WILF EQUIVALENCES AND STANLEY-WILF LIMITS FOR PATTERNS IN ROOTED LABELED FORESTS

MICHAEL REN

ABSTRACT. Building off recent work of Garg and Peng, we continue the investigation into classical and consecutive pattern avoidance in rooted forests, resolving some of their conjectures and questions and proving generalizations whenever possible. Through extensions of the forest Simion-Schmidt bijection introduced by Anders and Archer, we demonstrate a new family of forest-Wilf equivalences, completing the classification of forest-Wilf equivalence classes for sets consisting of a pattern of length 3 and a pattern of length at most 5. We also find a new family of nontrivial c-forest-Wilf equivalences between single patterns using the forest analogue of the Goulden-Jackson cluster method, showing that a $(1 - o(1))^n$-fraction of patterns of length n satisfy a nontrivial c-forest-Wilf equivalence and that there are c-forest-Wilf equivalence classes of patterns of length n of exponential size. Additionally, we consider a forest analogue of super-strong-c-Wilf equivalence, introduced for permutations by Dwyer and Elizalde, showing that super-strong-c-forest-Wilf equivalences are trivial by enumerating linear extensions of forest cluster posets. Finally, we prove a forest analogue of the Stanley-Wilf conjecture for avoiding a single pattern as well as certain other sets of patterns. Our techniques are analytic, easily generalizing to different types of pattern avoidance and allowing for computations of convergent lower bounds of the forest Stanley-Wilf limit in the cases covered by our result. We end with several open questions and directions for future research, including some on the limit distributions of certain statistics of pattern-avoiding forests.

CONTENTS

1. Introduction

A sequence of distinct integers is said to *avoid* a permutation, or pattern, $\pi = \pi(1) \cdots \pi(k)$ of $[k] = \{1, \ldots, k\}$ if it contains no subsequence that is in the same relative order as π. The study of pattern avoidance in permutations of $[n]$ was started in 1968 by Knuth in [24], where stack sorting was linked to permutations avoiding the pattern 231. Since then, pattern avoidance has blossomed into a very active area of research, with many connections made to classical and contemporary results in enumerative and algebraic combinatorics [21]. Different variants of permutation pattern avoidance, for example avoidance of consecutive patterns [8] and of generalized patterns [32], have also been extensively studied, along with notions of pattern avoidance in other combinatorial structures such as binary trees [29] and posets [20].

The variant of pattern avoidance that we investigate is in rooted labeled forests, a notion introduced in 2018 by Anders and Archer in [1]. Here, we consider unordered (non-planar) rooted forests on n vertices such that each vertex has a different label in $[n]$, which we call *rooted forests* on $[n]$. Such a forest is then said to *avoid* a pattern π if the sequence of labels from the root to any leaf avoids π in the sense described in the previous paragraph. As a special case, this includes the aforementioned case of permutation pattern avoidance when the forest is taken to be a path. Anders and Archer find the number $f_n(S)$ of forests avoiding a set S of patterns in [1] for certain sets S. They also study *forest-Wilf equivalence*, the phenomenon when $f_n(S) = f_n(S')$ for different sets S and S' of patterns and all $n \in \mathbb{N}$. Their work was continued by Garg and Peng in [17] where the authors posed several open questions, some of which we resolve in this paper.

A classical result in enumerating pattern-avoiding permutations is that for any pattern π of length 3, the number $|\mathrm{Av}_\pi(n)|$ of permutations of $[n]$ avoiding π is the nth Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$, shown for example through the *Simion-Schmidt bijection* in [30]. The independence of this count from the specific pattern being avoided launched the study of *Wilf equivalence*, the phenomenon when $|\mathrm{Av}_S(n)| = |\mathrm{Av}_{S'}(n)|$ for different sets S and S' of patterns. Work of Stankova and West [31] along with generalizations of the Simion-Schmidt bijection by Backelin, West, and Xin [4] show the existence of infinite families of Wilf equivalences that explain nearly all known Wilf equivalences between single patterns. Using a forest variant of the Simion-Schmidt bijection, Anders and Archer proved the forest-Wilf equivalence of 123 and 132 in [1], which was generalized by Garg and Peng in [17] through the following theorem.

Theorem 1.1 ([17, Theorem 1.3])**.** *Suppose that the patterns $\pi_1, \ldots, \pi_m$ satisfy the property that $\pi_i(k_i - 1) = k_i - 1$ and $\pi_i(k_i) = k_i$ where $\pi_i = \pi_i(1) \cdots \pi_i(k_i)$. Then $\{\pi_1, \ldots, \pi_m\}$ and $\{\widetilde{\pi}_1, \ldots, \widetilde{\pi}_m\}$ are forest-Wilf equivalent.*

Here, we denote by $\widetilde{\pi}$ the *twist* of the pattern $\pi = \pi(1) \cdots \pi(k)$, which we define to be the resulting pattern from switching the last two elements so that $\widetilde{\pi} = \pi(1) \cdots \pi(k-2)\pi(k)\pi(k-1)$. Through a computer calculation done by Garg and Peng, it can be verified that Theorem 1.1 explains all nontrivial single pattern forest-Wilf equivalences between patterns of length at most 5 [18]. Here, a *trivial* forest-Wilf equivalence is one between the sets $S = \{\pi_1, \ldots, \pi_k\}$ and $\overline{S} = \{\overline{\pi}_1, \ldots, \overline{\pi}_k\}$ where $\overline{\pi} = k + 1 - \pi(1), \ldots, k + 1 - \pi(k)$ is the *complement* of $\pi = \pi(1) \cdots \pi(k)$. Such an equivalence is trivial because reversing the order of the labels bijects forests avoiding S with forests avoiding $\overline{S}$. However, Theorem 1.1 does not explain the following forest-Wilf equivalences conjectured by Garg and Peng in [17].

Conjecture 1.2 ([17, Conjecture 7.1])**.** *The following pairs of sets are forest-Wilf equivalent:*

(i) $\{123, 2413\}$ *and* $\{132, 2314\}$*;*
(ii) $\{123, 3142\}$ *and* $\{132, 3124\}$*;*
(iii) $\{213, 4123\}$ *and* $\{213, 4132\}$*.*

By restricting the forest Simion-Schmidt bijection, we prove parts (i) and (ii) of this conjecture in Section 3. We also prove the following generalization of part (iii) by recursively applying the forest Simion-Schmidt bijection. Along with Theorem 1.1 and parts (i) and (ii) of Conjecture 1.2, Theorem 1.3 explains all nontrivial forest-Wilf equivalences between sets consisting of a pattern of length 3 and a pattern of length $k \leq 5$, again by a computer calculation of Garg and Peng [18].

Theorem 1.3. *If* $\pi(k) = \pi(k-1)+1 = \pi(k-2)+2$ *in the pattern* $\pi = \pi(1)\cdots\pi(k)$*, then* $\{213, \pi\}$ *and* $\{213, \widetilde{\pi}\}$ *are forest-Wilf equivalent.*

Garg and Peng also considered *consecutive* pattern avoidance in forests. A sequence of distinct integers *consecutively avoids* a pattern π if it contains no consecutive subsequence that is in the same relative order as π. The analogues of Wilf equivalence and forest-Wilf equivalence, which we call *c-Wilf equivalence* and *c-forest-Wilf equivalence*, are defined in the same way using the consecutive notion of pattern avoidance instead.

The systematic study of consecutive pattern avoidance in permutations was started in 2003 by Elizalde and Noy in [10], where the number of permutations of $[n]$ avoiding π was enumerated for short patterns π using analytic techniques. Elizalde and Noy further developed these techniques in [11], where they applied the Goulden-Jackson cluster method from [16] to asymptotically enumerate more classes of permutations that avoid consecutive patterns. Garg and Peng adapted the cluster method to forests that avoid consecutive patterns in [17], proving the following intriguing result.

Theorem 1.4 ([17, Theorem 1.4])**.** *The patterns* 1324 *and* 1423 *are c-forest-Wilf equivalent but not c-Wilf equivalent.*

In contrast, for the classical case all known forest-Wilf equivalences are also Wilf equivalences, and Garg and Peng even conjectured in [17, Conjecture 7.3] that forest-Wilf equivalence implies Wilf equivalence. Garg and Peng also asked whether there exist more nontrivial single pattern c-forest-Wilf equivalences after checking all patterns of length at most 5 with a computer and finding no more [18]. We answer their question in the affirmative by exhibiting a rich infinite family of c-forest-Wilf equivalences. Not only are there more nontrivial c-forest-Wilf equivalences, but a relatively large proportion of patterns satisfy such an equivalence. Furthermore, there are exponentially large c-forest-Wilf equivalence classes. In the process of our proof, we somewhat elucidate what properties of forests allow for such c-forest-Wilf equivalences between patterns that are not c-Wilf equivalent as in Theorem 1.4. Our construction is rather long and is detailed in Section 4, so for now we summarize its key properties in the following theorems.

Theorem 1.5. *For all* n*, there exists a c-forest-Wilf equivalence class containing at least* 2^{n-4} *patterns of length* n*.*

Theorem 1.6. *For any constant* $c < 1$*, there exists an integer* n_c *such that for all* $n > n_c$*, at least* $c^n n!$ *patterns of length* n *satisfy a nontrivial c-forest-Wilf equivalence.*

In 2018, Dwyer and Elizalde introduced in [6] the notion of *super-strong c-Wilf equivalence*, refining c-Wilf equivalence for permutations. Roughly speaking, two patterns π and π' of length k are said to be super-strongly c-Wilf equivalent if for all n and $S \subseteq [n-k+1]$, the number of permutations of $[n]$ whose occurrences of π are exactly indexed by S is equal to the number of permutations of $[n]$ whose occurrences of π' are exactly indexed by S. For example, in the permutation 124365, the occurrences of the pattern 132 are indexed by the set $\{2, 4\}$, corresponding to the substrings 243 and 365 that begin at indices 2 and 4. Dwyer and Elizalde show that among a special class of permutations, the *nonoverlapping permutations*, c-Wilf equivalence is equivalent to super-strong c-Wilf equivalence. By the work of Lee and Sah in 2018 in [25], c-Wilf equivalence among nonoverlapping patterns is completely understood and determined by the first and last element. Bna has shown that an asymptotically constant proportion of permutations of $[n]$ are nonoverlapping in [5], yielding a large number of nontrivial super-strong c-Wilf equivalences. By adapting the techniques for enumerating linear extensions of cluster posets from [6, 25], we show that the complete opposite is true for the forest analogue.

Theorem 1.7. *Any super-strong c-forest-Wilf equivalence between two patterns is trivial.*

Finally, we consider the asymptotics of the number of forests on $[n]$ avoiding certain sets S of patterns. The asymptotics of permutations that classically avoid a pattern π is governed by the *Stanley-Wilf conjecture*, which states that $\lim_{n\to\infty} |\mathrm{Av}_n(\pi)|^{1/n}$, the *Stanley-Wilf limit*, exists and is finite for all patterns π. In 1999, Arratia proved in [2] that the limit exists through a supermultiplicativity argument and in 2000 Klazar proved in [22] that the finiteness of the limit follows from the Füredi-Hajnal conjecture, which was proven in 2004 by Marcus and Tardos in [26]. Since then, much work has been done to study the value of various Stanley-Wilf limits. In 2013, Fox disproved in [13] the widely believed conjecture that the limits are always quadratic in the pattern length, instead showing that they are generally exponential. For permutations that consecutively avoid a pattern π, a 2011 result of Ehrenborg, Kitaev, and Perry from [9] shows using spectral theoretic methods that the proportion of permutations of $[n]$ avoiding a pattern π is $c\lambda^n + O(r^n)$ for positive constants $c, \lambda > r$ only depending on π. It is then natural to ask how the asymptotics behave for pattern avoidance in forests, and Garg and Peng made the following forest analogue of the Stanley-Wilf conjecture in [17] with respect to classical avoidance.

Conjecture 1.8 ([17, Conjecture 7.2])**.** *For any set S of patterns, let $f_n(S)$ and $t_n(S)$ denote the number of rooted forests and trees on $[n]$ avoiding S, respectively. Then*

$$\lim_{n\to\infty} \frac{f_n(S)^{1/n}}{n} \text{ and } \lim_{n\to\infty} \frac{t_n(S)^{1/n}}{n}$$

exist and are equal.

Here, a rooted tree on $[n]$ is just a connected rooted forest on $[n]$. Notably, the finiteness of the limit immediately follows from Cayley's formula: the number of rooted labeled forests and trees are $(n+1)^{n-1}$ and n^{n-1}, respectively, so the limit is automatically bounded above by 1. We resolve this conjecture in the positive for a large class of sets which includes all singleton sets.

Theorem 1.9. *For any set S of patterns in which no $\pi \in S$ begins with 1 or in which no $\pi \in S$ begins with its largest element,*

$$\lim_{n\to\infty} \frac{f_n(S)^{1/n}}{n} \text{ and } \lim_{n\to\infty} \frac{t_n(S)^{1/n}}{n}$$

exist and are equal.

Our methods are quite different from those previously used to prove the analogous results for permutations and relies on analytically interpreting the relationship between t_n and f_n, i.e. the forest structure. The proof is quite robust and immediately generalizes to sets S of consecutive or generalized patterns in which the condition in the theorem statement is satisfied. The key use of the pattern avoidance condition is to establish the inequality $t_{n+1} \geq f_n$, after which the rest of the proof is purely analytic. For this reason, we believe that our methods may be applicable to the asymptotic enumeration of other classes of labeled forests, perhaps unrelated to pattern avoidance. Additionally, our proof allows us to compute convergent lower bounds for the forest Stanley-Wilf limits, and we do so for several patterns in Section 5 with the help of formulas and recurrences shown in [1, 17].

The rest of the paper is organized as follows. In Section 2, we record all of the preliminary definitions that are necessary for the rest of the paper. In Section 3, we consider classical forest-Wilf equivalences, proving Conjecture 1.2 and Theorem 1.3. In Section 4, we examine c-forest-Wilf equivalences, first defining the *grounded permutations* that form the basis of our proofs to Theorems 1.5 and 1.6. We then find constraints on strong c-Wilf equivalences between grounded permutations along the lines of [17, Theorem 1.5] and prove Theorem 1.7, both by enumerating clusters. In Section 5, we discuss the asymptotics of pattern-avoiding forests and give the proof of Theorem 1.9 along with computed lower bounds for forest Stanley-Wilf limits. We show that the limit, when it exists, is always in $\{0\} \cup [e^{-1}, 1]$ and classify the sets that achieve 0 and the sets covered by Theorem 1.9 that achieve e^{-1}. In Section 6, we pose questions, conjectures, and potential future directions of research, including some on various limiting statistics of pattern-avoiding forests.

2. Definitions and Notations

We begin by defining all of the notions of pattern avoidance and rooted forests that we will use throughout this paper.

Definition 2.1. A *rooted labeled forest* on a set S of integers is a forest on $|S|$ vertices labeled with the elements of S in which every connected component has a distinguished root vertex. Each component then has the structure of an unordered rooted tree, and each vertex has a unique label in S. Here, by unordered we mean that the tree only has the structure of a graph with a distinguished vertex, and the order in which the neighbors of any vertex are drawn is irrelevant. In a rooted forest F on S, we let $L_F(v)$ denote the label of vertex v and suppress the subscript if it is clear from context.

The same definition can be made for rooted labeled trees. For the sake of brevity, we will oftentimes refer to rooted labeled trees and forests as trees and forests, respectively, and we will always specify when we refer to other types of trees or forests.

We will make use of some standard terminology for rooted trees and forests. In a rooted tree, the *root* is the distinguished vertex. For each non-root vertex v, the *parent* of v is the vertex directly before v in the path from the root to v, and every non-root vertex is a *child* of its parent. The *ancestors* of a vertex v are the vertices on the path from the root to v, and every vertex is a *descendant* of its ancestors. A *strict* descendant of v is a descendant of v that is not equal to v, and similarly a strict ancestor of v is an ancestor of v that is not equal to v. Each vertex v in the tree has a *depth*, defined as the number of vertices on the path from the root to v. For example, the root has depth 1. The *depth* of a rooted tree T is the maximal depth of a vertex in T. A *leaf* is a vertex with no children. A *subtree* of a rooted tree T is a connected induced subgraph of the T, which can be viewed as a rooted tree whose root is the vertex of minimal depth in T. The *subtree rooted at* v is the subtree consisting of all of the descendants of v. We may sometimes add modifiers to this term to talk about different subtrees with root v. For example, the subtree of T rooted at v consisting of the non-leaf descendants of v refers to the induced subgraph of T on the descendants of v that are not leaves of T, which can be viewed as a rooted tree with root v.

All of these terms naturally carry over to rooted forests. When we draw rooted forests, we will connect the roots of each connected component to an extra unlabeled vertex and refer to this vertex as the *root* of the forest. The root of the forest can be thought of as the parent of the roots of its connected components, though it is only drawn for visualization purposes and is not actually in the forest or counted when computing the depth of a vertex. In our drawings of rooted forests and trees, the root will be drawn at the top and each vertex will be drawn above its children. In this way, the path from a vertex to any of its descendants is a downward path, and we can refer to vertices as being *higher* or *lower* than other vertices along such a path. For example, the lowest vertex on a given path in the rooted forest is the vertex in the path of maximal depth.

We view rooted trees as trees with a distinguished vertex and rooted forests as a set of rooted trees. Thus, forests may be empty (have 0 vertices), but trees will always be nonempty. An *increasing* forest is a rooted labeled forest in which every vertex has a smaller label than its children so that the sequence of labels along any downward path in the forest is increasing. The same definition extends to trees, and we can define decreasing forests and trees analogously.

Definition 2.2. An *instance* of a pattern $\pi = \pi(1) \cdots \pi(k)$ in a rooted forest F is a sequence of vertices $v_1, \ldots, v_k$ such that v_i is an ancestor of v_{i+1} for all $1 \le i < k$ and $L(v_1), \ldots, L(v_k)$ is in the

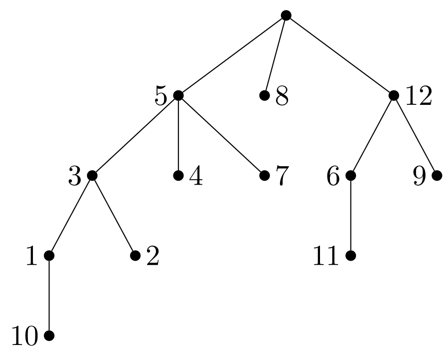

FIGURE 1. A rooted labeled forest on [12]. The root vertex is drawn but not in the forest. We generally draw forests so that the labels of the children of every vertex are sorted in increasing order.

same relative order as π. We define a *consecutive instance* in the same way, except we require that v_i is a parent of v_{i+1} instead of an ancestor so that $v_1, \ldots, v_k$ forms a downward path in the forest.

