leqslant \rho$. Since $\mu(C_1^{(\ell)}) \geqslant 2\varepsilon + \varepsilon^{1/3}$, repeatedly applying Lemma 4.6 shows that $d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(R^{d_{i-1}}B) \gtrsim d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(B)$ for any set B. In particular,

$$\begin{aligned} & d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S) \\ \gtrsim\ & d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S \cup R^{d_i}S \cup R^{d_i+d_{i+1}}S \cup \cdots \cup R^{d_i+\cdots+d_{\rho-1}}S) \\ \gtrsim\ & (\rho - i + 1) \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}). \end{aligned}$$

Consider the sets $S^{(\ell)}$ and $R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S$. First note that $R^{d_{i-1}+\cdots+d_j} \subset S_{(d_{i-1}+\cdots+d_j)A_\ell}$. Hence, if the sets $S^{(\ell)}$ and $R^{d_{i-1}+\cdots+d_j}S$ meet, then $S^{(\ell)} \cap S_{1+(d_{i-1}+\cdots+d_j)A_\ell} \neq \varnothing$. By the proposition statement, this implies that either $S^{(\ell)} \cap S_k \neq \varnothing$ or $S^{(\ell)} \cap S_{A_{\ell'}} \neq \varnothing$ (for some $\ell' < \ell$). The k-product-freeness of S rules out the former and the definition of $S^{(\ell)}$ the latter. Therefore, $S^{(\ell)}$ and $R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S$ are disjoint and so,

$$d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)} \cup R^{d_{i-1}}S \cup R^{d_{i-1}+d_i}S \cup \cdots \cup R^{d_{i-1}+\cdots+d_{\rho-1}}S) \gtrsim (\rho - i + 2) \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}).$$

Finally, noting that $S^{(\ell)} \subset S$ completes the induction. The $i = 1$ case gives

$$1 \geqslant d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S \cup R^{d_1}S \cup R^{d_1+d_2}S \cup \cdots \cup R^{d_1+\cdots+d_{\rho-1}}S) \gtrsim \rho \cdot d^{I_\infty}_{C_1^{(\ell)}\mathcal{F}}(S^{(\ell)}).$$

The claim then follows from Observation 4.4. □

We next show that very little of S has not been captured by the previous claim.

Claim 4.7.2. *For all $\ell \in [m]$ and all large n, $\mu(S(n) \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})) \leqslant m\varepsilon$.*

Proof. For each ℓ, $(C_t^{(\ell)})$ is an ε-capturing steeplechase for $R^{(\ell)}$ and so, for all large n,

$$\mu(R^{(\ell)}(n) \setminus C_1^{(\ell)}\mathcal{F}) \leqslant \varepsilon.$$

Hence, it is enough to show that $S \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F}) \subset \bigcup_\ell (R^{(\ell)} \setminus C_1^{(\ell)}\mathcal{F})$. Fix $w \in S \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})$. Let ℓ' be maximal with $w \in S^{(\ell')}$ (such an ℓ' exists as $S^{(1)} = S$). Since $w \in S \setminus \bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})$, we have $w \notin C_1^{(\ell')}\mathcal{F}$. We claim that $w \in S_{A_{\ell'}}$. If $\ell' = m$, then this is immediate since $S_{A_m} = S_1 = S$. If $\ell' < m$, then, by the maximality of ℓ', we must have $w \in S_{A_{\ell'}} \cup C_1^{(\ell')}\mathcal{F}$ and so $w \in S_{A_{\ell'}}$. Thus, $w \in (S_{A_{\ell'}} \cap S^{(\ell')}) \setminus C_1^{(\ell')}\mathcal{F} = R^{(\ell')} \setminus C_1^{(\ell')}\mathcal{F}$, as required. □

We now note that $C_1^{(1)}\mathcal{F}, \ldots, C_1^{(m)}\mathcal{F}$ are pairwise disjoint. If not, then some $w_i \in C_1^{(i)}$ is a prefix of some $w_j \in C_1^{(j)}$ (for $i \neq j$). Now, by construction, $\min C_1^{(\ell)} > \max C_1^{(\ell-1)}$ for all ℓ and so $i < j$. On the other hand, $w_j \in C_1^{(j)} \subset R^{(j)} \subset S^{(j)}$ and so $w_j \notin C_1^{(i)}\mathcal{F}$, a contradiction. In particular, $C_1^{(1)}, \ldots, C_1^{(m)}$ are pairwise disjoint and their union is prefix-free. Since they are all subsets of $S \subset w\mathcal{F}$, we have $\sum_\ell \mu(C_1^{(\ell)}) \leqslant \mu(w)$.

We can now show that $d_{w\mathcal{F}}^{I_\infty}(S) \leqslant 1/\rho$. Summing Claim 4.7.1 over ℓ and then applying Claim 4.7.2 gives

$$\mu(w) \gtrsim \rho \cdot d^{I_\infty}\left(\bigcup_\ell (S^{(\ell)} \cap C_1^{(\ell)}\mathcal{F})\right) \gtrsim \rho \cdot d^{I_\infty}(S).$$

By the definition of $d_{w\mathcal{F}}^{I}$ and the fact that $S \subset w\mathcal{F}$, we have $d_{w\mathcal{F}}^{I_\infty}(S) = \mu(w)^{-1} d^{I_\infty}(S) \lesssim 1/\rho$. Since ε is arbitrarily small, we obtain $d_{w\mathcal{F}}^{I_\infty}(S) \leqslant 1/\rho$, as required. □

We now show that there is always a sequence of sets satisfying Proposition 4.7. The sequence chosen here is motivated by the proofs of Propositions 3.9 and 3.10.