Definition 2.3. A forest F (*consecutively*) *contains* a pattern π if there exists a (consecutive) instance of π in F, and it (*consecutively*) *avoids* a set S of patterns if it does not contain any (consecutive) instance of π for all $\pi \in S$.

We will oftentimes drop the braces when referring to containing or avoiding a specific singleton set S. The word classically may be used to describe non-consecutive avoidance or containment, so to classically avoid a set is to avoid a set as in Definition 2.3. For example, the forest in Figure 1 contains 213 through the instance $5, 3, 10$, and it consecutively contains 312 through the consecutive instance $12, 6, 11$. It avoids 123 and consecutively avoids 213, but it does not classically avoid 213. A forest F avoiding a set S of patterns can be viewed as having the property that for every path from the root of F to a leaf of F, the sequence of labels avoids S in the sense of pattern avoidance for permutations and sequences. In Sections 3 and 4, we exclusively work with one of non-consecutive and consecutive avoidance, so in those sections and in general when the context is clear, we omit the word consecutive.

Definition 2.4. Two sets S and S' of patterns are (*c*-)*forest-Wilf equivalent* if for all $n \geq 0$, the number of forests on $[n]$ (consecutively) avoiding S is equal to the number of forests on $[n]$ (consecutively) avoiding S'. Two patterns π and π' are *strongly c-forest-Wilf equivalent* if for all $m, n \geq 0$, the number of forests on $[n]$ with exactly m consecutive instances of π is equal to the number of forests on $[n]$ with exactly m consecutive instances of π'.

As before, we will never work with forest-Wilf equivalence and c-forest-Wilf equivalence in the same section, so for the sake of brevity we will often refer to these concepts as equivalence when the context is clear. However, we will never omit the word strong(ly) or super-strong(ly). (We will define super-strong c-forest-Wilf equivalence in Section 4.) We will respectively denote by $S \sim S'$, $S \overset{c}{\sim} S'$, $S \overset{sc}{\sim} S'$, and $S \overset{ssc}{\sim} S'$ the forest-Wilf equivalence, c-forest-Wilf equivalence, strong c-forest-Wilf equivalence, and super-strong c-forest-Wilf equivalence of two sets S and S' of patterns. We also extend this notation to $\pi \sim \pi'$, $\pi \overset{c}{\sim} \pi'$, $\pi \overset{sc}{\sim} \pi'$, and $\pi \overset{ssc}{\sim} \pi'$ for equivalences between two individual patterns π and π'.

Recall from Section 1 that the complement $\overline{\pi}$ and twist $\widetilde{\pi}$ of a pattern $\pi = \pi(1) \cdots \pi(k)$ are $k+1-\pi(1), \ldots, k+1-\pi(k)$ and $\pi(1), \ldots, \pi(k-2), \pi(k), \pi(k-1)$, respectively. In other words, the complement is obtained by reversing the order of the elements and the twist is obtained by switching the last two elements. As noted in [1, Proposition 1], given a rooted forest F on $[n]$, we may consider the rooted forest $\overline{F}$ defined as follows: the underlying unlabeled forest structure will be the same, but $L_{\overline{F}}(v) = n+1-L_F(v)$ for all vertices v. In other words, for all $a \in [n]$, we switch the labels a and $n+1-a$. Note that any instance of a pattern π in F will become an instance of $\overline{\pi}$ in $\overline{F}$ under this relabelling, so we automatically have that $\{\pi_1, \ldots, \pi_m\}$ and $\{\overline{\pi}_1, \ldots, \overline{\pi}_m\}$ are equivalent in all senses described before, by complementation of forest labels and patterns. Such an equivalence between π and π' where $\pi = \pi'$ or $\pi' = \overline{\pi}$ is *trivial*.

3. Classical Wilf Equivalences in Forests

In this section we focus on classical forest-Wilf equivalences, the goal being to prove Conjecture 1.2 along with its generalizations. We will consider parts (i) and (ii) in Subsection 3.1 and part (iii) in Subsection 3.2. The general strategy in both cases is to make use of the *forest Simion-Schmidt bijection*, a forest analogue of the classical bijection between permutations avoiding 123 and 132 given by Simion and Schmidt in 1985 in [30]. An analogue of this bijection for posets arose as a special case of the work [20] of Hopkins and Weiler on pattern avoidance in posets. It was first formulated specifically for rooted labeled forests by Anders and Archer in [1]. Garg and Peng used a variant of this bijection in [17] to prove Theorem 1.1, introducing the *shuffle* and *antishuffle* maps on forests. We first recall a special case of their bijection and refer the reader to [17, Section 4.1] for the proofs of the following statements.

Definition 3.1. In a rooted forest on $[n]$, a *top-down minimum* is a vertex v such that $L(u) \geq L(v)$ for all ancestors u of v.

Note that a forest avoids 123 if and only if for all vertices v of the forest that are not top-down minima, $L(v) \geq L(w)$ for all descendants w of v. Similarly, a forest avoids 132 if and only if for all vertices v of the forest that are not top-down minima, if u is the lowest ancestor of v that is a top-down minimum, then $L(v) \leq L(w)$ or $L(u) \geq L(w)$ for all descendants w of v. Figure 2 shows an example of a 123-avoiding forest and a 132-avoiding forest.

Definition 3.2. Suppose that v is a vertex in a labeled forest on $[n]$ that is not a top-down minimum. The *shuffle* and *antishuffle* operations on v are permutations of the labels among the descendants of v such that the relative order of the labels among the strict descendants of v is preserved. This is accomplished by replacing the label of v with the label of one of its descendants and relabelling the strict descendants of v to preserve the relative order. For a shuffle, we relabel v with the maximum value of $L(w)$ among descendants w of v. For an antishuffle, we relabel v with the minimum value of $L(w)$ among descendants w of v that is greater than $L(u)$, where u is the lowest ancestor of v that is a top-down minimum.

Note that shuffles fix the labels and structure of a 123-avoiding forest and antishuffles fix the labels and structure of a 132-avoiding forest. The operation α on a 132-avoiding forest F on $[n]$ is defined as follows: shuffle the vertices of F that are not top-down minima one by one, ordering the vertices by depth from highest to lowest and breaking ties arbitrarily. In this way, we shuffle some leaf of the tree first and the root of the tree last, and we never change which vertices in the tree are top-down minima. The operation β on a 123-avoiding forest F on $[n]$ is defined as follows: antishuffle the vertices of F that are not top-down minima one by one, ordering the vertices by depth from lowest to highest and breaking ties arbitrarily. In this way, we antishuffle the root of the tree first and some leaf of the tree last. The order in which we shuffle or antishuffle vertices of some fixed depth does not matter, as their subtrees do not interact with each other when applying the shuffles or antishuffles. Thus, α and β are well-defined. We note that α maps 132-avoiding forests to 123-avoiding forests and β maps 123-avoiding forests to 132-avoiding forests. Furthermore, α and β are inverses. In fact, the corresponding steps of α and β are all inverses. By this, we mean that if in a step of α we shuffle a vertex v so that the subtree T rooted at v yields a tree T' that only differs from T by labels, the corresponding step of β that antishuffles v is applied to T' and results in T, and vice versa. The maps α and β are also clearly structure-preserving, in that they only permute the labels of the underlying rooted forest. Figure 2 shows an example of an application of α and β with two shuffles or antishuffles. This case is particularly simple because the only nontrivial shuffles and antishuffles are at the blue and red vertices, and shuffles and antishuffles applied to their descendants do not change the forest.

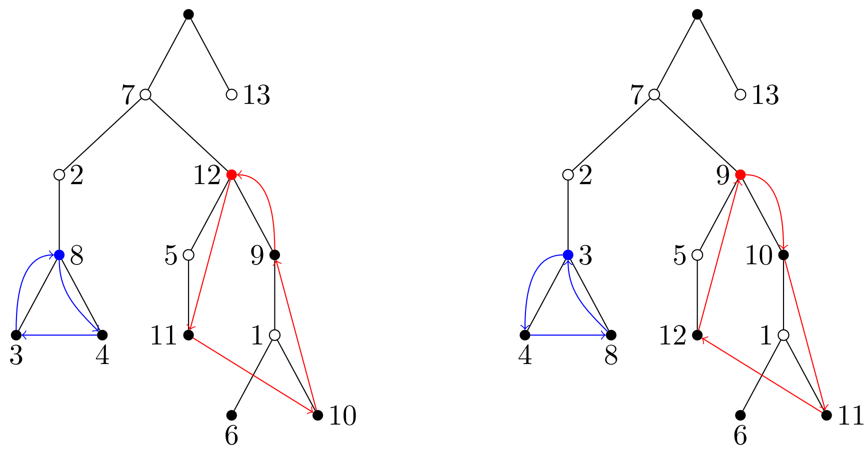

FIGURE 2. The forest on the left avoids 123, and the forest on the right avoids 132. The top-down minima have been drawn with unfilled vertices and are the same for both forests. The blue and red cycles show how vertex labels are permuted when the blue and red vertices are shuffled or antishuffled. These forests correspond to each other under α and β.

Finally, we record the following result, which is a special case of [17, Theorem 4.8].

Theorem 3.3. *The maps α and β restrict to maps of a bijection between rooted forests on $[n]$ avoiding $\{213, 123\}$ and rooted forests on $[n]$ avoiding $\{213, 132\}$.*

3.1. **Restricting the forest Simion-Schmidt bijection.**
We will now prove parts (i) and (ii) of Conjecture 1.2.

Recall that we wish to show that $\{123, 2413\} \sim \{132, 2314\}$ and $\{123, 3142\} \sim \{132, 3124\}$. Our approach is to show that the maps α and β restrict to inverse maps of a bijection between rooted forests on $[n]$ avoiding $\{123, X\}$ and rooted forests on $[n]$ avoiding $\{132, Y\}$, where we have that $(X, Y) \in \{(2413, 2314), (3142, 3124)\}$. Once we do so, the conjecture follows. To show this, for each (X, Y) we identify a property P of rooted forests such that a 123-avoiding forest avoids X if and only if it satisfies P, a 132-avoiding forest avoids Y if and only if it satisfies P, and P is preserved by shuffles and antishuffles. This is clearly sufficient for demonstrating the following theorem.

Theorem 3.4. *We have that the classical forest-Wilf equivalences $\{123, 2413\} \sim \{132, 2314\}$ and $\{123, 3142\} \sim \{132, 3124\}$ hold.*

For the rest of this subsection, we will refer to vertices that are top-down minima as TDM vertices and other vertices as non-TDM vertices.

Definition 3.5. A non-TDM vertex v is *special* if the path from the root of the forest to v contains (not necessarily consecutive) vertices v_1, v_2, v_3, and v_4, in that order, such that v_1 and v_3 are TDM and v_2 and v_4 are non-TDM (for example, we may take $v_4 = v$).

Definition 3.6. Let the *ceiling* of a special vertex v be its lowest ancestor u such that the path from u to v contains (not necessarily consecutive) vertices v_1, v_2, v_3, and v_4, in that order, such that v_1 and v_3 are TDM and v_2 and v_4 are non-TDM (so we necessarily have $v_1 = u$ and $v_4 = v$).

Any path starting from the root of the forest will consist of vertices

$$v_{0,1}, \ldots, v_{0,m_0}, w_{0,1}, \ldots, w_{0,n_0}, v_{1,1}, \ldots, v_{1,m_1}, w_{1,1}, \ldots, w_{1,n_1}, \ldots$$

where $v_{i,j}$ and $w_{i,j}$ respectively denote TDM and non-TDM vertices. The special vertices along this path are $w_{i,j}$ for $i > 0$, and the special vertices $w_{i,1}, \ldots, w_{i,n_i}$ all have the same ceiling $v_{i-1,m_{i-1}}$.

Definition 3.7. A P_1*-forest* is a forest in which the following holds: for every special vertex v with ceiling u, $L(u) > L(v)$.

Here, P_1 is the key property P for $(X, Y) = (2413, 2314)$ mentioned at the beginning of this subsection.

Example 3.8. Figure 3 shows a forest with labels ommitted, where TDM vertices have been colored black and non-TDM vertices have been colored white. The special vertices are v_1 and v_2, with respective ceilings u_1 and u_2.

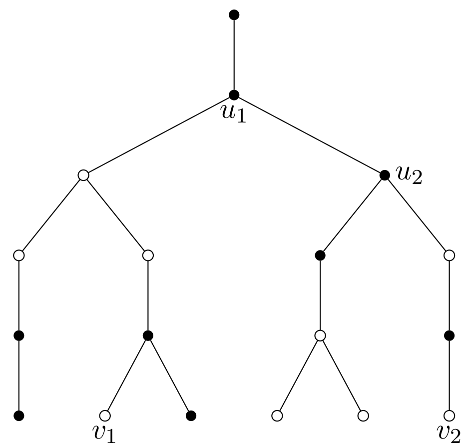

FIGURE 3. The special vertices v_1 and v_2 have respective ceilings u_1 and u_2.

Lemma 3.9. *A* 123*-avoiding forest avoids* 2413 *if and only if it is a* P_1*-forest.*

Proof. Recall that a forest avoids 123 if and only if every non-TDM vertex has the largest label among its descendants.

In one direction, we show that if a forest avoids 123 and contains 2413, then it contains a special vertex whose label is larger than its ceiling's. Suppose that v_1, v_2, v_3, v_4 is an instance of 2413 in the forest, so v_1, v_2, v_3, v_4 lie in that order on a path from the root and $L(v_3) < L(v_1) < L(v_4) < L(v_2)$. Since $L(v_1) < L(v_2)$ and $L(v_3) < L(v_4)$, v_2 and v_4 are non-TDM. Additionally, v_1 and v_3 are TDM since they are not greater than all of their strict descendants. Thus, v_4 is a special vertex, and its ceiling u is a TDM descendant of v_1. Then we have $L(u) < L(v_1) < L(v_4)$, so v_4 has a larger label than its ceiling, as desired.

In the other direction, we show that if a forest avoids 123 and contains a special vertex whose label is larger than its ceiling's, then it contains 2413. Let v be this special vertex with ceiling u, and let a and b be vertices along the path from u to v such that u and b are TDM, a and v are non-TDM, and u, a, b, v appear in that order. We may assume a is a child of u, as u's child is non-TDM and appears before b and v. As b is TDM, $L(b) < L(u), L(a)$, and as a is non-TDM, $L(a) > L(b), L(v)$. Since a is non-TDM and the child of the non-TDM vertex u, $L(u) < L(a)$. By assumption, $L(v) > L(u)$, so $L(a) > L(v) > L(u) > L(b)$. Hence, u, a, b, v form an instance of 2413, as desired. □

Lemma 3.10. *A* 132*-avoiding forest avoids* 2314 *if and only if it is a* P_1*-forest.*

Proof. Recall that a forest avoids 132 if and only if every non-TDM vertex has the smallest label among its descendants that are greater than its lowest TDM ancestor.

In one direction, we show that if a forest avoids 132 and contains 2314, then it contains a special vertex whose label is larger than its ceiling's. Suppose that v_1, v_2, v_3, v_4 is an instance of 2314 in the forest, so v_1, v_2, v_3, v_4 lie in that order on a path from the root and $L(v_3) < L(v_1) < L(v_2) < L(v_4)$. Since $L(v_2) > L(v_1)$ and $L(v_4) > L(v_3)$, v_2 and v_4 are non-TDM. Let v be the lowest TDM ancestor of v_2, so $L(v) < L(v_2)$. As $L(v_3) < L(v_2)$, we must have that $L(v) > L(v_3)$ or else v_2 would not have the smallest possible label greater than $L(v)$. Since $L(v_3) < L(v) < L(v_2)$, we may assume that $v_1 = v$, so v_1 is the lowest TDM ancestor of v_2. Since $L(v_1) > L(v_3)$, there exists a TDM vertex on the path from v_1 to v_3 different from v_1. This TDM vertex must come after v_2, so we may assume that it is v_3. Thus, v_1 and v_3 are TDM and v_2 and v_4 are non-TDM, which means that v_4 is a special vertex, and its ceiling u is a TDM descendant of v_1. Then we have $L(u) < L(v_1) < L(v_4)$, so v_4 has a larger label than its ceiling, as desired.

In the other direction, we show that if a forest avoids 132 and contains a special vertex whose label is larger than its ceiling's, then it contains 2314. Let v be this special vertex with ceiling u, and let a and b be vertices along the path from u to v such that u and b are TDM, a and v are non-TDM, and u, a, b, v appear in that order. We may take a to be u's child, as u's child is non-TDM and appears before b and v. As b is TDM, $L(b) < L(u), L(a)$. By assumption, $L(u) < L(v)$, so since a is non-TDM and its parent u is TDM, $L(u) < L(a) < L(v)$. Putting this together, we have $L(v) > L(a) > L(u) > L(b)$, so u, a, b, v form an instance of 2314, as desired. ☐

Lemma 3.11. *Applying a shuffle or antishuffle to a* P_1*-forest results in a* P_1*-forest.*

Proof. For each TDM vertex v, we define the subtree T_v as the maximal subtree of the forest rooted at v whose only TDM vertex is v, where maximal is in the sense of containment. For each non-TDM vertex u, we define the subtree T_u as the tree consisting of the descendants of u in T_v, where v is the lowest TDM ancestor of u. We may view T_u as a subtree of T_v, and T_u only consists of non-TDM vertices. Note that this definition only depends on which vertices of the forest are TDM and which are non-TDM, and that the forest is then partitioned into the trees T_v for TDM v. Figure 4 gives an example of how T_u and T_v are defined, where TDM vertices have been colored black and non-TDM vertices have been colored white.

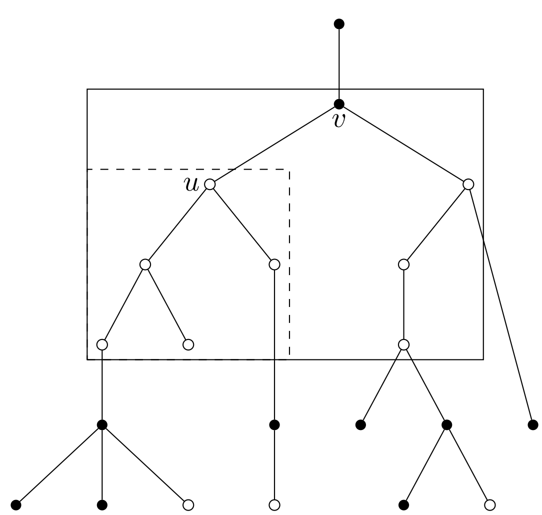

FIGURE 4. The solid box encloses T_v, and the dashed box encloses T_u.