Proposition 4.8. *For every integer $k \geqslant 2$ there is a sequence $A_1, \ldots, A_m$ satisfying the hypothesis of Proposition 4.7.*

Proof. We deal with the cases $k = 2, 3, 5, 7, 13$ first.

- $k = 2$: take $A_1 = \{1\}$ (with $d_1 = 1$),
- $k = 3$: take $A_1 = \{1, 2\}$ (with $d_1 = d_2 = 1$) and $A_2 = \{1\}$ (with $d_1 = d_2 = 1$),
- $k = 5$: take $A_1 = \{1, 3\}$ (with $d_1 = d_2 = 2$) and $A_2 = \{1\}$ (with $d_1 = d_2 = 2$),
- $k = 7$: take

$$\begin{aligned} A_1 &= \{1,3\} && (\text{with } d_1 = d_2 = d_3 = 2), \\ A_2 &= \{1,2\} && (\text{with } d_1 = d_2 = d_3 = 1), \\ A_3 &= \{1,4\} && (\text{with } d_1 = d_2 = d_3 = 1), \text{ and} \\ A_4 &= \{1\} && (\text{with } d_1 = d_2 = d_3 = 1), \end{aligned}$$

- $k = 13$: take

$$\begin{aligned} A_1 &= \{1,4\} && (\text{with } d_1 = d_2 = d_3 = d_4 = 3), \\ A_2 &= \{1,2\} && (\text{with } d_1 = d_2 = d_3 = d_4 = 3), \\ A_3 &= \{1,3,5,7\} && (\text{with } d_1 = d_2 = d_3 = d_4 = 1), \\ A_4 &= \{1,3\} && (\text{with } d_1 = d_2 = d_3 = d_4 = 1), \\ A_5 &= \{1,5\} && (\text{with } d_1 = d_2 = d_3 = d_4 = 1), \text{ and} \\ A_6 &= \{1\} && (\text{with } d_1 = d_2 = d_3 = d_4 = 1). \end{aligned}$$

We now turn to $k \notin \{2, 3, 5, 7, 13\}$. For positive integers d and t, let $B_{d,t} := \{1, d+1, 2d+1, \ldots, td+1\}$. We construct $A_1, A_2, \ldots$ by taking all the sets $B_{d,t}$ for $1 \leqslant d \leqslant \rho - 1$ and $1 \leqslant t < (k-1)/d$ in the order of decreasing d and then decreasing t, and add the set $\{1\}$ to the end.

Consider a set $A_\ell = B_{d,t}$. We need to show that A_ℓ satisfies Proposition 4.7. Let $s \in \mathbb{Z}^+$ be maximal such that $ds \leqslant \rho - 1$, and $\alpha \in \mathbb{Z}^+$ be minimal such that $\alpha d(d+1) \geqslant k-1$. For $1 \leqslant i \leqslant \rho - 1$, define

$$d_i = \begin{cases} \alpha d & \text{if } i \equiv 0 \bmod s+1 \text{ and } \alpha \neq 2, \\ d & \text{otherwise.} \end{cases}$$

For any $1 \leqslant i \leqslant j \leqslant \rho - 1$, we have that $d_i + d_{i+1} + \cdots + d_j = \beta d$ for some integer β satisfying $1 \leqslant \beta \leqslant \rho - 1 + (\alpha - 1)(\rho - 1)/(s+1)$. Moreover, by definition of the d_i, either $\beta \geqslant \alpha$ or $\beta \leqslant s$. Note that

$$1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell = \{1 + \beta d, 1 + (\beta+1)d, \ldots, 1 + \beta d(td+1)\}.$$

If $\beta = 1$, then $B_{d,t+1} \subset \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$. Now, either $k \in B_{d,t+1}$ or $B_{d,t+1} = A_{\ell'}$ for some $\ell' < \ell$ and so A_ℓ satisfies Proposition 4.7.

If $1 < \beta \leqslant s$, then $B_{\beta d,1} \subset \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$. Since $d < \beta d \leqslant ds \leqslant \rho - 1$, it holds that $B_{\beta d,1} = A_{\ell'}$ for some $\ell' < \ell$ and so A_ℓ satisfies Proposition 4.7.

If $\beta \geqslant \alpha$, we claim that $k \in 1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell$. Since $k-1$ is a multiple of d, it suffices to show that $\beta d \leqslant k - 1 \leqslant \beta d(td+1)$. Firstly,

$$\beta d(td+1) \geqslant \alpha d(d+1) \geqslant k-1.$$

For the second inequality, if $d \leqslant \rho - 2$, it holds that $\alpha d(d+1) = k-1$ as observed in the proof of Proposition 3.10. Furthermore, we have

$$\beta d \leqslant (\rho + (\alpha-1)\rho/(s+1))d \leqslant (\rho + (\alpha-1)d)d.$$

where the second inequality follows from $d(s+1) \geqslant \rho$. This is less than $k-1 = \alpha d(d+1)$ if $d(\alpha-1) + \rho \leqslant \alpha(d+1)$, which is true for $k \notin \{2,3,5,7,13\}$ as shown in the proof of Proposition 3.10. On the other hand, if $d = \rho - 1$, we have

$$\beta d \leqslant (d + (\alpha-1)d/(s+1))d \leqslant ((\alpha+1)/2)d^2 \leqslant ((\alpha+1)/2)d(d+1).$$