The condition that the forest is a P_1-forest is equivalent to the condition that for all non-TDM vertices u with lowest TDM ancestor v, the descendants of u in T_u have greater labels than the label of v, which is greater than the labels of the descendants of u not in T_u. In other words, T_u consists exactly of the descendants w of u with $L(w) > L(v)$. Indeed, if a vertex w of T_u has label smaller than v, then it would be TDM, a contradiction. If a vertex w not in T_u has label greater than v, then the path from u to w contains a TDM vertex, say x. Then v, u, x, w appear on a path in that order such that v and x are TDM, u and w are non-TDM, and $L(w) > L(v)$. This contradicts the fact that the forest is a P_1-forest, as the label of the ceiling of w is at most the label of v. Every non-TDM descendant w of u that is not in T_u is a special vertex, and every special vertex arises in this way. We then have $L(w) < L(v)$, so the special vertex w has a smaller label than its ceiling, which has label at most $f(v)$. The condition that the forest is a P_1-forest is then satisfied, so the two conditions are equivalent.

Now, note that when we shuffle or antishuffle a non-TDM vertex u, we only permute labels within T_u. Indeed, T_u contains all of the descendants of u with labels greater than $L(v)$, which are the only vertices that are permuted in any shuffle or antishuffle. Since we only permute within T_u, after the shuffle or antishuffle, the equivalent condition that the forest is a P_1-forest remains satisfied, as desired. □

Definition 3.12. For each non-TDM vertex v, let the *segment* of v be the set of TDM ancestors u of v with $L(u) < L(v)$. Let the *top* and *bottom* of the segment of v be the vertices of the segment with the least and greatest labels.

We may refer to these as the top and bottom of v instead of the segment of v. Note that the segment of v necessarily consists of the TDM vertices along the path from the top of v to v.

Definition 3.13. We say that a pair of vertices is *comparable* if one is an ancestor of the other.

We may view a forest as a Hasse diagram for a poset, which inspires the term comparable.

Definition 3.14. A P_2-*forest* is a forest in which the following holds: whenever two comparable non-TDM vertices have segments that intersect, the top of their segments coincide.

Here, P_2 is the key property P for $(X, Y) = (3142, 3124)$ mentioned at the beginning of this subsection. One way to visualize segments is to plot the labels along a path from the root to a leaf. For example, Figure 5 shows such a plot when if the label sequence is $8, 10, 6, 4, 7, 3, 9, 2, 5, 1$, with the top-down minima, now left-right minima, colored black. The segment of a non-TDM vertex consists of the TDM vertices that have a smaller x- and y-coordinate in the plot. The forest that the label sequence in Figure 5 comes from is not a P_2-forest.

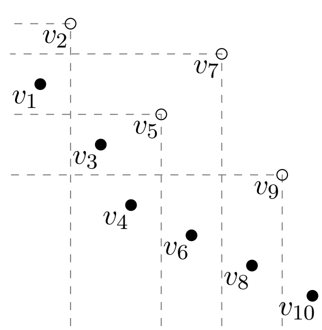

FIGURE 5. The segments of v_5 and v_9 intersect, but their tops v_3 and v_4 are different.

Lemma 3.15. *A* 123*-avoiding forest avoids* 3142 *if and only if it is a* P_2*-forest.*

Proof. As before, a forest avoids 123 if and only if every non-TDM vertex has the largest label among its descendants.

In one direction, we show that if a forest avoids 123 and contains 3142, then it contains a pair of comparable non-TDM vertices whose segments intersect but do not have the same top. Suppose that v_1, v_2, v_3, v_4 is an instance of 3142 in the forest, so v_1, v_2, v_3, v_4 lie in that order on a path from the root and $L(v_2) < L(v_4) < L(v_1) < L(v_3)$. Since $L(v_3), L(v_4) > L(v_2)$, v_3 and v_4 are non-TDM. Since $L(v_1), L(v_2) < L(v_3)$, v_1 and v_2 are not greater than all of their strict descendants and are thus TDM. Now the segment of v_3 contains v_1 and v_2, but the segment of v_4 contains v_2 but not v_1, so v_3 and v_4 form the desired pair of comparable non-TDM vertices.

In the other direction, we show that if a forest avoids 123 and contains a pair of comparable non-TDM vertices whose segments intersect but do not have the same top, then it contains 3142. Suppose that u and v form such a comparable pair with u an ancestor of v. Then, since u is non-TDM, we have that $L(u) > L(v)$, which means that u's top is an ancestor of v's top. As u and v have intersecting segments, v's top must be contained in u's segment. Let a and b be the tops of u and v, respectively, so since a and b are TDM and a is b's ancestor, $L(a) > L(b)$. By assumption, a does not lie in v's segment and b lies in v's segment, so $L(a) > L(v) > L(b)$. Also, a and b are in u's segment, so $L(u) > L(a), L(b)$. Consequently, a, b, u, v appear in that order along a path from the root in the forest with $L(u) > L(a) > L(v) > L(b)$, so they form an instance of 3142, as desired. □

Lemma 3.16. *A* 132*-avoiding forest avoids* 3124 *if and only if it is a* P_2*-forest.*

Proof. As before, a forest avoids 132 if and only if every non-TDM vertex has the smallest label among its descendants that are greater than the label of its lowest TDM ancestor, i.e. its bottom.

In one direction, we show that if a forest avoids 132 and contains 3124, then it contains a pair of comparable non-TDM vertices whose segments intersect but do not have the same top. Suppose that v_1, v_2, v_3, v_4 is an instance of 3124 in the forest, so v_1, v_2, v_3, v_4 lie in that order on a path from the root and $L(v_2) < L(v_3) < L(v_1) < L(v_4)$. Since $L(v_3) > L(v_2)$ and $L(v_4) > L(v_1)$, v_3 and v_4 are non-TDM. Because $L(v_1) > L(v_3)$ and v_3 is non-TDM, there exists a TDM vertex v on the path from v_1 to v_3. Indeed, the bottom of v_3 necessarily lies strictly between v_1 and v_3. Suppose that there are no TDM vertices on the path between v and v_3 other than v. We then have that $L(v) < L(v_3)$ so we may assume that $v_2 = v$ is the lowest TDM ancestor of v_3. Furthermore, if v_1 is non-TDM then since $L(v_3) < L(v_1)$, $L(v_3)$ must be less than the label of v_1's bottom, which means that we can replace v_1 with its bottom while preserving the 3124 order of v_1, v_2, v_3, v_4. Thus, we may assume that v_1 and v_2 are TDM and v_3 and v_4 are non-TDM, which means that v_3's segment contains v_2 but not v_1 while v_4's segment contains v_1 and v_2. Hence v_3 and v_4 form the desired pair of comparable non-TDM vertices.

In the other direction, we show that if a forest avoids 132 and contains a pair of comparable non-TDM vertices whose segments intersect but do not have the same top, then it contains 3124. Suppose that u and v forms such a comparable pair with u an ancestor of v. We must have that $L(u) < L(v)$, as if $L(u) > L(v)$ then $L(v)$ must be smaller than the label of u's bottom, which means that v's top has a label smaller than u's bottom so u and v have disjoint segments, a contradiction. As u and v have intersecting segments, u's top must be contained in v's segment. Let a and b be the tops of u and v, respectively. By assumption, b does not lie in u's segment and a lies in u's segment, so $L(a) < L(u) < L(b)$. Also, a and b are in v's segment, so $L(v) > L(a), L(b)$. Consequently, b, a, u, v appear in that order along a path from the root in the forest and satisfy the inequalities $L(v) > L(b) > L(u) > L(a)$, so they form an instance of 3124, as desired. □

Lemma 3.17. *Applying a shuffle or antishuffle to a* P_2*-forest results in a* P_2*-forest.*

Proof. It suffices to show that the segment of each non-TDM vertex does not change when we shuffle or antishuffle any non-TDM vertex, as shuffles and antishuffles do not change which vertices are TDMs. Since the segment of a non-TDM vertex only depends on its top, it suffices to show that the top of each non-TDM vertex does not change when we shuffle or antishuffle any non-TDM vertex.

When we shuffle a non-TDM vertex v, we replace its label $L(v)$ with the largest label $L(w)$ among the descendants of v. Note that if $v \neq w$, then the bottom of v is in the segment of w. Thus, v and w have intersecting segments, so they have the same tops. It follows that when we replace the label of v with $f(w)$, the top of v does not change. The same argument works for all of the vertices that change labels during the shuffle. They are the descendants of v whose labels lie between $L(v)$ and $L(w)$. All of these vertices have the same top as v and w, and this does not change after the shuffle relabeling. The segments of the other non-TDM descendants of v do not change either.

When we antishuffle a non-TDM vertex v, we replace its label $L(v)$ with the smallest label $L(w)$ among the descendants of v that is greater than the label of the bottom of v. Note that if $v \neq w$, then the bottom of v is in the segment of w. Thus, v and w have intersecting segments, so they have the same tops. It follows that when we replace the label of v with $L(w)$, the top of v does not change. The same argument works for all of the vertices that change labels during the shuffle. They are the descendants of v whose labels lie between $L(v)$ and $L(w)$. All of these vertices have the same top as v and w, and this does not change after the antishuffle relabeling. The segments of the other non-TDM descendants of v do not change either. □

Remark 3.18. The properties P_1 and P_2 we constructed can be viewed as properties on sequences of distinct integers that we enforce on the label sequence of every path from the root of the forest to a leaf. Indeed, from this perspective the concept of top-down minima translates to the concept of left-right minima in the sequence, and properties P_1 and P_2 retain their definitions along with TDM and non-TDM vertices, now terms in the sequence. Lemmas 3.9, 3.10, 3.15, and 3.16 show that sequences avoiding 123 or 132 avoid some another pattern we are interested if and only if the analogous property P_1 or P_2 holds, and their proofs are local in the sense that only vertices along one path is considered at a time. On the other hand, Lemmas 3.11 and 3.17 work globally with the whole forest, establishing that shuffles and antishuffles preserve properties P_1 and P_2 along each path in the forest.

With all of these lemmas, we can finish the proof of Theorem 3.4.

Proof of Theorem 3.4. As mentioned before, to show $\{123, X\} \sim \{132, Y\}$ for some patterns X and Y, it suffices to show that the maps α and β of Garg and Peng in [17] defined earlier restrict to maps between forests avoiding $\{123, X\}$ and forests avoiding $\{132, Y\}$. To do so, it suffices to exhibit a property P of forests such that a 123-avoiding forest avoids X if and only if it satisfies P, a 132-avoiding forest avoids Y if and only if it satisfies P, and P is preserved by shuffles and antishuffles. For $(X, Y) = (2413, 2314)$, the property is P_1, demonstrated by Lemmas 3.9, 3.10, and 3.11, and for $(X, Y) = (3142, 3124)$, the property is P_2, demonstrated by Lemmas 3.15, 3.16, and 3.17. □

Since α and β restrict to bijections between forests avoiding $\{123, 2413\}$ and forests avoiding $\{132, 2314\}$ as well as between forests avoiding $\{123, 3142\}$ and forests avoiding $\{132, 3124\}$, they also restrict to bijections between the intersection of these two restrictions.

Corollary 3.19. *We have the forest-Wilf equivalence* $\{123, 2413, 3142\} \sim \{132, 2314, 3124\}$.

3.2. **Forests avoiding 213 and another pattern.**

We will now prove Theorem 1.3 by generalizing the bijection given by Theorem 3.3.

Throughout this subsection, let $\pi = \pi(1) \cdots \pi(k)$ denote a pattern of length k that also satisfies $\pi(k) = \pi(k-1) + 1 = \pi(k-2) + 2$. Suppose that $\pi(j) = 1$. Note that Theorem 1.3 in this case is trivial if π contains 213. Thus, suppose that it does not, so $\pi(1), \dots, \pi(j-1) > \pi(j), \dots, \pi(k)$.

Definition 3.20. We say that a positive integer $i < j$ is a *checkpoint* if the numbers $\pi(1), \dots, \pi(i)$ are all greater than the numbers $\pi(i+1), \dots, \pi(j)$.

If the points $P_i = (i, \pi(i))$ are plotted, then i is a checkpoint if we can draw two lines parallel to the coordinate axes such that $P_1, \dots, P_i$ lie in the upper left quadrant and $P_{i+1}, \dots, P_k$ lie in the lower right quadrant formed by the lines.

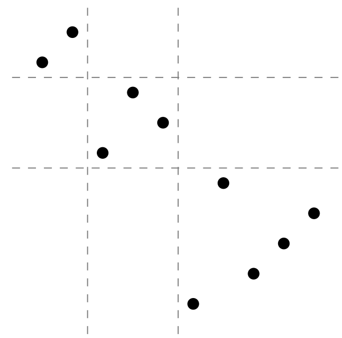

FIGURE 6. The checkpoints of the pattern $9, 10, 6, 8, 7, 1, 5, 2, 3, 4$ are 2 and 5.

Suppose that the checkpoints of π are $a_1 < \cdots < a_m$, which is an empty list when $\pi(1) = 1$ but always contains $j-1$ when $j > 1$. For convenience, we define $a_0 = 0$.

Definition 3.21. Let the *rank* of a vertex v in a forest with respect to π be the greatest positive integer $i \leq m$ such that the path from the root of the forest to v contains the pattern $\pi(1) \cdots \pi(a_i)$. If π has no checkpoints, or if such an integer i does not exist, then we define the rank to be 0.

All trees and forests considered in this subsection will avoid 213. A forest avoiding 213 is a set of trees that avoid 213. In a tree that avoids 213, the labels greater than the label of the root must come before any labels less than the root on any path from the root to a leaf, a fact shown for example in [17, Section 3.2]. Thus, we may decompose the tree into the tree of labels that are at least the label of the root, which we refer to as the *large tree*. The rest of the vertices necessarily form a forest, which we refer to as the *small forest*. Note that the small forest is empty when the root has the smallest label in the tree, but the large tree is always nonempty. We will also consider for a given vertex v in the large tree the vertices in the small forest whose lowest ancestor in the large tree is v. These descendants of v form a forest, which we refer to as the *small subforest of* v.

The rank of a vertex v in the large tree of a tree T measures the "progress" that the path from the root of T to v makes in containing an instance of π by recording the index of the furthest checkpoint the label sequence in this path gets past. Figure 7 gives an example of the decomposition of a 213-avoiding forest along with ranks of vertices in the large tree with respect to $\pi = 54123$. Continuing into the small subforest of v, the only thing that matters towards containing π is how much progress was already made in the large tree of T, i.e. the rank of v. Avoiding or containing π now becomes a matter of avoiding or containing a truncation of π in the small subforest of v. This is formalized in the following rank additivity lemma.

Lemma 3.22. *Suppose that a vertex v of the large tree of a tree T has rank r with respect to T and π. Let F be the small subforest of v with respect to T, and suppose that a vertex w of F has rank s with respect to F and the pattern $\pi(a_r + 1) \cdots \pi(k)$. Then w has rank $r + s$ with respect to T and π. Furthermore, the union of the path from the root of T to v and F avoids π if and only if F avoids the pattern $\pi(a_r + 1) \cdots \pi(k)$.*

Proof. We first take care of the case that both r and s are positive. As defined before, we suppose that π has checkpoints at $a_1, \dots, a_m$. Note that $\pi(a_r+1) \cdots \pi(k)$ then has checkpoints at the indices

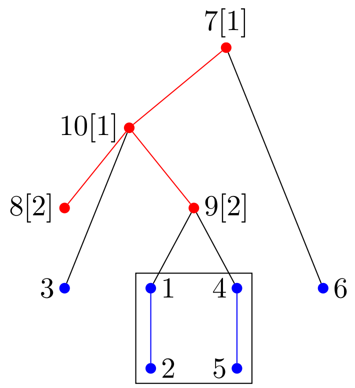

FIGURE 7. A 213-avoiding tree on [10]. The large tree on $\{7, 8, 9, 10\}$ is colored red, the small forest on $\{1, 2, 3, 4, 5, 6\}$ is colored blue, and the small subforest of the vertex labeled 9 is boxed. Vertices in the large tree are also marked by their ranks with respect to the pattern 54123 in brackets.

$a_{r+1} - a_r, \ldots, a_m - a_r$. Suppose that $v_1, \ldots, v_{a_r}$ is an instance of $\pi(1)\cdots\pi(a_r)$ on the path from the root of T to v and that $w_1, \ldots, w_{a_{r+s}-a_r}$ is an instance of $\pi(a_r+1)\cdots\pi(a_{r+s})$ on the path from v to w, not including v. Then as a_r is a checkpoint, we know that $\pi(1), \ldots, \pi(a_r)$ are all greater than $\pi(a_r+1), \ldots, \pi(a_{r+s})$. As $v_1, \ldots, v_{a_r}$ are ancestors of v in the large tree and $w_1, \ldots, w_{a_{r+s}-a_r}$ are in the small subforest of v, we have that the labels of $v_1, \ldots, v_{a_r}$ are greater than the labels of $w_1, \ldots, w_{a_{r+s}-a_r}$. Thus, $v_1, \ldots, v_{a_r}, w_1, \ldots, w_{a_{r+s}-a_r}$ is an instance of $\pi(1)\cdots\pi(a_{r+s})$ and the rank of w with respect to T and π is at least $r+s$. Conversely, suppose that $r+s < m$ (note that we are done if $r+s = m$) and that the rank of w with respect to T and π is more than $r+s$. Then, there is an instance $v_1, \ldots, v_x, w_1, \ldots, w_y$ of $\pi(1)\cdots\pi(a_{r+s+1})$ on the path from the root of T to w, where the v_i are ancestors of v and the w_i are in the small subforest of v. As the labels of $v_1, \ldots, v_x$ are greater than the labels of $w_1, \ldots, w_y$, x must be a checkpoint of π. As the rank of v with respect to T and π is r, we have that $x \le a_r$, so $w_1, \ldots, w_y$ contains an instance of $\pi(a_r+1)\cdots\pi(a_{r+s+1})$. But $a_{r+s+1} - a_r$ is the $(s+1)$-st checkpoint of $\pi(a_r+1)\cdots\pi(k)$, contradicting the fact that the rank of w with respect to F and $\pi(a_r+1)\cdots\pi(k)$ is s. Thus, the rank of w with respect to T and π is $r+s$.