If $\alpha \geqslant 3$, this is at most $(\alpha-1)d(d+1) < k-1$ as required. If $\alpha \leqslant 2$, we can observe that $\beta \leqslant \rho - 1$ to obtain $\beta d \leqslant \rho(\rho-1) \leqslant k-1$ where the last inequality was shown in the proof of Proposition 3.9. In all cases, we have $\beta d \leqslant k-1$ as required. Hence, $k \in 1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell$, and so A_ℓ satisfies Proposition 4.7.

Finally, for $A_\ell = \{1\}$, we can simply pick $d_1 = \cdots = d_{\rho-1} = 1$. We then get that $B_{j-i+1,1} \subset \{1\} \cup (1 + (d_i + d_{i+1} + \cdots + d_j)A_\ell)$. Since $1 \leqslant j-i+1 \leqslant \rho - 1$, it holds that $B_{j-i+1,1} = A_{\ell'}$ for some $\ell' < \ell$ and so A_ℓ satisfies Proposition 4.7. □

Propositions 4.7 and 4.8 combine to give Theorem 2.6(b) and so we have indeed proved Theorem 2.6 in this section, as promised.

5 Product-free sets in the free group

We now adapt our methods to the free group and prove Theorem 1.8. Throughout, $\mathcal{G}$ denotes the free group on a finite alphabet $\mathcal{A}$ and $S \subset \mathcal{G}$ is a k-product-free set whose density we want to bound. We always assume that all words are in reduced form. Moreover, AB denotes the product of two sets $A, B \subset \mathcal{G}$ without cancellation, that is

$$AB := \{w \in \mathcal{G} : \text{ there is a substring decomposition } w = ab \text{ with } a \in A \text{ and } b \in B\}.$$

For example, if $A = \{ba^{-1}, c\}$ and $B = \{a, d\}$, then $AB = \{ba^{-1}d, ca, cd\}$ does not contain b even though $b = ba^{-1} \cdot a$, since b cannot be decomposed into two substrings where the first is in A and the second is in B. In particular, $C\mathcal{G}$ consists of all words with a prefix in C. We equip $\mathcal{G}$ with the measure μ defined as $\mu(w) = 1/|\mathcal{G}(|w|)|$. If $\mathbf{W} = \alpha_1\alpha_2\cdots$ is a random infinite word where each α_{i+1} is an independent uniformly random letter other than α_i^{-1}, then

$$\mu(w) = \mathbb{P}(\mathbf{W} \text{ hits } w) = \begin{cases} (2|\mathcal{A}|)^{-1}(2|\mathcal{A}|-1)^{-(|w|-1)} & \text{if } w \text{ is not the empty word,} \\ 1 & \text{if } w \text{ is the empty word.} \end{cases}$$

As before, for $B \subset \mathcal{G}$, $\mu(B) = \sum_{w \in B} \mu(w)$ is the expected number of times that $\mathbf{W}$ hits B. So, we can make the following observations corresponding to Observations 2.1 and 2.2.

Observation 5.1. *If $C \subset \mathcal{G}$ is prefix-free, then $\mu(C) \leqslant 1$, $\mu(C\mathcal{G}(n)) \leqslant \mu(C)$ for all n, and $\mu(C\mathcal{G}(n)) = \mu(C)$ if C is finite and $n \geqslant \max C$.*

We now define the relative density of subsets of $\mathcal{G}$ as follows.

Definition 5.2 (relative density). Let $B \subset \mathcal{G}$, and $\mathcal{H}$ be a subsemigroup of $\mathcal{G}$ with $\mathcal{H}(n) \neq \varnothing$ for all sufficiently large n. If I is an interval, then the *relative density of B in $\mathcal{H}$ on interval I* is

$$d^I_{\mathcal{H}}(B) := \frac{\mu(B(I) \cap \mathcal{H})}{\mu(\mathcal{H}(I))} = \mu(\mathcal{H}(I))^{-1} \sum_{n \in I} \mu(B(n) \cap \mathcal{H}).$$

If $\mathcal{H}(I) = \varnothing$, we will take the relative density to be 0 by convention, and $d^I(B) := d^I_{\mathcal{G}}(B)$.

The upper Banach density of B is then $d^*(B) = \limsup_{I \to \infty} d^I(B)$. At this point, we can again diagonalise to obtain a sequence (I_j) such that, for every $\mathcal{H}$ and B that we consider in our proofs, $(d^{I_j}_{\mathcal{H}}(B))$ converges to some limit $d^{I_\infty}_{\mathcal{H}}(B)$, and $d^{I_\infty}(S) = d^*(S)$. These limits are again additive. Also note that $d^I_{\mathcal{H}}(B) = d^I(B \cap \mathcal{H})/d^I(\mathcal{H})$ and so $d^{I_\infty}_{\mathcal{H}}(B) = d^{I_\infty}(B \cap \mathcal{H})/d^{I_\infty}(\mathcal{H})$ if $d^{I_\infty}(\mathcal{H}) > 0$. We define sup density as follows.