The other cases are trivial or follow similarly. If $r = 0$ because π has no checkpoints (so $\pi(1) = 1$), then all vertices of T have rank 0. We then have that s is the rank of w with respect to π in F, which is 0 as π has no checkpoints. If $r = 0$ because in the path from the root of T to v, there is no instance of $\pi(1)\cdots\pi(a_1)$, then the proof is the same as in the case $r, s > 0$. If $s = 0$ because $\pi(a_r+1)\cdots\pi(k)$ has no checkpoints, then the rank of v with respect to π and T is $r = m$. Then all descendants of v also have rank m, and all vertices in F have rank 0 with respect to F and $\pi(a_r+1)\cdots\pi(k)$ since it has no checkpoints. If $s = 0$ because there is no instance of $\pi(a_r+1)\cdots\pi(a_{r+1})$ in the path from v to w excluding v, then the proof is the same as the proof for $r, s > 0$.

The proof of the second part of the lemma is essentially the same. Any instance of π of the form $v_1, \ldots, v_x, w_1, \ldots, w_y$ with v_i in the large tree and w_i in the small subforest of v must satisfy the property that the labels of $v_1, \ldots, v_x$ are greater than the labels of $w_1, \ldots, w_y$. Thus, x is a checkpoint of π and is at most a_r since the rank of v with respect to π and T is r. Then, $w_1, \ldots, w_y$ must contain an instance of $\pi(a_r+1)\cdots\pi(k)$. Conversely, if $w_1, \ldots, w_{k-a_r}$ form an instance of $\pi(a_r+1)\cdots\pi(k)$ in F, then concatenating with $v_1, \ldots, v_{a_r}$, an instance of $\pi(1)\cdots\pi(a_r)$ on the path from the root of T to v, gives an instance of π. $\square$

The key idea behind the proof for Theorem 1.3 is a recursive bijection between forests avoiding $\{213, \pi\}$ and forests avoiding $\{213, \widetilde{\pi}\}$ that preserves the structure of the forest (it only permutes the labels) and rank of each vertex. Lemma 3.22 allows us to transfer between avoiding π and avoiding truncations of π in small subforests, which allows us to inductively define such a bijection, with Theorem 3.3 forming our base case. We denote the structure- and rank-preserving map from forests avoiding $\{213, \pi\}$ to forests avoiding $\{213, \widetilde{\pi}\}$ by f_π and the inverse map by $f_{\widetilde{\pi}}$. Note that it suffices to define such a bijection between trees avoiding $\{213, \pi\}$ and trees avoiding $\{213, \widetilde{\pi}\}$, as we can then apply this bijection to each tree in the forest. The ranks of the vertices only depend on the relative ordering of the labels within a tree, so the bijection extends properly. We define this bijection recursively on trees T with the following cases, noting that whenever we write f_σ or $f_{\widetilde{\sigma}}$ for some pattern σ, the last three terms of σ are increasing consecutive integers. We may refer to applying f_σ to a subforest F of T, by which we mean to replace the labels of F by the labels of $f_\sigma(F)$ (note this does not change the label set of F). We define f_π and $f_{\widetilde{\pi}}$ by splitting into the following cases, where we are allowed to apply f_σ to the forest in cases (3) and (6) by Lemma 3.22:

(1) If T has less than three vertices, we define $f_\pi(T) = f_{\widetilde{\pi}}(T) = T$ to be the identity map.
(2) If $\pi = 123$, we define f_π and $f_{\widetilde{\pi}}$ to be β and α, respectively, where α and β are as defined in the beginning of this section.
(3) If $\pi(1) = 1$ and the label of the root of the tree is 1, we define $f_\pi(T)$ and $f_{\widetilde{\pi}}(T)$ to be the result of applying f_σ and $f_{\widetilde{\sigma}}$ to the strict descendants of the root, which is a forest, where $\sigma = \pi(2) - 1, \ldots, \pi(k) - 1$.
(4) If $\pi(1) > 1$ and the label of the root of the tree is 1, we define $f_\pi(T)$ and $f_{\widetilde{\pi}}(T)$ to be the result of applying f_π and $f_{\widetilde{\pi}}$ to the strict descendants of the root, which is a forest.
(5) If $\pi(1) = 1$ and the label of the root is greater than 1, we define $f_\pi(T)$ and $f_{\widetilde{\pi}}(T)$ to be the result of applying f_π and $f_{\widetilde{\pi}}$ to both the large tree and small forest of T, respectively.
(6) If $\pi(1) > 1$ and the label of the root is greater than 1, we define $f_\pi(T)$ and $f_{\widetilde{\pi}}(T)$ as follows. We respectively apply f_π and $f_{\widetilde{\pi}}$ to the large tree of T. Then for each vertex v of the large tree of T, we do the following. If the rank of v with respect to T and v is r, then we respectively apply f_σ and $f_{\widetilde{\sigma}}$ to the small subforest of v in T, where $\sigma = \pi(a_r + 1) \cdots \pi(k)$.

Proposition 3.23. *The maps f_π and $f_{\widetilde{\pi}}$ are structure- and rank-preserving and are inverses. Furthermore, f_π maps $\{213, \pi\}$-avoiding forests to $\{213, \widetilde{\pi}\}$-avoiding forests and $f_{\widetilde{\pi}}$ maps $\{213, \widetilde{\pi}\}$-avoiding forests to $\{213, \pi\}$-avoiding forests.*

Proof. We proceed using induction on k (the length of π) and the number of vertices in the tree to show that the defined maps are inverses, noting that the bijection clearly extends to forests with all of the claimed properties. Note that the notion of rank does not depend on $\pi(k-1)$ or $\pi(k)$, so the rank with respect to π is the same as the rank with respect to $\widetilde{\pi}$.

For the base case, if the number of vertices is less than three, then the two maps are both the identity map and are obviously inverses of each other that fix structure and rank. If $\pi = 123$, then Theorem 3.3 along with the definitions of α and β implies that the defined maps are inverses of each other and preserve structure. All vertices have rank 0 when $\pi = 123$, so the base case follows.

For the inductive step, note first that the maps we define are always inverses of each other, preserve structure, and do not change the large tree of T. Indeed, we always either apply f_π or $f_{\widetilde{\pi}}$ to a smaller forest (note the different treatment of the cases when the root has label 1, corresponding to when the small forest is empty) or apply f_σ or $f_{\widetilde{\sigma}}$, where σ is shorter than π. We always apply f_σ to disjoint subforests of T. Hence, f_π and $f_{\widetilde{\pi}}$ always invert each other. Indeed, we also see by induction that whether the root of the tree is 1 or not is also fixed by our maps, so all of the cases of the recursive definitions imply that the maps are inverses of each other. This settles all but the last case defined in the recursion, where the smaller map applied depends on the rank of vertices in T. Once we show that the maps preserve rank, this case also follows, and all of the claimed properties of the maps will be proved.

For the inductive step that our defined maps preserve vertex ranks, we split into the cases in our definition of f_π. Note that there is nothing to show when $\pi(1) = 1$ (Cases (3) and (5)), as all vertices have rank 0 in that case. If $\pi(1) > 1$ and the label of the root of the tree is 1 (Case (4)), then note that the rank of the root is 0 and the rank of a vertex v with respect to π and T is equal to the rank of v with respect to T with the root removed by Lemma 3.22. This follows because when the root has label 1 and $\pi(1) > 1$, the root is never part of an instance of π. If $\pi(1) > 1$ and the label of the root of the tree is greater than 1 (Case (6)), then note that the ranks of all vertices in the large tree are preserved by induction, as the large tree has fewer vertices than T. But the ranks with respect to all of the the small subforests and their corresponding truncated patterns are also preserved by induction, so by Lemma 3.22, the rank of a vertex of the small forest with respect to T and π is preserved as the sum of its rank with respect to the small subforest and the rank of its lowest ancestor in the large tree.

It remains to show that f_π maps a tree avoiding $\{213, \pi\}$ to a tree avoiding $\{213, \widetilde{\pi}\}$. This inductive step argument is essentially the same as the previous one. When the root of T has label 1 and $\pi(1) > 1$ (Case (4)), avoiding π is equivalent to avoiding π in the forest resulting from removing the root, as the root is never involved in instances of π. When the root of T has label 1 and $\pi(1) = 1$ (Case (3)), avoiding π is equivalent to avoiding $\pi(2)\cdots\pi(k)$ in the forest resulting from removing the root, as every instance of π in T gives an instance of $\pi(2)\cdots\pi(k)$ in the forest, while every instance of $\pi(2)\cdots\pi(k)$ in the forest gives an instance of π in T by appending the root to the beginning. When the root of T has label greater than 1 and $\pi(1) = 1$ (Case (5)), note that instances of π must be entirely contained in the large tree or the small forest, since the first vertex of the occurrence must have the smallest label. Finally, when the root of T has label greater than 1 and $\pi(1) > 1$ (Case (6)), instances of π occur if and only if the small subforests do not contain the corresponding truncations of π by Lemma 3.22. We have shown avoidance of π to be equivalent to a combination of avoidance of π and truncations of π in disjoint subforests, and the recursion we have defined gives exactly the same equivalent avoidances in the corresponding subforests, so we are done. □

Proof of Theorem 1.3. Proposition 3.23 shows that the bijection we have defined exists, preserves structure and rank, and maps $\{213, \pi\}$-avoiding forests on $[n]$ to $\{213, \widetilde{\pi}\}$-avoiding forests on $[n]$, so we have that $\{213, \pi\} \sim \{213, \widetilde{\pi}\}$. □

4. Consecutive Wilf Equivalences in Forests

In this section, we discuss various notions of c-forest-Wilf equivalence. As mentioned in Section 2, whenever we refer to instances and equivalences throughout this section, we mean consecutive instances and c-forest-Wilf equivalences, respectively. In Subsection 4.1, we consruct a large family of such c-forest-Wilf equivalences between certain types of patterns we call *grounded permutations* to prove Theorems 1.5 and 1.6. Subsection 4.2 then discusses some necessary conditions for two grounded permutations to be strongly c-forest-Wilf equivalent, following the methodology of Garg and Peng in their proof the following theorem from [17].

Theorem 4.1 ([17, Theorem 1.5])**.** *If π and π' are patterns of length k such that $\pi \overset{sc}{\sim} \pi'$, then $\pi(1) = \pi'(1)$ or $\pi(1) + \pi'(1) = k + 1$.*

Finally, we investigate super-strong c-forest-Wilf equivalence in Subsection 4.3 and prove Theorem 1.7.

We begin by defining some notations and terms specific to the theory of consecutive avoidance. The following notion of a *forest cluster* was introduced by Garg and Peng in [17].

Definition 4.2 ([17, Definition 6.1])**.** A *m-cluster of size n* with respect to a pattern π is a rooted tree on $[n]$ (a rooted forest on $[n]$ with only one component) with m distinct *highlighted instances* of π such that every vertex is a part of a highlighted instance and there is no proper nonempty subset V of vertices such that all highlighted instances either lie completely in V or do not intersect V. Informally, the highlighted instances in the cluster are "connected."

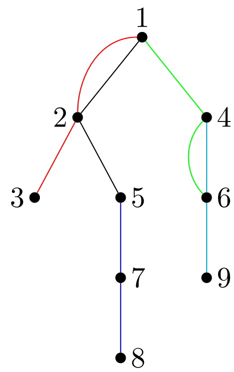

FIGURE 8. A 5-cluster of size 9 with respect to 123. The instances are the colored paths on $\{1,2,3\}$, $\{1,2,5\}$, $\{1,4,6\}$, $\{4,6,9\}$, and $\{5,7,8\}$.

Figure 8 shows a forest cluster with respect to the pattern 123, where we use colored paths to denote the different highlighted instances. We do not require that every instance is highlighted. Indeed, note that the instance $2,5,7$ is not highlighted. If it were, then we would have a different forest cluster.

Forest clusters are the analogue of permutation clusters for forests. General cluster methods were introduced by Goulden and Jackson in [16] and applied to permutation consecutive pattern avoidance by Elizalde and Noy in [11]. Forest clusters were introduced by Garg and Peng in [17], where their connection to strong equivalence was shown through the following theorem.

Theorem 4.3 ([17, Theorem 6.2])**.** *Two patterns π and π' are strongly equivalent if and only if for all $m, n \geq 0$, the number of m-clusters of size n with respect to π is equal to the number of m-clusters of size n with respect to π'.*

Equivalences between patterns are typically proven by showing strong equivalence via the cluster method for both permutations and forests. The proof of Theorem 1.4 in [17] proceeds by showing that the number of clusters for 1324 and 1423 satisfy the same recursion. Our results in this section also follow this method. It would be interesting to see a proof of equivalence without appealing to strong equivalence.

We can refine the notion of equivalence by specifying not only the number of vertices in the cluster but the structure of the tree.

Definition 4.4. Two patterns π and π' satisfy *cluster structure equivalence* if for all (unlabeled) rooted trees T and integers $m \geq 0$, the number of clusters on T with m highlighted instances of π is equal to the the number of clusters on T with m highlighted instances of π'.

We can also extend the notion of *super-strong* equivalence, introduced by Dwyer and Elizalde for permutations in [6], to forests.

Definition 4.5. Two patterns π and π' of length k are *super-strongly* equivalent if for all unlabeled rooted forests F on n vertices with a set S of highlighted paths of length k, the number of ways to label the vertices of F with distinct labels in $[n]$ such that S is the set of instances of π in F is equal to the number of ways to label the vertices of F with distinct labels in $[n]$ such that S is the set of instances of π' in F.

4.1. **A large family of nontrivial equivalences.**

In this section, we prove Theorems 1.5 and 1.6, which state that there is an equivalence class of size at least 2^{n-4} for all n and that a $(1-o(1))^n$-fraction of patterns of length n satisfy a nontrivial equivalence. We do so by first demonstrating a family of nontrivial equivalences through a series of lemmas. Then, we use the properties of this family to obtain Theorems 1.5 and 1.6 as a result.

First, we define some notions related to a certain class of permutations we will consider. Throughout this subsection, we assume that $\pi = \pi(1)\cdots\pi(k)$ is a pattern of length k that is not the identity.

Definition 4.6. The *streak* of a pattern $\pi = \pi(1)\cdots\pi(k)$ is the largest integer m such that $\pi(i) = i$ for all $1 \le i \le m$.

Note that if $\pi(1) \neq 1$, then the streak of π is 0.

Definition 4.7. For $1 < i \le k$, let the *height* of $\pi(i)$ in π be the largest integer $m < i$ such that $\pi(i-m+1) < \ldots < \pi(i)$, i.e. the length of the longest consecutive increasing subsequence beginning after $\pi(1)$ and ending at $\pi(i)$. For $\pi(1) \neq j \in [k]$, let the height of j in π be the height of $\pi(i)$ in π, where $\pi(i) = j$. Let the *max height* of π, be the maximum of the height of $\pi(i)$ over all $1 < i \le k$.

Example 4.8. The pattern $\pi = 123485679$ has streak 4 and max height 4, since the heights of 8 and 9 in π are 4 due to the maximal increasing consecutive subsequences 2348 and 5679. Note that 12348 is an increasing consecutive subsequence of length 5, but we do not count it because it starts at the first element of π.

We will be working with a more general class of clusters. Note that the definition of clusters extends to clusters with respect to a set of more than one pattern in the obvious way: we have a rooted tree with some highlighted instances of patterns in the set such that it is not possible to partition the vertices of the tree into two parts so that each highlighted instance has vertices in at most one part.

Definition 4.9. A *p-pseudo m-cluster of size n* with respect to π is an m-cluster of size n with respect to the patterns $\{\pi(1)\cdots\pi(k), \pi(2)\cdots\pi(k), \ldots, \pi(p)\cdots\pi(k)\}$ such that each highlighted instance of a pattern that is not π contains the root of the cluster when the cluster is viewed as a tree.

Informally, p-pseudo m-clusters of size n for π can be thought of as the subtree of a vertex v in a cluster for π, where the parts of the highlighted instances containing the ancestors of v are "cut off" and no such instance has more than p vertices cut off.

Definition 4.10. A *primitive* m-cluster of size n with respect to π is an m-cluster of size n in which all highlighted instances of π contain the root of the cluster. A *primitive* p-pseudo m-cluster of size n with respect to π is a p-pseudo m-cluster of size n in which all highlighted instances of a pattern contains the root of the cluster.

When speaking generally, we may omit some of m, n, p. We note that (primitive) 1-pseudo clusters are simply (primitive) clusters, and that for $i < j$, all (primitive) i-pseudo clusters are (primitive) j-pseudo clusters. In a primitive m-cluster, there are exactly m leaves, all at depth k.

Definition 4.11. In a primitive pseudo cluster C, for each vertex v that is not the root of the cluster, the *height* of v in C is the largest integer m such that there exist vertices $v_1, \ldots, v_m = v$, none of which is the root of C, such that the labels of $v_1, \ldots, v_m$ are increasing and for all $1 \le i < m$, v_{i+1} is v_i's child. In other words, the height of a vertex is the length of the longest consecutive sequence of vertices different from the root of C with increasing labels.

Note that if v is the ith vertex in a highlighted instance of π in the cluster, then the height of v in C is the same as the height of $\pi(i)$ in π. However, the notions of height do not necessarily coincide for vertices not in a highlighted instance of π.

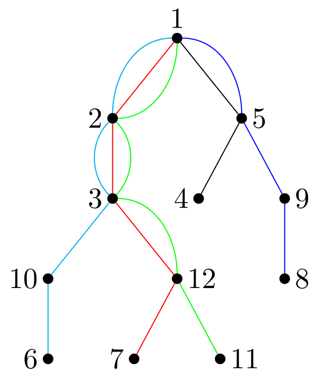

Figure 9. A primitive 3-pseudo 5-cluster of size 12 with respect to 12354. The instances are the colored paths on $\{1,2,3,10,6\}$, $\{1,2,3,12,7\}$, $\{1,2,3,12,11\}$, $\{1,5,4\}$, and $\{1,5,9,8\}$. The blue and black paths, respectively on $\{1,5,9,8\}$ and $\{1,5,4\}$, are instances of 2354 and 354.

Definition 4.12. Two permutations π and π' of length k are *p-equivalent* if for all nonnegative integers ℓ, m, n with $\ell \in [n]$ and functions $f : [n] \setminus \{\ell\} \to [k]$, the number of primitive p-pseudo m-clusters C of size n whose root is labeled ℓ such that for all $i \in [n] \setminus \{\ell\}$, $f(i)$ is equal to the height of the vertex with label i in C, is equal for π and π'.