Definition 5.3 (sup density). For a set B, the *sup density of B in $\mathcal{H}$* is

$$d^{I_\infty}_{\sup \mathcal{H}}(B) := \sup_{w \in \mathcal{H}} d^{I_\infty}_{w\mathcal{G} \cap \mathcal{H}}(B).$$

From now on, let $\mathcal{H} = \mathcal{G}^{\alpha\beta} \subset \mathcal{G}$ be the subsemigroup of $\mathcal{G}$ consisting of all words starting with α and ending in β where $\alpha, \beta \in \mathcal{A} \cup \mathcal{A}^{-1}$ and $\alpha \neq \beta^{-1}$. A random sequence argument shows the following.

Observation 5.4. *Let $C \subset \mathcal{H}$ be finite and prefix-free. Then, for all $n \geqslant \max C + 2$,*

$$\mu((C\mathcal{G} \cap \mathcal{H})(n)) \geqslant \frac{\mu(C)}{(2|\mathcal{A}|-1)^2}.$$

In particular, $d^{I_\infty}(w\mathcal{G} \cap \mathcal{H}) > 0$ for all $w \in \mathcal{H}$. As in Lemma 2.4, subtree densities of G satisfy the property that we may strip away prefixes.

Lemma 5.5. *If $w, v \in \mathcal{H}$, then $d^{I_\infty}_{wv\mathcal{G}\cap\mathcal{H}}(wB) = d^{I_\infty}_{v\mathcal{G}\cap\mathcal{H}}(B)$.*

Proof. For $u \in \mathcal{H}$, note that $\mu(wu) = a \cdot \mu(u)$ where $a = (2|A| - 1)^{-|w|}$. So, if $X \subset \mathcal{H}$ is finite, then $\mu(wX) = a \cdot \mu(X)$. Let I be any interval with $\min I > |wv|$. Then

$$d^{I}_{wv\mathcal{G}\cap\mathcal{H}}(wB) = \frac{\mu((wB)(I) \cap wv\mathcal{G})}{\mu((wv\mathcal{G} \cap \mathcal{H})(I))} = \frac{\mu(B(I - |w|) \cap v\mathcal{G})}{\mu((v\mathcal{G} \cap \mathcal{H})(I - |w|))}.$$

where we used that $wv\mathcal{G} \cap \mathcal{H} = w(v\mathcal{G} \cap \mathcal{H})$. For any $X \subset \mathcal{H}$, the fact that $\mu(X(n)) \in [0, 1]$ implies that

$$|\mu(X(I)) - \mu(X(I - |w|))| = \left| \sum_{n \in I} \mu(X(n)) - \sum_{n \in I - |w|} \mu(X(n)) \right| \leqslant |w|.$$

Therefore,

$$\frac{\mu(B(I) \cap v\mathcal{G}) - |w|}{\mu((v\mathcal{G} \cap \mathcal{H})(I)) + |w|} \leqslant d^{I}_{wv\mathcal{G}\cap\mathcal{H}}(wB) \leqslant \frac{\mu(B(I) \cap v\mathcal{G}) + |w|}{\mu((v\mathcal{G} \cap \mathcal{H})(I)) - |w|}.$$

Set $I = I_j$ and take j to infinity. From $d^{I_\infty}(v\mathcal{G} \cap \mathcal{H}) > 0$ it follows $\mu((v\mathcal{G} \cap \mathcal{H})(I_j)) \to \infty$. Hence, both bounds above tend to $d^{I_\infty}_{v\mathcal{G}\cap\mathcal{H}}(B)$ and so $d^{I_\infty}_{wv\mathcal{G}\cap\mathcal{H}}(wB) = d^{I_\infty}_{v\mathcal{G}\cap\mathcal{H}}(B)$. □

We can also obtain the following analogues to Observations 4.4 and 4.5.

Observation 5.6. *Let $C \subset \mathcal{H}$ be finite and prefix-free. Then*

$$d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) = \sum_{c \in C} \frac{d^{I_\infty}(c\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(C\mathcal{G} \cap \mathcal{H})} \cdot d^{I_\infty}_{c\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B \cap C\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(C\mathcal{G} \cap \mathcal{H})}.$$

Proof. Because $d^{I_\infty}(c\mathcal{G} \cap \mathcal{H}) > 0$ for all $c \in C$, and therefore also $d^{I_\infty}(C\mathcal{G} \cap \mathcal{H}) > 0$, it holds that

$$d^{I_\infty}_{c\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B \cap c\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(c\mathcal{G} \cap \mathcal{H})} \quad \text{and} \quad d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B \cap C\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(C\mathcal{G} \cap \mathcal{H})}.$$

Since d^{I_∞} is additive, this implies that

$$d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) = \frac{d^{I_\infty}(B \cap C\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(C\mathcal{G} \cap \mathcal{H})} = \sum_{c \in C} \frac{d^{I_\infty}(B \cap c\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(C\mathcal{G} \cap \mathcal{H})} = \sum_{c \in C} \frac{d^{I_\infty}(c\mathcal{G} \cap \mathcal{H})}{d^{I_\infty}(C\mathcal{G} \cap \mathcal{H})} \cdot d^{I_\infty}_{c\mathcal{G}\cap\mathcal{H}}(B).$$

□

Observation 5.7. *The first equality of Observation 5.6 implies that $d^{I_\infty}_{C\mathcal{G}\cap\mathcal{H}}(B) \leqslant d^{I_\infty}_{\sup \mathcal{H}}(B)$.*

Steeplechases in the free group can be defined exactly as for the free semigroup.

Definition 5.8 (steeplechase)**.** An infinite sequence (C_t) of subsets of $\mathcal{H}$ is a *steeplechase* if, for each positive integer t,

- each C_t is prefix-free and finite,

- every word in C_{t+1} has a proper prefix in C_t (in particular, $C_{t+1}\mathcal{G} \subset C_t\mathcal{G}$).

Steeplechase (C_t) is *spread* if $\max C_t < \min C_{t+1}$ for all t and is ε-*tight* if, for all m, n, $|\mu(C_m) - \mu(C_n)| \leqslant \varepsilon$.

The following lemma is an analogue to Lemma 4.2.

Lemma 5.9. *Let $\varepsilon > 0$ and $B \subset \mathcal{H}$. There is an ε-tight spread steeplechase (C_t) such that*

- $C_1 \cup C_2 \cup \cdots \subset B$,
- *for all t and all large n (in terms of t), $\mu((B \setminus C_t\mathcal{G})(n)) \leqslant \varepsilon$,*
- *for all t, $\mu(C_t)/d^{I_\infty}(\mathcal{H}) \geqslant d^{I_\infty}_{\mathcal{H}}(B) - \varepsilon$.*

Proof. This is very similar to the proof of Lemma 4.2. Recall that the headcount, $h(x)$, of $x \in B$ is the number $b \in B$ that are prefixes of x. For each positive integer t, let $D_t = \{x \in B \colon h(x) = t\}$. Iteratively do the following for each positive integer t.

1. Let ℓ_t be such that $\mu(D_t(\{\ell_t + 1, \ell_t + 2, \ldots\})) \leqslant \varepsilon \cdot d^{I_\infty}(\mathcal{H})/2^t$.
2. Let $C_t = D_t(\{1, 2, \ldots, \ell_t\})$.
3. Remove $(D_t \setminus C_t)\mathcal{G}$ from B (including from all later D_i).

Then (C_t) is a steeplechase and $C_1 \cup C_2 \cup \cdots \subset B$. Fix t and let $n > \max\{\ell_1, \ldots, \ell_t\}$. Then,

$$\mu((B \setminus C_t\mathcal{G})(n)) \leqslant \sum_r \mu(((D_r \setminus C_r)\mathcal{G})(n)) \leqslant \sum_r \mu(D_r \setminus C_r) \leqslant \varepsilon \cdot d^{I_\infty}(\mathcal{H}) \leqslant \varepsilon.$$

Finally, this implies that $\mu(B(n)) \leqslant \mu(C_t\mathcal{G}(n)) + \varepsilon \cdot d^{I_\infty}(\mathcal{H}) = \mu(C_t) + \varepsilon \cdot d^{I_\infty}(\mathcal{H})$. Averaging this over $n \in I_j$ and taking $j \to \infty$ gives $d^{I_\infty}(B) \leqslant \mu(C_t) + \varepsilon \cdot d^{I_\infty}(\mathcal{H})$ and therefore $d^{I_\infty}_{\mathcal{H}}(B) = d^{I_\infty}(B)/d^{I_\infty}(\mathcal{H}) \leqslant \mu(C_t)/d^{I_\infty}(\mathcal{H}) + \varepsilon$. By passing to a subsequence, we may assume that (C_t) is spread and ε-tight. □

We call a steeplechase (C_t) given by Lemma 5.9 an *ε-capturing steeplechase for B*. There is also an analogue to Lemma 4.3.

Lemma 5.10. *Let (C_t) be an ε-tight spread steeplechase. For every $w \in \mathcal{H}$ there is an N such that the following holds. If $C = C_1 \cup C_2 \cup \cdots \cup C_N$ and $n \geqslant \max C_N + |w|$, then*

$$\mu(((C_1\mathcal{G} \cap \mathcal{H}) \setminus (Cw\mathcal{G} \cap \mathcal{H}))(n)) \leqslant 2\varepsilon.$$

Proof. Note that $(C_1\mathcal{G} \cap \mathcal{H}) \setminus (Cw\mathcal{G} \cap \mathcal{H}) = (C_1\mathcal{G} \setminus Cw\mathcal{G}) \cap \mathcal{H} \subset C_1\mathcal{G} \setminus Cw\mathcal{G}$, and so it suffices to show that $\mu((C_1\mathcal{G} \setminus Cw\mathcal{G})(n)) \leqslant 2\varepsilon$.

We can show this exactly as in the proof of Lemma 4.3, we only need $\mathbf{W}$ to be the random walk from the beginning of this section. As a consequence, if $\mathbf{W}$ hits C_i, the probability that the next $|w|$ letters of $\mathbf{W}$ spell w is $(2|\mathcal{A}| - 1)^{-|w|}$. Importantly, this uses the fact that the last letter of a word in C_i is β and the first letter of w is α, and $\alpha \neq \beta^{-1}$. □

Now we can prove the key technical lemma for our density results, corresponding to Lemma 4.6.