Definition 4.13. An *ordered primitive cluster* is a primitive cluster whose underlying rooted tree is an ordered rooted tree, in the sense that the children of any vertex are ordered, so if labels between the children of a vertex are permuted a different labeled rooted tree is obtained.

Thus far, we have been working with unordered forests. We will only consider ordered (i.e. planar) forests in the context of ordered primitive clusters. Note that given the structure of the underlying tree, the ratio of the number of ordered clusters to the number of unordered clusters is the number of automorphisms of the tree, so the notions are closely related.

Definition 4.14. Two permutations π and π' of length k are *primitive structure equivalent* if for all nonnegative integers m and n and unlabeled rooted trees T on n vertices with m leaves, the number of labellings of T that are ordered primitive clusters for π is equal to the number of labellings of T that are ordered primitive clusters for π'.

The family of equivalences we will describe involve a class of permutation we call the *grounded permutations*.

Definition 4.15. A permutation $\pi = \pi(1)\cdots\pi(k) \neq 1\cdots k$ is *grounded* if there exists a positive integer m such that $\pi(1)\cdots\pi(m) = 1\cdots m$ and $\pi(2)\cdots\pi(k)$ avoids the pattern $1\cdots(m+1)$ in the consecutive sense of permutation pattern avoidance.

In other words, a pattern π is grounded if it is not the identity and its max height is at most its streak. Since the max height of π is at least its streak when $\pi \neq 1\cdots k$, we can equivalently say that π is grounded if it is not equal to $1\cdots k$ and its max height is equal to its streak. The heights of the elements in π are then bounded above by its streak, hence the term grounded.

The description of the family of equivalences proceeds with a long series of lemmas. We first describe a decomposition of clusters for grounded π into pseudo clusters.

Definition 4.16. Suppose that π is a grounded pattern with streak s. Then for any $1 \leq p \leq s$, any p-pseudo m-cluster C with respect to π, we define the *standard decomposition* of C as $(P, \mathcal{Q})$, where

(1) P is the primitive p-pseudo cluster consisting of all highlighted instances of patterns that contain the root of C, and
(2) $\mathcal{Q}$ is the collection of the pseudo clusters C_v over all non-root vertices v of P. Here, we define C_v to be the h-pseudo cluster rooted at v consisting of the vertices in C whose lowest ancestor in P is v, where H is the height of v in P.

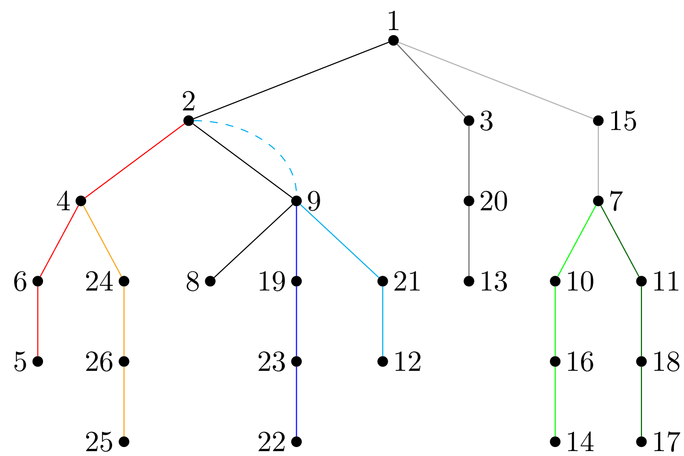

FIGURE 10. The standard decomposition of a 2-pseudo cluster with respect to 1243. The highlighted instances on $\{1,2,9,8\}$, $\{1,3,20,13\}$, and $\{1,15,7\}$ in the primitive portion P are colored with different shades of grey. The C_v are colored in different shades of red, blue, and green. The red pseudo cluster consists of instances $2,4,6,5$ and $4,24,26,25$, the blue pseudo cluster consists of instances $9,19,23,22$ and $9,21,12$, and the green pseudo cluster consists of instances $7,10,16,14$ and $7,11,18,17$. Note that the vertex labeled 2 is not in the blue pseudo cluster corresponding to the vertex labeled 9 in the decomposition, but it is a part of the highlighted instance $2,9,21,12$ in the whole pseudo cluster.

We will show the existence and uniqueness of the standard decomposition in Lemma 4.18, along with some other properties. To do so, we first show an intermediate result.

Lemma 4.17. *Let π be a grounded pattern with streak s. In any p-pseudo cluster with respect to π with $p \le s$, the root has label* 1.

Proof. Suppose that we have two intersecting instances $v_1,\dots,v_k$ and $w_1,\dots,w_k$ of π, where v_i, w_i are labeled vertices and $v_1,\dots,v_k$ and $w_1,\dots,w_k$ are paths. Suppose that $v_i = w_j$ is the last vertex that the two instances share in common, i.e. the largest i,j with $v_i = w_j$ so that $v_{i+1} \ne w_{j+1}$, if both v_{i+1} and w_{j+1} exist. Without loss of generality, suppose that $i \ge j$, which means that $w_1 \in \{v_1,\dots,v_k\}$. We assume that $i > j$, or equivalently that $v_1 \ne w_1$. In this case, we must have that j is at most the height h of $\pi(i)$ in π. Indeed, note that we must have $j \le s$ because the first $s+1$ terms of π are increasing. If $j > s$, then $w_1,\dots,w_{s+1}$ would be a consecutive subsequence of $v_2,\dots,v_k$ with increasing labels, so some v_q has height at least $s+1$, a contradiction. Since the first s terms of π are increasing, $w_1,\dots,w_j$ have increasing labels, so since $v_i = w_j$, j is upper bounded by the height of $\pi(i)$ in π. Since $j \le s$ and the first s terms of π are $1,\dots,s$, this means that the labels of $w_{j+1},\dots,w_k$ are greater than the labels of $w_1,\dots,w_j$, which are greater than the labels of $v_1,\dots,v_{\min(s,i-1)}$ (which are in turn smaller than the labels of $v_{\min(s+1,i)},\dots,v_k$). Thus, the smallest labels among the two instances are the labels of $v_1,\dots,v_{\min(s,i-1)}$. In particular, this means that if we instead took an instance $v_p, \dots, v_k$ of $\pi(p) \cdots \pi(k)$ for $p \le s$ and an instance $w_1, \dots, w_k$ of π such that $w_1 \in \{v_p, \dots, v_k\}$, then v_p would have the smallest label among all vertices. Also, if we instead took an instance $v_{p_1}, \dots, v_k$ of $\pi(p_1) \cdots \pi(k)$ and an instance $w_{p_2}, \dots, w_k$ of $\pi(p_2) \cdots \pi(k)$ with $p_1, p_2 \le s$ with $v_{p_1} = w_{p_2}$, then $v_{p_1} = w_{p_2}$ would have the smallest label among all vertices. In other words, whenever two possibly truncated (by no more than s elements in the beginning) instances of π intersect, where truncated instances must begin at the root of the tree formed by the two overlapping paths, the root will have the smallest label. This automatically implies that in any p-pseudo cluster with $p \le s$, the root has the smallest label. □

Lemma 4.18. *Let π be a grounded pattern with streak s, C be a p-pseudo m-cluster with respect to π, and r be the root vertex of C. Then the standard decomposition $(P, \mathcal{Q})$ of C exists and is unique. Furthermore, every vertex of C lies in P or a pseudo cluster in $\mathcal{Q}$, the pseudo clusters in $\mathcal{Q}$ are disjoint, and the only vertex that lies in both P and C_v is v for all $v \in P \setminus \{r\}$. Every highlighted instance of a pattern in C corresponds to a highlighted instance of a pattern in exactly one of P and the pseudo clusters in $\mathcal{Q}$, and corresponding patterns end on the same vertex. The total number of highlighted instances of patterns throughout P and the pseudo clusters in $\mathcal{Q}$ is thus m. Finally, in P and the pseudo clusters in $\mathcal{Q}$, the root of the pseudo cluster has the smallest label.*

Proof. Note that P is uniquely determined by its definition, as it is the cluster consisting of the highlighted instances of patterns that contain the root of C. For each non-root vertex v of P, we consider the tree T_v of elements whose lowest ancestor that lies in P is v. In this way, all of the T_v are disjoint, all vertices of C lie in one of P and T_v for $v \notin P$, and the only vertex that lies in both P and T_v is v. The T_v will serve as the underlying trees of the pseudo clusters C_v.

We now highlight the instances of patterns in P and the T_v. Every highlighted instance of a pattern in C is contained entirely in one of P and the T_v or contains vertices in both P and T_v for some v. If a highlighted instance of a pattern in C is contained entirely in one of P and the T_v, then we highlight it in the corresponding tree, whichever one of P and the T_v it is. Note that this is consistent with the way we defined P to be the pseudo cluster of highlighted instances of patterns containing the root of C, as no highlighted instance of a pattern that does not contain the root of C can be entirely contained in P. It is also clear that P is a primitive p-pseudo cluster. For every other highlighted instance $u_1, \dots, u_k$ of a pattern, which is necessarily a highlighted instance of π since it does not contain the root of C, we consider the maximal index j such that $u_j \in P$. By the same arguments as in the previous paragraph, we must have that j is at most the height h of u_j in P. We then highlight the instance $u_j, \dots, u_k$ of $\pi(j) \cdots \pi(k)$ in T_{u_j}. As $j \le h$, this does not violate the condition that C_{u_j} is to be an h-pseudo cluster. Furthermore, it is also clear that the highlighted instances of patterns in T_v are all connected, which means that they form a cluster C_v. In this way, all highlighted instances of patterns in C have a corresponding highlighted instance of a pattern in one of P and the pseudo clusters in $\mathcal{Q}$, and corresponding instances end on the same vertex. By Lemma 4.17, in all of P and the pseudo clusters in $\mathcal{Q}$, the root has the smallest label, so all of the claimed conditions have been proven.

Finally, note that given any standard decomposition $(P, \mathcal{Q})$, it is easy to recover the original pseudo cluster C. The underlying tree of C is simply the union of the underlying trees of P and the pseudo clusters in $\mathcal{Q}$. We keep any highlighted instances of patterns containing the root or highlighted instances of π. All other highlighted instances of patterns are instances $v_j, \dots, v_k$ of $\pi(j) \cdots \pi(k)$ that begin at a vertex v_j of P. In this case, we have by construction that v_j is the only vertex in the instance that is in P and that j is at most the height h of v_j in P, as C_{v_j} is an h-pseudo cluster. Thus, if $u_1, \dots, u_h = v_j$ is the path in P of increasing labels, we can highlight $u_{h-j+1}, \dots, u_h = v_j, \dots, v_k$ as the corresponding highlighted instance of π in C. Doing so for all highlighted instances of patterns in P and the T_v recovers the original pseudo cluster C. Consequently, p-pseudo m-clusters C are in bijection with standard decompositions $(P, \mathcal{Q})$. □

The next step is to use Lemma 4.18 to show that p-equivalence implies strong equivalence for grounded patterns. Though the proof is via an explicit recursion, it is possible to make it bijective. Informally, Lemma 4.19 is true because clusters can be decomposed into primitive pseudo clusters, and p-equivalence implies that the number of primitive pseudo clusters and the number of ways to "glue together" primitive pseudo clusters to form a cluster are the same.

Lemma 4.19. *Suppose that π and π' are grounded patterns of length k of the same streak s. If π and π' are p-equivalent for all $1 \le p \le s$, then the number of p-pseudo m-clusters of size n with respect to π and π' are equal for all $m, n, 1 \le p \le s$. In particular, by taking $p = 1$ we have that π and π' have the same forest cluster numbers and are thus strongly equivalent by Theorem 4.3.*

Proof. Let $r_{m,n,p}$ be the number of p-pseudo m-clusters of n vertices with respect to π, and let $r'_{m,n,p}$ denote the same with respect to π'. Our approach, mimicking the proof of Theorem 4.3 that was given in [17], is to show that $r_{m,n,p}$ and $r'_{m,n,p}$ satisfy the same recursion, or equivalently by induction on n. The base cases of $n = 0$ and $n = 1$ are clear, as for $n = 0$ there is only the empty pseudo cluster and for $n = 1$ we have that $r_{m,1,p} = r'_{m,1,p} = 0$. For the recursion or the inductive step, we make use of the standard decomposition.

For $m, n \ge 0, 1 \le p \le s, f : \{2, \dots, n\} \to [s]$, let $B(m, n, p, f)$ denote the number of primitive p-pseudo m-clusters C of size n such that for all $1 < i \le n$, the height of the vertex with label i in C is $f(i)$. Note here that we have shown before that such a pseudo cluster must have the smallest label at its root and that the height of every non-root vertex is at most s, so it suffices to take f to be a function from $\{2, \dots, n\}$ to $[s]$, as otherwise $B(m, n, p, f) = 0$. We define $B'(m, n, p, f)$ analogously for π' and note that $B(m, n, p, f) = B'(m, n, p, f)$ by assumption of the p-equivalence of π and π'.

Suppose that L_P is a subset of $\{2, \dots, n\}$ with $|L_P| = q$ and $L_P = \{\ell_1, \dots, \ell_q\}$ with $\ell_1 < \cdots < \ell_q$. Let $L_1 \sqcup \cdots \sqcup L_q$ be a partition of $\{2, \dots, n\}$ such that $L_i \subseteq \{\ell_i, \dots, n\}$ and $L_P \cap L_i = \{\ell_i\}$ for all $1 \le i \le q$. For simplicity, let $\mathcal{L}_n$ be the set of all pairs $(L_P, (L_1, \dots, L_q))$ satisfying these conditions. Given integers $m, n > 0$, $1 \le p \le s$, $(L_P, (L_1, \dots, L_q)) \in \mathcal{L}_n$, nonnnegative integers $m_0, \dots, m_q$ summing to m, and a function $f : L_P \to [s]$, let $N(m, n, p, L_P, (L_1, \dots, L_q), (m_0, \dots, m_q), f)$ denote the number of p-pseudo m-clusters C of size n with respect to π such that if the standard decomposition of C is $(P, \mathcal{Q})$, then

- P has label set $L_P \cup \{1\}$ (the label 1 is automatically the root, as all pseudo clusters have root label 1), the height of the vertex labeled with ℓ_i in P is $f(\ell_i)$, and there are exactly m_0 highlighted instances of a pattern completely in P, and
- a pseudo cluster C_v in $\mathcal{Q}$ with root v has label set L_i if v has label ℓ_i and there are exactly m_i highlighted instances of a pattern in C_v,

and let $N'(m, n, p, L_P, (L_1, \dots, L_q), (m_0, \dots, m_q), f)$ denote the analogous quantity for π'.

By summing over all possible $L_p, (L_1, \dots, L_q), (m_0, \dots, m_q), f$, we have that

$$r_{m,n,p} = \sum_{(L_P,(L_1,\dots,L_q))\in\mathcal{L}_n} \sum_{m_0+\cdots+m_q=m} \sum_{f:L_P\to[s]} N(m, n, p, L_P, (L_1, \dots, L_q), (m_0, \dots, m_q), f).$$

On the other hand, we have that

$$N(m, n, p, L_P, (L_1, \dots, L_q), (m_0, \dots, m_q), f) = B(m_0, q + 1, p, f \circ g_{L_P}) \prod_{|L_i|>1 \text{ or } m_i>0} r_{m_i,|L_i|,f(i)}$$

where $g_{L_P} : \{2, \dots, q + 1\} \to L_P$ is the order-preserving bijection between $\{2, \dots, q + 1\}$ and L_P. Indeed, by Lemma 4.18, pseudo clusters C satisfying the conditions in the bullets above are in bijection with the standard decompositions $(P, \mathcal{Q})$ that satisfy the conditions. We may then independently decide the pseudo cluster P and the pseudo clusters in $\mathcal{Q}$ and multiply the results together. For P, we want to count the number of primitive p-pseudo m_0-clusters of size $q+1$ with height function $f \circ g_{L_P}$ as only the relative order matters for counting pseudo clusters. For $C_v \in \mathcal{Q}$

with root v, when $|L_i| = 1$ and $m_i = 0$, this corresponds to not attaching any extra highlighted instances of patterns to v, of which there is only one way. Otherwise, this is by definition counted by $r_{m_i,|L_i|,f(i)}$, demonstrating the identity.

Consequently we have that

$$r_{m,n,p} = \sum_{(L_P,(L_1,\dots,L_q))\in\mathcal{L}_n} \sum_{m_0+\cdots+m_q=m} \sum_{f:L_P\to[s]} B(m_0, q+1, p, f\circ g_{L_P}) \prod_{|L_i|>1 \text{ or } m_i>0} r_{m_i,|L_i|,f(i)}$$

as well as

$$r'_{m,n,p} = \sum_{(L_P,(L_1,\dots,L_q))\in\mathcal{L}_n} \sum_{m_0+\cdots+m_q=m} \sum_{f:L_P\to[s]} B'(m_0, q+1, p, f\circ g_{L_P}) \prod_{|L_i|>1 \text{ or } m_i>0} r'_{m_i,|L_i|,f(i)}$$

by applying the same arguments to π'. We have assumed that π and π' are p-equivalent. Thus, $B(m_0, q+1, p, f\circ g_{L_P}) = B'(m_0, q+1, p, f\circ g_{L_P})$ for all $m_0, q, 1 \le p \le s, f\circ g_{L_P} : \{2,\dots,q+1\} \to [s]$, so $r_{m,n,p}$ and $r'_{m,n,p}$ satisfy the same recursion and the lemma follows. □

The next step is show a way to "boost" 1-equivalence between certain (not necessarily grounded) patterns to p-equivalence between longer patterns. It allows us to more easily find examples of the stronger condition of p-equivalence.

Lemma 4.20. *Suppose that π and π' are 1-equivalent permutations of length k with the property that $\pi(1) > \pi(2)$ and $\pi'(1) > \pi'(2)$. Then π and π' have the same max height h, and for all ℓ, $\sigma = 1,\dots,\ell,\pi(1)+\ell,\dots,\pi(k)+\ell$ and $\sigma' = 1,\dots,\ell,\pi'(1)+\ell,\dots,\pi'(k)+\ell$ are p-equivalent for all $1 \le p \le \ell$. By Lemma 4.19, if $\ell \ge h$, then σ and σ' are strongly equivalent.*

Proof. To see that π and π' have the same max height, we consider clusters on k vertices with 1 instance of π or π'. These clusters must consist of a path with labels in the order of π or π'. The 1-equivalence of π and π' then implies that π and π' start with the same number a and that for all $i \neq a$, the height of i in π is equal to the height of i in π', so π and π' have the same max height.