Lemma 5.11. *Let $\varepsilon > 0$ and let $A, B \subset \mathcal{H}$. If (C_t) is an ε-capturing steeplechase for A with $\mu(C_1) \geqslant (2|\mathcal{A}|-1)^2(2\varepsilon+\varepsilon^{1/3})$, then*

$$d^{I_\infty}_{C_1\mathcal{G}\cap\mathcal{H}}(AB) \geqslant d^{I_\infty}_{\sup\mathcal{H}}(B) - 3\varepsilon^{1/3} \geqslant d^{I_\infty}_{C_1\mathcal{G}\cap\mathcal{H}}(B) - 3\varepsilon^{1/3}.$$

Proof. Observation 5.4 implies that $\mu((C_1\mathcal{G}\cap\mathcal{H})(n)) \geqslant \mu(C_1)/(2|\mathcal{A}|-1)^2 \geqslant 2\varepsilon+\varepsilon^{1/3}$ for $n \geqslant \max C_1 + 2$. We proceed as in the proof of Lemma 4.6. Let $w \in \mathcal{H}$ be such that

$$d^{I_\infty}_{w\mathcal{G}\cap\mathcal{H}}(B) \geqslant d^{I_\infty}_{\sup\mathcal{H}}(B) - \varepsilon^{1/3}.$$

Apply Lemma 5.10 to (C_t) and w to give an N such that letting $C = C_1 \cup \cdots \cup C_N$, if $n \geqslant \max C_N + |w|$, then $\mu(((C_1\mathcal{G}\cap\mathcal{H}) \setminus (Cw\mathcal{G}\cap\mathcal{H}))(n)) \leqslant 2\varepsilon$. We may greedily choose $\widetilde{C} \subset C$ such that $\widetilde{C}w$ is prefix-free and $\widetilde{C}w\mathcal{G} = Cw\mathcal{G}$. Note that

$$\begin{aligned} 2\varepsilon &\geqslant \mu(((C_1\mathcal{G}\cap\mathcal{H}) \setminus (\widetilde{C}w\mathcal{G}\cap\mathcal{H}))(n)) \\ &\geqslant \mu((C_1\mathcal{G}\cap\mathcal{H})(n)) - \mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(n)) \\ &\geqslant 2\varepsilon + \varepsilon^{1/3} - \mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(n)) \end{aligned}$$

and so $\mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(n)) \geqslant \varepsilon^{1/3}$ as well as $\mu((C_1\mathcal{G}\cap\mathcal{H})(n)) \leqslant \mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(n)) + 2\varepsilon$.

Let I be an interval with $\min I \geqslant \max C_N + |w|$, so $\mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(I)) \geqslant |I|\varepsilon^{1/3}$ and $\mu((C_1\mathcal{G}\cap\mathcal{H})(I)) \leqslant \mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(I)) + |I|2\varepsilon$. Let $X \subset \mathcal{H}$. Note that $\widetilde{C}w\mathcal{G} \subset C_1\mathcal{G}$ and so

$$d^{I}_{C_1\mathcal{G}\cap\mathcal{H}}(X) = \frac{\mu(X(I)\cap C_1\mathcal{G})}{\mu((C_1\mathcal{G}\cap\mathcal{H})(I))} \geqslant \frac{\mu(X(I)\cap\widetilde{C}w\mathcal{G})}{\mu((C_1\mathcal{G}\cap\mathcal{H})(I))} \geqslant \frac{\mu(X(I)\cap\widetilde{C}w\mathcal{G})}{\mu((\widetilde{C}w\mathcal{G}\cap\mathcal{H})(I)) + |I|2\varepsilon}.$$

Using the fact that $x/(y+|I|2\varepsilon) \geqslant x/y - |I|2\varepsilon x/y^2 \geqslant x/y - 2\varepsilon^{1/3}$ for $\varepsilon > 0$, $0 \leqslant x \leqslant |I|$, and $y \geqslant |I|\varepsilon^{1/3}$, we have

$$d^{I}_{C_1\mathcal{G}\cap\mathcal{H}}(X) \geqslant d^{I}_{\widetilde{C}w\mathcal{G}\cap\mathcal{H}}(X) - 2\varepsilon^{1/3}.$$

Setting $X = AB$, $I = I_j$, and taking j to infinity gives

$$d^{I_\infty}_{C_1\mathcal{G}\cap\mathcal{H}}(AB) \geqslant d^{I_\infty}_{\widetilde{C}w\mathcal{G}\cap\mathcal{H}}(AB) - 2\varepsilon^{1/3}. \tag{3}$$

Now,

$$\begin{aligned} d^{I_\infty}_{\widetilde{C}w\mathcal{G}\cap\mathcal{H}}(AB) &= \sum_{c\in\widetilde{C}} \frac{d^{I_\infty}(cw\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(\widetilde{C}w\mathcal{G}\cap\mathcal{H})} \cdot d^{I_\infty}_{cw\mathcal{G}\cap\mathcal{H}}(AB) \\ &\geqslant \sum_{c\in\widetilde{C}} \frac{d^{I_\infty}(cw\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(\widetilde{C}w\mathcal{G}\cap\mathcal{H})} \cdot d^{I_\infty}_{cw\mathcal{G}\cap\mathcal{H}}(cB) \\ &= \sum_{c\in\widetilde{C}} \frac{d^{I_\infty}(cw\mathcal{G}\cap\mathcal{H})}{d^{I_\infty}(\widetilde{C}w\mathcal{G}\cap\mathcal{H})} \cdot d^{I_\infty}_{w\mathcal{G}\cap\mathcal{H}}(B) \\ &= d^{I_\infty}_{w\mathcal{G}\cap\mathcal{H}}(B) \geqslant d^{I_\infty}_{\sup\mathcal{H}}(B) - \varepsilon^{1/3}, \end{aligned}$$

where the first equality used Observation 5.6, the first inequality used the fact that $c \in \widetilde{C} \subset C \subset A$, the second equality used Lemma 5.5, and the second inequality is due to the choice of w. Combining this with (3) gives the first inequality claimed, while Observation 5.7 gives the second. □

At this point, we have recovered all important technical results that we needed to bound the density of (strongly) k-product-free sets in the free semigroup. We can now simply use exactly the same arguments as in Section 4.2. We only need to replace $w\mathcal{F}$ by $w\mathcal{G} \cap \mathcal{H}$, $\mu(w)$ by $d^{I_\infty}(w\mathcal{G} \cap \mathcal{H})$, $C_i\mathcal{F}$ by $C_i\mathcal{G} \cap \mathcal{H}$, $\mu(C_i)$ by $d^{I_\infty}(C_i\mathcal{G} \cap \mathcal{H})$, $d^{I_\infty}_{\sup}$ by $d^{I_\infty}_{\sup\mathcal{H}}$, and all references by references to the corresponding results in this section to prove the following analogue of Theorem 2.6.

Theorem 5.12. *Let $k \geqslant 2$ be an integer, $\mathcal{A}$ be a finite set, $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$, and $\mathcal{H} = \mathcal{G}^{\alpha\beta}$ be the subsemigroup of $\mathcal{G}$ consisting of all words starting with α and ending with β where $\alpha, \beta \in \mathcal{A} \cup \mathcal{A}^{-1}$ and $\alpha \neq \beta^{-1}$. Let (I_j) be a sequence of intervals, with $I_j \to \infty$, on which all relevant densities converge.*

(a) *If $S \subset \mathcal{H}$ is strongly k-product-free, then $d^{I_\infty}_{\sup\mathcal{H}}(S) \leqslant 1/k$.*

(b) *If $S \subset \mathcal{H}$ is k-product-free, then $d^{I_\infty}_{\sup\mathcal{H}}(S) \leqslant 1/\rho(k)$.*

The arguments from Ortega, Rué, and Serra [ORS24] show that a density bound on (strongly) k-product-free sets in $\mathcal{G}^{\alpha\beta}$ immediately translates to a density bound in $\mathcal{G}$. Therefore, Theorems 1.3 and 1.8 are immediate corollaries of Theorem 5.12.

6 Open problems

A first natural problem left open is to determine the structure of the extremal k-product-free sets for $k \in \{3, 5, 7, 13\}$. For $k = 5, 7, 13$, we conjecture that the extremal sets are exactly as in Theorem 1.6. The extremal sets for $k = 3$ will be slightly more complicated. Indeed, while $1 + 3\mathbb{Z}_{\geqslant 0}$ and $2 + 3\mathbb{Z}_{\geqslant 0}$ are both maximal 3-sum-free subsets of the non-negative integers of density $1/3$, so are both $\{1, 2\} + 6\mathbb{Z}_{\geqslant 0}$ and $\{5, 6\} + 6\mathbb{Z}_{\geqslant 0}$ (Łuczak and Schoen [ŁS97] showed that there are no others). We conjecture the corresponding result holds for the free semigroup.

Conjecture 6.1. *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. If $S \subset \mathcal{F}$ is 3-product-free and $d^*(S) = 1/3$, then one of the following hold. Either it is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/3\mathbb{Z}$ such that S is a subset of*

$$\{w \in \mathcal{F} \colon \text{the sum of the labels of letters in } w \text{ is } 1 \bmod 3\},$$

or it is possible to label each letter of $\mathcal{A}$ with a label in $\mathbb{Z}/6\mathbb{Z}$ such that S is a subset of

$$\{w \in \mathcal{F} \colon \text{the sum of the labels of letters in } w \text{ is } 1, 2 \bmod 6\}.$$

Łuczak [Łuc95] proved that every sum-free subset of the non-negative integers with density greater than $2/5$ is a subset of the odd integers (Łuczak and Schoen [ŁS97] proved similar results for (strongly) k-sum-free sets). Such strengthenings for strongly k-product-free subsets of the free semigroup are false as the constant $1/k$ in Theorem 1.4 cannot be replaced by anything smaller. For example, let T be the set of words of length $1 \bmod k$, let x be any word, and let $p > k$ be a prime. Define

$$\begin{aligned} T' :=& \{w \in T \colon \text{neither } x \text{ nor } w \text{ is a prefix or suffix of the other}\} \\ &\cup \{xwx \colon xwx \text{ has length } 0 \bmod k \text{ and } 1 \bmod p\}. \end{aligned}$$

Then T' is strongly k-product-free, has density at least $1/k - 2 \cdot |\mathcal{A}|^{-|x|}$, and there is no labelling of the letters of $\mathcal{A}$ by $\mathbb{Z}/k\mathbb{Z}$ in which every word of T' has sum 1 mod k (in fact, a set of positive density would need to be removed before this happens). Nonetheless, T' is a small perturbation from the set T. Hence, it is natural to ask whether there is some form of stability. Here is the conjecture for $k = 2$.