The 1-equivalence of π and π' is equivalent to the existence of a root- and height-preserving bijection between primitive m-clusters of size n for π and primitive m-clusters of size n for π'. Here, root-preserving means that the label of the root is fixed and height-preserving means that for each label different from the label of the root, the height of the vertex having that label in the tree is fixed (though the vertex that has that label and even the underlying unlabeled forest structure may change). We specify that such a bijection should be both root-preserving and height-preserving because the height is technically only defined for non-root vertices. Let α be a bijection mapping primitive m-clusters C for π to primitive m-clusters C' for π'. By assumption, α has an inverse β, so α and β are inverse maps between primitive clusters on π and π'.

To prove the lemma, we will demonstrate a root- and height-preserving bijection between primitive p-pseudo m-clusters of size n for σ and primitive p-pseudo m-clusters of size n for σ' for all $1 \le p \le \ell$ based on α and β. Before we explain the bijection, we first make some observations about primitive pseudo clusters for σ and σ'. First, in any primitive pseudo cluster it is unnecessary to highlight instances of patterns, because the highlighted instances are always going to consist of the paths from the root to the leaves. In this way, highlighted instances correspond directly to the leaves of the underlying tree. Consider a primitive p-pseudo cluster C for $\sigma = \sigma(1)\cdots\sigma(k+\ell)$ with $1 \le p \le \ell$. For each instance $v_i,\dots,v_{k+\ell}$ of $\sigma(i)\cdots\sigma(k+\ell)$, color v_j with the color j for all $i \le j \le k+\ell$, so vertices may receive multiple colors. Note that all vertices colored with a number greater than $\ell+1$ are only colored by that number. Indeed, suppose that such a vertex v was colored by both a and b, where $a > b$ and $a > \ell+1$. Then note that the parent of v is colored by both $a-1$ and $b-1$. By taking parents and going up the tree, we may assume that $a = \ell+2$. We never end up at the root during this process since $p \le \ell$, so the root is only colored with numbers at most ℓ. Note that $\sigma(1) < \cdots < \sigma(\ell+1) > \sigma(\ell+2)$. Thus, if v is colored with $\ell+2$ then its label is smaller than its parent's label. But if v is colored with $b \leq \ell + 1$ then its label is greater than its parent's label, a contradiction.

Define the *fixed tree* T of C as the subtree of vertices whose colors are all at most $\ell + 1$. Let the *seeds* of T be the vertices that are labeled $\ell + 1$. Note that the root cannot be a seed as $p \leq \ell$. By considering the colors of the vertices, for each vertex u not in T, the lowest ancestor of u in T is a seed. For each seed v, we let T_v be the subtree of C consisting of v and the vertices outside of T for which v is the lowest ancestor in T, which necessarily contains vertices other than v. Then, C is the union of T and T_v for seeds $v \in T$, where the T_v are disjoint, and the only vertices in both T and T_v is v. Let $\mathcal{U}$ be the collection of T_v over all seeds $v \in T$, and call the decomposition of C into $(T, \mathcal{U})$ the *primitive decomposition* of C. Note that it is well-defined for both σ and σ', as σ' also satisfies all of the properties that we used throughout this paragraph.

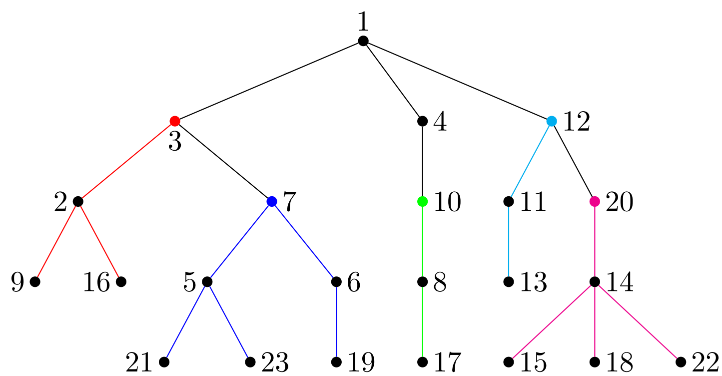

FIGURE 11. The primitive decomposition of a pseudo cluster with respect to 12435. The fixed tree consists of the vertices labeled with $1, 3, 4, 7, 10, 12, 20$. We have not highlighted pattern instances as they are just the paths from the root to the leaves. Instead, we have colored the primitive 435-clusters and their seeds. The red cluster is rooted at 3 and has instances on $\{3, 2, 9\}$ and $\{3, 2, 16\}$, the blue cluster is rooted at 7 and has instances on $\{7, 5, 21\}$, $\{7, 5, 23\}$, and $\{7, 6, 19\}$, the green cluster is rooted at 10 and has an instance on $\{10, 8, 17\}$, the cyan cluster is rooted at 12 and has an instance on $\{12, 11, 13\}$, and the magenta cluster is rooted at 20 and has instances on $\{20, 14, 15\}$, $\{20, 14, 18\}$, and $\{20, 14, 22\}$.

The last observation we need is that for a primitive p-pseudo cluster C for σ with $p \leq \ell$ with primitive decomposition $(T, \mathcal{U})$, all of the clusters T_v are primitive clusters for π. This is clear from the definition of the primitive decomposition. Now, given a primitive p-pseudo m-cluster C of size n for σ, we consider its primitive decomposition $(T, \mathcal{U})$. For each $T_v \in \mathcal{U}$, we apply the map α (using the relative order of the labels in T_v). The result is a primitive p-pseudo m-cluster C' of size n for σ'. This is as for each seed v, the labels of v's strict ancestors are all less than the labels of v's descendants. By replacing each root-to-leaf instance of π in T_v with a root-to-leaf instance of π', we replace each root-to-leaf instance of $i, \ldots, \ell, \pi(1) + \ell, \ldots, \pi(k) + \ell$ in C with a root-to-leaf instance of $i, \ldots, \ell, \pi'(1) + \ell, \ldots, \pi'(k) + \ell$. Since α fixes the root and the number of vertices and number of leaves by definition, the result is a p-pseudo m-cluster C' of size n for σ'. Furthermore, it is clear that C' has the same fixed tree (including labels) and seeds as C. Thus, if $(T, \mathcal{U}')$ is the primitive decomposition of C' where $\mathcal{U}' = \bigcup_{v \text{ seed}} T'_v$, then applying β to each T'_v transforms them back into T_v, giving the original cluster C. Thus, we have defined a bijection between pseudo clusters for σ and pseudo clusters for σ'. It remains to show that this bijection fixes the root and height of each vertex. The root is clearly fixed since it is a part of the fixed tree. Vertices and labels in the fixed tree are fixed, so their heights are as well. To finish, we note that for a vertex $w \neq v$ in T_v where v is a seed of the fixed tree T of C, the height of w in C is equal to the height of w in T_v. This is as $\sigma(\ell+1) > \sigma(\ell+2)$ due to $\pi(1) > \pi(2)$, so all of the children of a seed v in T_v have smaller labels than v. Thus, as α and β preserve the heights, so does the bijection we have defined. Finally, note that σ and σ' are grounded patterns of streak ℓ, so by Lemma 4.19 they are strongly equivalent. □

The next step is to obtain 1-equivalences from primitive structure equivalences. The rough idea is that if $x, \dots, y$ all have the same height in a pattern π, are not consecutive, and appear after $x-1$ and $y+1$, then we are able to permute the vertices in the layers of any primitive cluster corresponding to the positions of $x, \dots, y$ in π following primitive structure equivalence. This follows because the appearance of $x-1$ and $y+1$ before $x, \dots, y$ "pin down" the values of those layers so that they are able to be freely permuted without affecting the relative order with the rest of the forest.

Lemma 4.21. *Let π be a permutation of length k, and consider indices $1 < a_1 < \dots < a_i \leq k$ with $a_{j+1} - a_j > 1$ for all $1 \leq j < i$ (i.e. the indices are not consecutive) such that the heights of $\pi(a_j)$ in π are equal for all j. Suppose that*

- *$\{\pi(a_1), \dots, \pi(a_i)\} = \{x, \dots, y\}$ for some $x < y$ and that*
- *$x-1$ and $y+1$ (if they are in $[n]$) appear before $\pi(a_1)$ in π.*

Let σ and σ' be permutations of length i that are primitive structure equivalent. Suppose that σ and $\pi(a_1), \dots, \pi(a_i)$ are in the same relative order, and let $\pi'(a_1), \dots, \pi'(a_i)$ be the permutation of $\pi(a_1), \dots, \pi(a_i)$ that is in the same relative order as σ'. Then π is 1-equivalent to π', where π' is π with the subsequence $\pi(a_1), \dots, \pi(a_i)$ is replaced by $\pi'(a_1), \dots, \pi'(a_i)$, i.e. define $\pi'(a_j)$ as before and $\pi'(j) = \pi(j)$ for $j \notin \{a_1, \dots, a_i\}$.

Proof. We will be defining a root- and height-preserving bijection between primitive clusters for π and π'. In fact, our bijection will be structure-preserving, so it will only permute certain non-root labels, specifically the labels of the vertices with depth in $\{a_1, \dots, a_i\}$. Note that the conditions in the lemma statement are symmetric with respect to π and π', i.e. they are satisfied for π if and only if they are satisfied for π' (this holds for any permutation π' that starts with π and permutes some of $\pi(a_1), \dots, \pi(a_i)$). This is because the elements of π outside of $\pi(a_1), \dots, \pi(a_i)$ are either greater than all of $\pi(a_1), \dots, \pi(a_i)$ or less than all of $\pi(a_1), \dots, \pi(a_i)$. None of the elements $\pi(a_1), \dots, \pi(a_i)$ are adjacent, so permuting the $\pi(a_j)$ does not impact the equal heights condition. It in fact preserves the heights of all $\pi(j)$ in π. All other conditions are defined symmetrically so are always satisfied when $\pi(a_1), \dots, \pi(a_i)$ are permuted.

Note that by the conditions in the statement, any structure-preserving bijection between primitive clusters for π and π' of the form described in the previous paragraph (where only labels of vertices with depth in $\{a_1, \dots, a_i\}$ are permuted) will preserve the height of all vertices. Let φ be the structure-preserving bijection from ordered primitive clusters for σ to ordered primitive clusters for σ'. For each vertex v of depth a_1 in a primitive cluster C of π, we will permute labels within the subtree of C rooted at v. The conditions in the lemma statement guarantee that if S is the set of all of the descendants of v that have depth in $\{a_1, \dots, a_i\}$, then any descendant or ancestor w of v will either be in S, have label either greater than all of the labels of vertices in S, or have label less than all of the labels of vertices in S. Indeed, suppose that the label of $w \notin S$ is greater than the label of v. Consider the ancestor u of v of depth j, where $\pi(j) = y+1$, which exists by the conditions in the statement. We know that the label of w is greater than the label of u. On the other hand, the label of u is greater than all of the labels in S. Thus, the label of w is greater than all of the labels in S. The case in which the label of w is less than the label of v is analogous. Thus, if the labels of the vertices in S are permuted so that every instance of σ (here we are looking at the vertices of depth $a_1, \dots, a_i$ in a path starting at v) is replaced with an instance of σ', then every instance of π with the a_1st vertex at v will be replaced with an instance of π'. If we do so for all v of depth a_1, we will have replaced all instances of π with an instance of π'.

If $a_i < k$, then this replacement is precisely given by φ. Indeed, if we connect every vertex u_1 and u_2 in S of depths a_j and a_{j+1} such that u_1 is an ancestor of u_2 for all $1 \le j < i$ (i.e. the contraction of the tree to S), then we give S the structure of an ordered primitive cluster for σ. This is because every path from a root to a leaf in this contraction is an instance of σ, but the children of each vertex are distinguished and their order matters. In the original graph, non-leaf vertices of S have children that are not in S. Such children that are not in S distinguish non-leaf vertices in from their siblings in S, so we indeed have an ordered primitive cluster for σ. Applying φ and φ^{-1} to each subtree rooted at depth a_1 gives us a desired bijection between primitive clusters for π and primitive clusters for π'.

If $a_i = k$, then we need to modify the previous argument to account for the fact that the leaves of the contraction to S do not have children in the original tree and may not have a distinguished order. The order of leaves $v_1, \dots, v_m$ in S does not matter if and only if $v_1, \dots, v_m$ have a the same parent in the original tree. Thus, given an unlabeled tree structure, there is a fixed number M that counts the number of ways to permute the labels of the leaves of a cluster that do not change the cluster. If we fix an unlabeled tree structure and enrich clusters for π and π' with one of the M orders for their leaves, then the bijection from the previous paragraph directly applies. Thus, given a fixed unlabeled tree structure, the order-enriched clusters for π and π' are in bijection. This means that the number of clusters for π with a given tree structure is equal to the number of clusters for π' with a given tree structure because the fixed number M of orders on the leaves is the same for π and π'. Summing over all possible unlabeled tree structures gives the result. □

Example 4.22. Figure 12 gives an example of the bijection in the proof of Lemma 4.21. In the pattern 1253764, 3 and 4 have the same height of 1, and 2 and 5 appear before them. Thus, we may swap them to get that 1253764 and 1254763 are 1-equivalent, because the patterns 12 and 21 are primitive structure equivalent. A structure-preserving bijection between ordered primitive clusters for 12 and 21 is given by swapping the largest and smallest labels. In the primitive cluster for 1253764, the vertices of depth 4 and 7, corresponding to the positions of 3 and 4 in the pattern, form ordered primitive clusters for 12, which are colored red and blue. When the bijection between ordered primitive clusters for 12 and 21 is applied, we obtain a corresponding primitive cluster for 1254763. Note that technically, the order of the vertices labeled 6 and 7 does not matter. If we enrich the cluster with an order between the vertices labeled 6 and 7, as well as an order between the vertices labeled 10 and 11, then each cluster for 1253764 is counted $M = 4$ times. Since the corresponding cluster for 1254763 has the same structure, each cluster for 1253764 of that structure is also counted $M = 4$ times.

The final step in our construction is to find examples of primitive cluster equivalence. Complementation and appending $k+1$ give some primitive cluster equivalences which can be explicitly described, but there is also a family of recursively defined equivalences constructed in the same way as in Lemma 4.21.

Lemma 4.23. *A permutation π is primitive structure equivalent to its complement. If π and π' are permutations of length k that are primitive structure equivalent, then $k+1, \pi$ and $k+1, \pi'$ are primitive structure equivalent. Finally, consider indices $1 < a_1 < \dots, a_i \le k$ with the property that $\{\pi(a_1), \dots, \pi(a_i)\} = \{x, \dots, y\}$ for some $x < y$ and such that $x-1$ and $y+1$ (if they are in $[n]$) appear before $\pi(a_1)$ in π. Let σ and σ' be permutations of length i that are primitive structure equivalent. Suppose that σ and $\pi(a_1), \dots, \pi(a_i)$ are in the same relative order, and let $\pi'(a_1), \dots, \pi'(a_i)$ be the permutation of $\pi(a_1), \dots, \pi(a_i)$ that is in the same relative order as σ'. Then π and π' are primitive structure equivalent, where π' is π but the subsequence $\pi(a_1), \dots, \pi(a_i)$ is replaced by $\pi'(a_1), \dots, \pi'(a_i)$, i.e. define $\pi'(a_j)$ as before and $\pi'(j) = \pi(j)$ for $j \notin \{a_1, \dots, a_i\}$.*

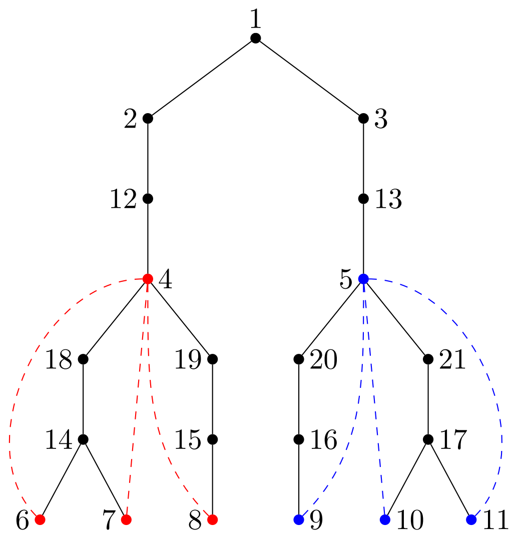

FIGURE 12. The patterns 1253764 and 1254763 are 1-equivalent because the patterns 12 and 21 are primitive structure equivalent.

Proof. The first part follows trivially by complementation of the labels of each ordered primitive cluster. For the second part, suppose we have a structure-preserving bijection φ between ordered primitive clusters for π and ordered primitive clusters for π'. Any ordered primitive cluster C for $k+1, \pi$ or $k+1, \pi$ must have the largest label at the root. We can apply φ to every subtree rooted at a child of the root of C, which is an ordered primitive cluster for π. This gives a structure-preserving bijection between ordered primitive clusters for $k+1, \pi$ and $k+1, \pi'$, as desired. The proof of third part is essentially the same as the proof of Lemma 3.17. We permute the labels of vertices of depth in $\{a_1, \ldots, a_i\}$, where each subtree of each vertex of depth a_1 is permuted separately. Each subtree is permuted by considering the contraction of the tree to the vertices of depth $\{a_1, \ldots, a_i\}$ and applying the structure-preserving bijection given by the primitive structure equivalence of σ and σ'. In contrast with the previous part, we do not require that the indices are not consecutive, because we do not need the vertex heights to be preserved in this case. Since we are already working with ordered primitive clusters, the vertices of depths in $\{a_1, \ldots, a_i\}$ already have the structure of an ordered rooted forest, and so there is no issue when permuting vertices of consecutive depths. □

Note that by induction, the first and second parts of Lemma 4.23 give the primitive structure equivalence between all permutations of $[k]$ in which every number is either greater than or less than all of the numbers that come after it. These can alternatively be described as the $\{213, 231\}$ avoiding permutations, in the classical non-consecutive sense for permutations, and there are 2^{k-1} such permutations, as shown for example in [30]. The third part of the lemma gives other primitive structure equivalences in a recursive sort of construction, but it seems hard to explicitly describe in general the resulting primitive structure equivalences.

Lemmas 4.19, 4.20, and 4.21 reduce the problem of finding strong equivalences to the problem of finding p-equivalences for all $1 \le p \le s$ to the problem of finding 1-equivalences to the problem of finding primitive structure equivalences, many of which are described in Lemma 4.23. These primitive structure equivalences can be passed back up through the chain of lemmas to prove many strong equivalences. It is possible that there are permutations that satisfy the conditions of each lemma different from the ones given by the next lemma, but we are not aware of such methods. Doing so successfully would yield new strong equivalences.