Conjecture 6.2. *Let $\mathcal{A}$ be a finite set and $\mathcal{F}$ be the free semigroup on alphabet $\mathcal{A}$. For each $\delta > 0$, there is some $\varepsilon > 0$ such that if $S \subset \mathcal{F}$ is product-free and $d^*(S) > 1/2 - \varepsilon$, then there exists an odd-occurrence set $\mathcal{O}_\Gamma$ such that $d^*(S \setminus \mathcal{O}_\Gamma) < \delta$.*

We do not have examples corresponding to T' for k-product-free sets. It would be interesting to determine whether such examples exist or if Theorem 1.6 can be strengthened to densities below $1/\rho(k)$.

Theorems 1.4 and 1.6 give the structure of extremal (strongly) k-product-free sets in the free semigroup. In the free group, Theorems 1.3 and 1.8 solve the density problem, but the structure problem remains open. The simplest open case is the following.

Conjecture 6.3. *Let $\mathcal{A}$ be a finite set and $\mathcal{G}$ be the free group on alphabet $\mathcal{A}$. If $S \subset \mathcal{G}$ is product-free and $d^*(S) = 1/2$, then the following holds. It is possible to label each letter of $\mathcal{A} \cup \mathcal{A}^{-1}$ with a label in $\mathbb{Z}/2\mathbb{Z}$ such that the label of α^{-1} is the negation of the label of α for all $\alpha \in \mathcal{A}$ and S is a subset of*

$$T := \{w \in \mathcal{G} : \text{the sum of the labels of letters in } w \text{ is } 1 \bmod 2\}.$$

For strongly k-product-free we expect the above conjecture to hold with 2 replaced by k. For k-product-free we expect the behaviour to be the same as for the free semigroup.

We remark that our methods do give some structure. Similar arguments to Section 3.1 show there is a labelling of all words in the subsemigroup $\mathcal{G}^{\alpha\beta}$ (defined in Section 5) such that the label of a concatenation is the sum of the individual labels and all words in $S \cap \mathcal{G}^{\alpha\beta}$ have label 1 mod k. What is missing is an understanding of how the labellings interact when letters cancel during concatenation.

Acknowledgements. We thank the anonymous referees for their careful reviewing and helpful suggestions. A question by one of the referees inspired us to prove Corollary 2.7.

References

[BPR16] R. Balasubramanian, Gyan Prakash, and D. S. Ramana (Nov. 2016). Sum-free subsets of finite abelian groups of type III. *European Journal of Combinatorics* **58**, 181–202. 1

[BS85] László Babai and Vera T. Sós (Jun. 1985). Sidon sets in groups and induced subgraphs of Cayley graphs. *European Journal of Combinatorics* **6**(2), 101–114. 1

[CE96] Neil J. Calkin and Paul Erdős (Jul. 1996). On a class of aperiodic sum-free sets. *Mathematical Proceedings of the Cambridge Philosophical Society* **120**(1), 1–5. 2, 3

[Ebe16] Sean Eberhard (Feb. 2016). Product mixing in the alternating group. *Discrete Analysis* Paper No. 2, 19. 1

[Gow08] W. T. Gowers (May 2008). Quasirandom groups. *Combinatorics, Probability and Computing* **17**(3), 363–387. 1

[GR05] Ben Green and Imre Z. Ruzsa (Dec. 2005). Sum-free sets in abelian groups. *Israel Journal of Mathematics* **147**, 157–188. 1

[Ked09] Kiran S. Kedlaya (2009). Product-free subsets of groups, then and now. *Communicating mathematics*, *Contemporary Mathematics*, vol. 479, 169–177. 2

[KLM24] Peter Keevash, Noam Lifshitz, and Dor Minzer (Sep. 2024). On the largest product-free subsets of the alternating groups. *Inventiones Mathematicae* **237**, 1329–1375. 2

[Lev03] Vsevolod F. Lev (Feb. 2003). Sharp estimates for the number of sum-free sets. *Journal für die Reine und Angewandte Mathematik* **555**, 1–25. 14

[LLNW20] Imre Leader, Shoham Letzter, Bhargav Narayanan, and Mark Walters (Jan. 2020). Product-free sets in the free semigroup. *European Journal of Combinatorics* **83**, 103003. 2

[ŁS97] Tomasz Łuczak and Tomasz Schoen (Oct. 1997). On infinite sum-free sets of natural numbers. *Journal of Number Theory* **66**(2), 211–224. 2, 3, 6, 32

[ŁS01] Tomasz Łuczak and Tomasz Schoen (Oct. 2001). Sum-free subsets of right cancellative semigroups. *European Journal of Combinatorics* **22**(7), 999–1002. 6

[Łuc95] Tomasz Łuczak (May 1995). A note on the density of sum-free sets. *Journal of Combinatorial Theory. Series A* **70**(2), 334–336. 6, 32

[ORS24] Miquel Ortega, Juanjo Rué, and Oriol Serra (Jul. 2024). Product-free sets in the free group. *Mathematika* **70**(3), e12255. 2, 3, 32

[TV17] Terence Tao and Van Vu (Jun. 2017). Sum-free sets in groups: a survey. *Journal of Combinatorics* **8**(3), 541–552. 2