Example 4.24. Searching through all permutations of length $k \leq 5$ for 1-equivalences given by Lemmas 4.20, 4.21, and 4.23, we find the following 1-equivalences:

- 3142, 3241
- 31425, 31524, 32415, 32514
- 31452, 32451
- 31542, 32541
- 52314, 52413
- 43152, 43251
- 53142, 53241

These give rise to the following infinite families of equivalences by Lemma 3.16:

- $125364 \overset{c}{\sim} 125463, 1236475 \overset{c}{\sim} 1236574, \ldots$
- $1253647 \overset{c}{\sim} 1253746 \overset{c}{\sim} 1254637 \overset{c}{\sim} 1254736,$
  $12364758 \overset{c}{\sim} 12364857 \overset{c}{\sim} 12365748 \overset{c}{\sim} 12365847, \ldots$
- $12364785 \overset{c}{\sim} 12365784, 123475896 \overset{c}{\sim} 123476895, \ldots$
- $1253764 \overset{c}{\sim} 1254763, 12364875 \overset{c}{\sim} 12365874, \ldots$
- $1274536 \overset{c}{\sim} 1274635, 12385647 \overset{c}{\sim} 12385746, \ldots$
- $1265374 \overset{c}{\sim} 1265473, 12376485 \overset{c}{\sim} 12376584, \ldots$
- $1275364 \overset{c}{\sim} 1275463, 12386475 \overset{c}{\sim} 12386574, \ldots$

These are in fact strong equivalences, but we have written $\overset{c}{\sim}$ instead of $\overset{sc}{\sim}$. Notably, many of these pairs are not c-Wilf equivalent in terms of permutations, specifically the non-overlapping permutations with different ending terms due to [6, Theorem 1.9]. It seems difficult to describe all 1-equivalences given by the Lemma 4.21, but in principle all of the 1-equivalences between permutations of length k given by Lemma 4.21 can can be computed by checking all of the permutations of length k.

Remark 4.25. The equivalences we have obtained do not include the equivalence $1324 \overset{c}{\sim} 1423$ found by Garg and Peng. Indeed, 1324 and 1423 are not grounded since each of these patterns have streak 1, but the height of the last element in each pattern is 2. This equivalence appears to be of a more sporadic nature than the ones in our construction, and we were not able to find a way to generalize the proof of this equivalence given in [17]. There may be more sporadic equivalences, but we were not able to find any.

Remark 4.26. It is possible to describe explicit bijections between π and π' clusters by combining all of the bijections used in each step of the construction and modifying the proof to Lemma 4.19 into a recursive bijection. Such bijections are not structure-preserving and do not have analogues for c-Wilf equivalence in permutations, which somewhat explains this very different behavior that emerges in the case of forests. For example, the strong equivalence between 125364 and 125463 can be shown via a bijection between clusters. Informally, the bijection proceeds by "swapping 3 and 4" while also moving instances that are rooted at 3 or 4 along. Here, by 3 or 4 we mean the fourth or sixth vertices in highlighted instances corresponding to where 3 and 4 are located in the patterns. One way to view this is that we cut the branches of the tree stemming from 3 or 4, swap the labels of the 3 and 4, and replant the branches we cut to where the corresponding label moves. This process of cutting and replanting is done throughout the whole cluster, moving top-down, so that all instances of 125364 are replaced by instances of 125463 and vice-versa. The replanting step works because the branches we cut form a pseudo cluster whose root has the smallest label (Lemma 4.18). These branches can be replanted at any vertex of the same height, so long as the new location does not change the relative order of labels between the cut branch and the path from the root to the new location (Lemma 4.21). Figure 13 shows an example of this.

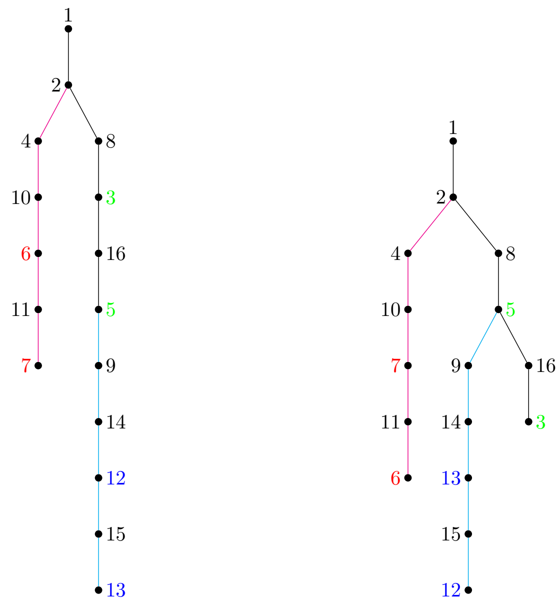

FIGURE 13. Corresponding clusters for 125364 and 125463. Corresponding instances and labels that are swapped are colored. The instances on $\{1, 2, 8, 3, 16, 5\}$, $\{2, 4, 10, 6, 11, 7\}$, and $\{5, 9, 14, 12, 15, 13\}$ on the left correspond to the instances on $\{1, 2, 8, 5, 16, 3\}$, $\{2, 4, 10, 7, 11, 6\}$, and $\{5, 9, 14, 13, 15, 12\}$ on the right. The swapped pairs of labels are 3 and 5, 6 and 7, and 12 and 13.

With this family of equivalences, we are now able to prove Theorems 1.6 and 1.5.

Proof of Theorem 1.5. We now construct a c-forest-Wilf equivalence class of size at least 2^{n-4} among patterns of length n for $n \geq 6$. The cases of $n \leq 5$ follow trivially by complementation. We first give a large class of 1-equivalences. Note that in a classically $\{213, 231\}$-avoiding permutation, every term is either greater than or less than all of the terms following it. If $n = 2k + 1$ is odd, we consider classically $\{213, 231\}$-avoiding permutations $a_1, \ldots, a_{k-1}$ and $b_1, \ldots, b_{k-1}$ of $1, \ldots, k-1$ and $k+1, \ldots, 2k-1$. Then note that in $k, a_1, b_1, \ldots, a_{k-1}, b_{k-1}$ the heights of $a_1, b_1, \ldots, a_{k-1}, b_{k-1}$ are $1, 2, \ldots, 1, 2$ so by Lemma 4.21, $k, a_1, b_1, \ldots, a_{k-1}, b_{k-1}$ over all classically $\{213, 231\}$-avoiding permutations $a_1, \ldots, a_{k-1}$ and $b_1, \ldots, b_{k-1}$ are all 1-equivalent. Thus, over all classically $\{213, 231\}$-avoiding permutations $a_1, \ldots, a_{k-1}$ and $b_1, \ldots, b_{k-1}$ of $1, \ldots, k-1$ and $k+1, \ldots, 2k-1$, the patterns $1, 2, k+2, a_1+2, b_1+2, \ldots, a_{k-1}+2, b_{k-1}+2$ all lie in the same c-forest-Wilf equivalence class by Lemma 4.20. There are $2^{k-2} \cdot 2^{k-2} = 2^{2k-4} = 2^{n-5}$ such permutations. If $n = 2k$ is even, we consider classically $\{213, 231\}$-avoiding permutations $a_1, \ldots, a_{k-1}$ and $b_1, \ldots, b_{k-2}$ of $1, \ldots, k-1$ and $k+1, \ldots, 2k-2$. By the same argument as before, $k, a_1, b_1, \ldots, a_{k-2}, b_{k-2}, a_{k-1}$ over all classically $\{213, 231\}$-avoiding permutations $a_1, \ldots, a_{k-1}$ and $b_1, \ldots, b_{k-2}$ are all 1-equivalent. Thus the patterns $1, 2, k+2, a_1+2, b_1+2, \ldots, a_{k-2}+2, b_{k-2}+2, a_{k-1}+2$ over all $\{213, 231\}$-avoiding permutations $a_1, \ldots, a_{k-1}$ and $b_1, \ldots, b_{k-2}$ of $1, \ldots, k-1$ and $k+1, \ldots, 2k-2$ all lie in the same c-forest-Wilf equivalence class. There are $2^{k-2} \cdot 2^{k-3} = 2^{2k-5} = 2^{n-5}$ such permutations. In both cases, we found 2^{n-5} such permutations, and taking complements gives a c-forest-Wilf equivalence class of size at least 2^{n-4}. Note that for $n = 6$ and $n = 7$ we get the equivalences from the first two examples in Example 4.24. $\square$

Notably in the case when n is even, all of the permutations are non-overlapping so we actually get permutations from $\frac{n-2}{2}$ different c-Wilf equivalence classes for permutations in the same equivalence class by [6, Theorem 1.9]. In contrast to classical forest-Wilf equivalence, beyond trivial symmetries c-forest-Wilf equivalence seems to be largely unrelated to c-Wilf equivalences for permutations.

Remark 4.27. We can use similar ideas to construct infinitely many families of c-forest-Wilf equivalence classes of exponential size. One way to do so is just to augment the above construction by increasing everything by 1 and appending a 1 at the beginning using Lemma 4.20. Another way is to divide into more "blocks." The above construction can be thought of as splitting $[2k+1]$ into the blocks $1,2 \mid 3,\dots,k+1 \mid k+2 \mid k+3,\dots,2k+1$. Then, we start the pattern with the size one blocks $1,2,k+2$ and alternate elements from the first large block and the second large block, placing numbers in the order of some pair of classically $\{213,231\}$-avoiding permutations. This can also be done with more blocks. For example, with three blocks the construction would take the following form: take $n=3k+1$ and split $[3k+1]$ into the blocks $1,2 \mid 3,\dots,k+1 \mid k+2 \mid k+3,\dots,2k+1 \mid 2k+2 \mid 2k+3,\dots,3k+1$. One can take classically $\{213,231\}$-avoiding permutations $a_1,\dots,a_{k-1}$ and $b_1,\dots,b_{k-1}$ and $c_1,\dots,c_{k-1}$ of the blocks $3,\dots,k+1$ and $k+3,\dots,2k+1$ and $2k+3,\dots,3k+1$, and then by Lemma 4.21, the patterns of the form $1,2,2k+2,k+2,a_1,c_1,b_1,\dots,a_{k-1},c_{k-1},b_{k-1}$ all belong to one c-forest-Wilf equivalence class. This gives a c-forest-Wilf equivalence class of size $2\cdot 2^{k-2}\cdot 2^{k-2}\cdot 2^{k-2}=2^{3k-5}=2^{n-6}$. This extends to all n by deleting last one or two elements based on the residue of n mod 3. One can also split $[3k+2]$ into the blocks $1,2,3 \mid 4,\dots,k+2 \mid k+3 \mid k+4,\dots,2k+2 \mid 2k+3 \mid 2k+4,\dots,3k+2$ and apply the same idea with the pattern $1,2,3,2k+3,k+3,a_1,b_1,c_1,\dots,a_{k-1},b_{k-1},c_{k-1}$. The reason we need $1,2,3$ in front now is that we have introduced elements of height 3 by putting it in the order a,b,c. This construction yields a c-forest-Wilf equivalence class of size $2^{3k-5}=2^{n-7}$. As we increase the number of blocks, we get more disjoint families of exponential size, but the constant factor decreases.

With our construction, we are now able to give the proof of Theorem 1.6. We first need the following lemmas.

Lemma 4.28. *For all $c<1$, there exists a positive integer k such that for sufficiently large n, at least $c^n n!$ permutations of $[n]$ avoid the consecutive pattern $1,\dots,k$.*

Proof. Divide a uniform random permutation of $[n]$ into consecutive blocks of k numbers, where the last block may have less than k numbers. If each block of k numbers (so the last block is excluded if it has less than k numbers) is not increasing, then the whole permutation avoids the consecutive pattern $1,\dots,2k$. The probability that this occurs is $\left(1-\frac{1}{k!}\right)^{\lfloor n/k\rfloor}\geq\left(1-\frac{1}{k!}\right)^{n/k}$. Since $\left(1-\frac{1}{k!}\right)^{1/k}\to 1$ as $k\to\infty$, the lemma follows. □

In order to make use of our construction, we need to find many 1-equivalences using Lemma 4.21. The easiest way to do so is if the numbers $3,1,2$ appear in that order and are not next to each other. In that case, we can then switch 1 and 2. A negligible number of permutations have two of $1,2,3$ next to each other, even when we restrict to $1\cdots k$ (consecutively) avoiding permutations. A third of the remaining permutations have 3 before 1 and 2, giving us enough permutations to work with. The next lemma formalizes this.

Lemma 4.29. *For all $c<1$, there exists a positive integer k such that for sufficiently large n, at least $c^n n!$ permutations of $[n]$ avoid the consecutive pattern $1,\dots,k$ and satisfy the additional property that $1,2,3$ are not consecutive in the permutation and 3 appears before 1 and 2.*

Proof. Let $a_{m,k}$ denote the number of permutations of $[n]$ that avoid the consecutive pattern $1\cdots k$. By a theorem of Ehrenborg, Kitaev, and Perry in [9], there exist positive constants $b_k, c_k > d_k$ with $\frac{a_{n,k}}{n!}=b_k c_k^n+O(d_k^n)$. By Lemma 4.28, we have that $c_k\to 1$ as $k\to\infty$.

We first show that there are at most $6a_{n-1,k}$ permutations on $[n]$ that avoid the consecutive pattern $1\cdots k$ but have two of $1,2,3$ next to each other. Indeed, note that if 1 and 2 are consecutive, then by deleting 2 and subtracting 1 from all other numbers, we obtain a permutation on $[n-1]$ that avoids the consecutive pattern $1\cdots k$. For each permutation π on $[n-1]$ that avoids the consecutive pattern $1\cdots k$, there are at most two such permutations on $[n]$ in which we could have deleted 2 to obtain π. To recover the original permutations, we just insert 2 in one of the two positions next to 1 and increment all other numbers by 1. If 2 and 3 are consecutive, we do the same thing but delete 3 and subtract 1 from all other numbers greater than 2, resulting in at most $2a_{n-1,k}$ other permutations. If 1 and 3 are consecutive but not adjacent to 2, then we do the same thing but delete 3 and subtract 1 from all other numbers greater than 3, resulting in at most $2a_{n-1,k}$ other permutations. Here, we are using the fact that when 2 is not adjacent to 1 and 3, we do not create any instances of $1\cdots k$ when we delete 3. This covers all cases, so there are at most $6a_{n-1,k}$ such permutations in total.

To finish, note that in a permutation that avoids the consecutive pattern $1\cdots k$ such that no two of $1,2,3$ are consecutive, we are free to permute $1,2,3$ without affecting this property. Thus, in exactly a third of these permutations does 3 appear before 1 and 2. It follows that the number of permutations with the desired property is at least $\frac{a_{n,k}-6a_{n-1,k}}{3} = b_k\left(\frac{1}{3}-\frac{2}{nc_k}\right)c_k^n n! + O(d_k^n n!)$. Since $c_k > d_k$ and $c_k \to 1$ as $k \to \infty$, the lemma follows. □

Theorem 1.6 then follows from combining Lemmas 4.20, 4.21, and 4.29 in the appropriate way.

Proof of Theorem 1.6. To complete the proof, we show that our construction for equivalences yields $c^n n!$ permutations of $[n]$ with nontrivial equivalences for all $c < 1$ and sufficiently large n when combined with Lemma 4.29.

Let $\epsilon > 0$ be a real number such that $c+\epsilon < 1$, and let k be a positive integer such that for all sufficiently large n, at least $(c+\epsilon)^n n!$ permutations of $[n]$ avoid the consecutive pattern $1\cdots k$ and satisfy the additional property that $1,2,3$ are not consecutive in the permutation and 3 appears before 1 and 2. Consider a permutation $\sigma = \sigma(1)\cdots\sigma(n-k-1)$ of $[n-k-1]$ with the aforementioned properties. We then consider the permutation $\pi = 1,\ldots,k,n,\sigma(1)+k,\ldots,\sigma(n-k-1)+k$ of $[n]$. Note that π is grounded (its streak is equal to its max height) with streak k. This is as $n > \sigma(1)+k$ and $\sigma(1)+k,\ldots,\sigma(n-k-1)+k$ avoids the consecutive pattern $1\cdots k$, so the height of all terms in the permutation are at most k. Let $\sigma(p)=1, \sigma(q)=2, \sigma(r)=3$, so $r < p,q$ and no two of p,q,r are consecutive. We then have that $\pi(k+1+p) = k+1$ and $\pi(k+1+q) = k+2$ are not consecutive in π and have the same height of 1. Since $\pi(k) = k$ and $\pi(k+1+r) = k+3$ appear before $k+1$ and $k+2$ in π, by our construction π is c-forest-Wilf equivalent to π', where $\pi' = \pi(1),\ldots,\pi(k+p),\pi(k+q+1),\pi(k+p+2),\ldots,\pi(k+q),\pi(k+p+1),\pi(k+q+2),\ldots,\pi(n)$, i.e. π but we switch $k+1$ and $k+2$. This is essentially an application of Lemmas 4.20 and 4.21, but we have written it in this way to count the number of pairs obtained. Since π and π' are not complements, π is a part of a nontrivial c-forest-Wilf equivalence. Each such σ results in a different π, and there are at least $(c+\epsilon)^{n-k-1}(n-k-1)! \geq \frac{(c+\epsilon)^n n!}{((c+\epsilon)n)^{k+1}}$ such σ. This is at least $c^n n!$ for sufficiently large n, as desired. □

Remark 4.30. We have made no efforts to improve the bound of $(1-o(1))^n n!$ that Theorem 1.6 gives for the number of patterns of length n that satisfy a nontrivial equivalence. It is possible that our construction yields many more nontrivial equivalences. However, finding an exact amount seems to be quite difficult due to the recursive examples of primitive structure equivalence given by Lemma 4.23. Nevertheless, our analysis above only uses Lemma 4.21 to swap 1 and 2 to find 1-equivalent patterns. There may be many more different patterns in which we can swap elements to obtain 1-equivalences. Furthermore, effective asymptotics for the number of permutations of $[n]$ that consecutively avoid $1\cdots k$ may lead to better bounds than given here. The result of Ehrenborg, Kitaev, and Perry we used gives the asymptotic $b_\pi c_\pi^n + O(d_\pi^n)$ for the number of permutations of $[n]$

that consecutively avoid any pattern π, where $b_\pi, c_\pi > d_\pi$ are positive constants that only depend on π. In the specific case of $\pi = 1 \cdots k$, more can be said about these constants. Elizalde and Noy in [10] find an explicit homogeneous linear differential equation with constant coefficients for the exponential generating function of the number of such permutations for each k. The characteristic polynomial of this equation is $\frac{x^k-1}{x-1}$, which suggests that the constants b_k and c_k may be explicitly found. The size of b_k and the rate of convergence of c_k to 1 can give more specific bounds on how many equivalences our construction produces.

4.2. **Necessary conditions for strong equivalences between grounded permutations.** Subsection 4.1 establishes some sufficient conditions for grounded permutations to be equivalent, and in this subsection we examine the other direction. Our goal is to determine some necessary conditions for two grounded permutations π and π' to be equivalent. We follow the same proof methodology of Theorem 4.1 given by Garg and Peng, finding formulas for specific cluster numbers and comparing them between π and π'. Note that if $\pi \overset{c}{\sim} \pi'$, then π and π' have the same length.

Proposition 4.31. *If π and π' are grounded permutations of length k with $\pi \overset{sc}{\sim} \pi'$, then π and π' have the same streak. Furthermore, the heights of i in π and π' are equal for all $i \in \{2, \dots, k\}$.*

Proof. To proceed, define the forest cluster numbers $r_{n,m}$ for π as the number of m-clusters of size n for π. Suppose that π has streak s. We will consider $r_{2k-h,2}$ for $h \leq s+1$. Note that 2-clusters of size $2k-h$ consist of two instances of π that intersect at h vertices. One way in which this occurs is that the two instances share the first h vertices. When $h = s+1$, there are no other ways in which this occurs, since the only increasing consecutive subsequence of π of length $s+1$ is at the beginning. For $h \leq s$, let $S_h \subseteq \{2, \dots, k\}$ denote the terms of π of height at most h (here we refer to the terms themselves, not the indices). Figure 14 shows examples of such clusters.

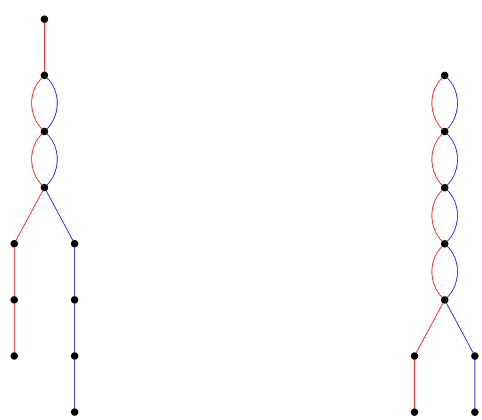

FIGURE 14. Clusters for a pattern of length 7 with two instances.

If the two instances of π do not share their first vertex, then one begins at the root of the cluster and the other begins at some non-root vertex of the first instance. Let v be the last vertex that the two instances share, and let i be the label of v. Note that all labels of the second instance that are descendants of v are at least i. It follows that the labels $1, \dots, i$ all lie in the first instance, so if v is the jth vertex in the first instance, then $\pi(j) = i$, and the height of i in π is at most h. The number of such clusters is then $\binom{2k-h-i}{k-h}$, as it suffices to determine the remaining $k-h$ labels in the second instance, where the possible values are in $\{i+1, \dots, 2k-h-i\}$. There are thus $\sum_{i \in S_h} \binom{2k-h-i}{k-h}$ such clusters.

If the two instances of π share their first vertex, then they share their first h vertices. For $h \le s$, the first h labels must be $1, \dots, h$. Then there are $\frac{1}{2}\binom{2k-2h}{k-h}$ ways to determine the rest of the labels, as we must decide which $k-h$ labels are in the first instance and then divide by two since the order of the instances does not matter when they share their first vertex. When $h = s+1$, there are $\frac{1}{2}\binom{2(\pi(s+1)-s-1)}{\pi(s+1)-s-1}\binom{2(k-\pi(s+1))}{k-\pi(s+1)}$ such clusters. Indeed, the labels of the first s vertices are $1, \dots, s$, and the label of the $(s+1)$st vertex v must be $s + 2(\pi(s+1) - s - 1) + 1$, since among the remaining vertices exactly $2(\pi(s+1) - s - 1)$ have labels less than v. It the remains to decide which $\pi(s+1) - s - 1$ of the elements of $\{s+1, \dots, s+2(\pi(s+1)-s-1)\}$ and which $k-\pi(s+1)$ of the elements of $\{s+2(\pi(s+1)-s-1)+2, \dots, 2k-s-1\}$ are labels in the first instance, and then divide by 2 due to the indistinguishability of the two instances.

Thus, $r_{2k-h,2} = \frac{1}{2}\binom{2k-2h}{k-h} + \sum_{i \in S_h}\binom{2k-h-i}{k-h}$ for $h \le s$, and for $h = s+1$ we have that

$$r_{2k-s-1,2} = \frac{1}{2}\binom{2(\pi(i+1)-s-1)}{\pi(s+1)-s-1}\binom{2(k-\pi(s+1))}{k-\pi(s+1)}.$$

Define the cluster numbers $r'_{n,m}$ and sets S'_h for π' analogously, and note that by Theorem 3.2, if π and π' are strong-c-forest-Wilf equivalent then $r_{n,m} = r'_{n,m}$ for all m, n. Suppose that the streak s' of π' is greater than s. We have that

$$\frac{1}{2}\binom{2(\pi(i+1)-s-1)}{\pi(s+1)-s-1}\binom{2(k-\pi(s+1))}{k-\pi(s+1)} = \frac{1}{2}\binom{2k-2s-2}{k-s-1} + \sum_{i \in S'_{s+1}}\binom{2k-s-1-i}{k-s-1}$$

since $r_{2k-s-1,2} = r'_{2k-s-1,2}$. But note that the right-hand side is strictly greater than the left-hand side. Indeed, the sequence $\binom{2a}{a}\binom{2n-2a}{n-a}$ is strictly decreasing for $a \le \frac{n}{2}$ as

$$\frac{\binom{2a+2}{a+1}\binom{2n-2a-2}{n-a-1}}{\binom{2a}{a}\binom{2n-2a}{n-a}} = \frac{(2a+1)(n-a)}{(a+1)(2n-2a-1)} < 1$$

for $n > 2a+1$. Then as S'_{i+1} is nonempty, the sum on the right-hand side is positive while the binomial on the right-hand side is at least the binomial on the left-hand side. Hence, it cannot happen that $r_{2k-s-1,2} = r'_{2k-s-1,2}$ if $s \ne s'$, so $s = s'$, as desired.

We then consider the equalities $r_{2k-h,2} = r'_{2k-h,2}$ for all $h \le s$. They tell us that

$$\frac{1}{2}\binom{2k-2h}{k-h} + \sum_{i \in S_h}\binom{2k-h-i}{k-h} = \frac{1}{2}\binom{2k-2h}{k-h} + \sum_{i \in S'_h}\binom{2k-h-i}{k-h},$$

so $\sum_{i \in S_h}\binom{2k-h-i}{k-h} = \sum_{i \in S'_h}\binom{2k-h-i}{k-h}$. We claim that this means that $S_h = S'_h$. Indeed, the elements of S_h and S'_h are greater than h, and for $i = h+1, \dots, k$ the value of $\binom{2k-h-i}{k-h}$ is among $\binom{2k-2h-1}{k-h}, \dots, \binom{k-h}{k-h}$. But each element in this sequence is more than twice the next element, as $\frac{\binom{k-h+i}{k-h}}{\binom{k-h+i-1}{k-h}} = \frac{k-h+i}{i} > 2$ for $i < k-h$. It follows that $\sum_{i \in S}\binom{2k-h-i}{k-h}$ takes on different values for different $S \subseteq \{h+1, \dots, k\}$, which means that $\sum_{i \in S_h}\binom{2k-h-i}{k-h} = \sum_{i \in S'_h}\binom{2k-h-i}{k-h}$ can only occur when $S_h = S'_h$. The sets of terms of height at most h are equal for π and π' for all $h \le s$. It follows that the heights of all $i \in \{2, \dots, k\}$ with respect to π and π' are equal, as desired. $\square$

Proposition 4.32. *If π and π' are grounded permutations of length k that are strong-c-forest-Wilf equivalent with streak s, then $\pi(s+1) = \pi'(s+1)$ or $\pi(s+1) + \pi'(s+1) = s+k+1$. Furthermore, if d is the least positive integer such that $\{\pi(d), \dots, \pi(k)\}$ does not contain two consecutive numbers and if d' is the least positive integer such that $\{\pi'(d'), \dots, \pi'(k)\}$ does not contain two consecutive numbers, then $d = d'$.*

Proof. We will prove this proposition by considering the equation $r_{2k-h,2} = r'_{2k-h,2}$ for $h > s$. Note that in a 2-cluster of size $2k-h$ for $h > s$, the two instances must share the first h vertices. By the same arguments as in Proposition 4.31, if $\{s+1,\dots,k\}\setminus\{\pi(s+1),\dots,\pi(h)\}$ has consecutive blocks of size $a_1,\dots,a_i$, then $r_{2k-h,2} = \frac{1}{2}\binom{2a_1}{a_1}\cdots\binom{2a_i}{a_i}$. By taking $h = s+1$, we have that $a_1 = \pi(i+1)-s-1$ and $a_2 = k-\pi(i+1)$. If we define $a'_1,\dots,a'_i$ analogously for π', we have that $a'_1 = \pi'(i+1)-s-1$ and $a'_2 = k-\pi'(i+1)$. As before, the sequence $\binom{2a}{a}\binom{2n-2a}{n-a}$ is strictly decreasing for $a \le \frac{n}{2}$, so $\frac{1}{2}\binom{2a_1}{a_1}\binom{2a_2}{a_2} = \frac{1}{2}\binom{2a'_1}{a'_1}\binom{2a'_2}{a'_2}$ and thus $\{a_1,a_2\} = \{a'_1,a'_2\}$ since $a_1+a_2 = a'_1+a'_2$. This implies that $\pi(s+1) = \pi'(s+1)$ or $\pi(s+1)+\pi'(s+1) = s+k+1$. Note that $r_{2k-h,2} = 2^{k-h-1}$ for all $h \ge d$, while $r_{2k-d+1,2} = 3\cdot 2^{k-d-1}$. The same holds true for r' and d', which implies the second part of the proposition. □

Note that the necessary conditions we have derived are quite far from the sufficient conditions we showed in Subsection 4.1. One may be able to show using these methods that the set of indices of height h in π is also equal to the set of indices of height h in π, which is consistent with the equivalences we showed in Subsection 4.1. However, the only information we get from these cluster numbers is a system of equations of the form $\prod\binom{2a_j}{a_j} = \prod\binom{2a'_j}{a'_j}$, which does not appear to be easy to solve.

4.3. **The triviality of super-strong equivalences.**

In this subsection, we prove Theorem 1.7. We follow the same methodology of counting linear extensions of cluster posets as in [6, 25]. We again consider two permutations $\pi = \pi(1)\cdots\pi(k)$ and $\pi' = \pi'(1)\cdots\pi'(k)$ of length k, assumed to satisfy $\pi \overset{ssc}{\sim} \pi'$. Let $m = \lfloor\frac{k}{2}\rfloor$.

For an unlabeled rooted forest F on n vertices with a set S of highlighted paths of length k, let $a_{F,S}$ be the number of ways to label the vertices of F with distinct labels in $[n]$ such that the instances of π are precisely the highlighted paths in S. If $a'_{F,S}$ is defined analogously for π', note that by definition, $\pi \overset{ssc}{\sim} \pi'$ if and only if $a_{F,S} = a'_{F,S}$ for all F and S.

Definition 4.33. An *unlabeled cluster* of order k is a rooted tree with a set of highlighted paths of length k such that every vertex belongs to a highlighted path and it is not possible to partition the vertex set into two parts such that no path contains vertices in both parts. In other words, an unlabeled cluster is simply a forest cluster without the labels. For each unlabeled cluster C of order k, define r_C and r'_C to be the number of ways to label the vertices in C that result in a forest cluster for π and π', respectively.

Definition 4.34. Given an unlabeled cluster C of order k, the *forest cluster poset* P_C is the poset on the vertex set of C generated by the following relations: for each highlighted path $v_1,\dots,v_k$, $v_{\pi^{-1}(1)} < \cdots < v_{\pi^{-1}(k)}$, where π^{-1} is the inverse of π in S_k. We define P'_C analogously for π'.

Note that r_C and r'_C are respectively the number of linear extensions of P_C and P'_C divided by the number of automorphisms of the underlying rooted tree of C.

We first prove the following relation between the refined forest cluster numbers r_C and super-strong equivalence for forests. Our proof mimics the analogous result of Dwyer and Elizalde for permutations given in [6].

Lemma 4.35. *We have that $\pi \overset{ssc}{\sim} \pi'$ if and only if $r_C = r'_C$ for all unlabeled clusters C of order k.*

Proof. Let $b_{F,S}$ be the number of ways to label F such that all paths in S are instances of π, though not necessarily all instances of π are in S, and define $b'_{F,S}$ analogously for π'. We have that $b_{F,S} = \sum_{S\subseteq T} a_{F,T}$ and $b'_{F,S} = \sum_{S\subseteq T} a'_{F,T}$. By the Principle of Inclusion-Exclusion, $a_{F,S} = a'_{F,S}$ for all F and S if and only if $b_{F,S} = b'_{F,S}$ for all F and S, so $\pi \overset{ssc}{\sim} \pi'$ if and only if $b_{F,S} = b'_{F,S}$ for all F and S.

It is clear that the equality $b_{F,S} = b'_{F,S}$ holding for all F and S implies $r_C = r'_C$ for all C, as C is an unlabeled cluster where the underlying forest is a tree. In the other direction, consider the finest partition of the vertex set of F such that any two adjacent vertices in a highlighted path are in the same part. In this way, we have partitioned the vertices into singleton sets and trees $T_1, \dots, T_p$ with sets $S_1, \dots, S_p$ of highlighted paths such that T_i and S_i form an unlabeled cluster C_i. If the number of vertices of T_i is n_i and the number of vertices of F is n, then in a uniform random labelling of F, the probability that all highlighted paths are instances of π is $\frac{b_{F,S}}{n!}$. On the other hand, it is also equal to $\prod_{i=1}^{p} \frac{r_{C_i}}{n_i!}$ since the events E_i that the paths in S_i are all instances of π are independent. We thus have that $\frac{b_{F,S}}{n!} = \prod_{i=1}^{p} \frac{r_{C_i}}{n_i!}$ and analogously that $\frac{b'_{F,S}}{n!} = \prod_{i=1}^{p} \frac{r'_{C_i}}{n_i!}$, which proves the reverse direction, establishing the lemma. □

Recall that by Theorem 4.1 and complementation, we may assume that $\pi(1) = \pi'(1)$ since $\pi \overset{ssc}{\sim} \pi'$. We first construct a certain family of clusters in Lemma 4.36 that prove that under this assumption, in most cases the two permutations must be the same, which would imply the theorem. The one caveat is that if $\pi(1) = \frac{k-1}{2}$, we are unable to distinguish between complements with only Lemma 4.36, so Lemma 4.37 is also required to prove the result. In contrast to [11, 25], our approach is not to show a contradiction of asymptotics but rather to use the polynomial identity theorem: two polynomials that agree on an infinite set must be equal. The extra flexibility of forests allows us to construct families of cluster posets whose linear extension counts are effectively given by polynomials. The key innovation in our proof is to rely on the algebraic properties of polynomials, such as their roots or the ability to cancel out common factors, that may not have an obvious combinatorial interpretation, instead of on how fast the linear extension counts grow.

Lemma 4.36. *Suppose that $\pi(1) = \pi'(1)$ and $\pi \overset{ssc}{\sim} \pi'$. If it is not the case that $\pi(1) = m$ and $k = 2m-1$, then $\pi = \pi'$. If $\pi(1) = m$ with $k = 2m-1$, then $\pi(i) = \pi'(i)$ or $\pi(i) + \pi'(i) = k+1$ for all i.*

Proof. Fix an $i > 1$ and let $q = \pi(1)$. Consider the unlabeled cluster C_n of order k defined as follows: we have one instance beginning at the root of C_n, and there are n disjoint instances starting from the ith vertex of this instance. Thus, the intersection of any pair of instances consists only of the ith vertex of the first instance. Suppose that $\pi(i) = p$ and $\pi'(i) = p'$. We now describe the cluster poset P_{C_n}. It consists of $n+1$ chains with one common element v. In one of these chains, v is the pth smallest element, and in the rest of the chains, v is the qth smallest element. This poset has $\binom{(q-1)n+p-1}{q-1,\dots,q-1,p-1}\binom{(k-q)n+k-p}{k-q,\dots,k-q,k-p}$ linear extensions, as the label of v is determined and the poset less than v and the poset greater than v all consist of n disjoint chains. Similarly, P'_{C_n} has $\binom{(q-1)n+p'-1}{q-1,\dots,q-1,p'-1}\binom{(k-q)n+k-p'}{k-q,\dots,k-q,k-p'}$ linear extensions. Figure 15 gives an example of an unlabeled cluster C_n and its cluster poset P_{C_n}.

If π and π' are super-strongly equivalent, then P_{C_n} and P'_{C_n} have the same number of linear extensions so $\binom{(q-1)n+p-1}{q-1,\dots,q-1,p-1}\binom{(k-q)n+k-p}{k-q,\dots,k-q,k-p} = \binom{(q-1)n+p'-1}{q-1,\dots,q-1,p'-1}\binom{(k-q)n+k-p'}{k-q,\dots,k-q,k-p'}$ for all n. This is equivalent to

$$\binom{(q-1)n+p-1}{p-1}\binom{(k-q)n+k-p}{k-p} = \binom{(q-1)n+p'-1}{p'-1}\binom{(k-q)n+k-p'}{k-p'}$$

holding for all n. This holds as a polynomial identity for all n, so it follows that

$$\binom{(q-1)x+p-1}{p-1}\binom{(k-q)x+k-p}{k-p} = \binom{(q-1)x+p'-1}{p'-1}\binom{(k-q)x+k-p'}{k-p'}$$

as polynomials. The roots of the left-hand side and right-hand side then must be the same. It is easy to see that the roots of the left-hand side are $-\frac{1}{q-1}, \dots, -\frac{p-1}{q-1}, -\frac{1}{k-q}, \dots, -\frac{k-p}{k-q}$ and the roots of the right-hand side are $-\frac{1}{q-1}, \dots, -\frac{p'-1}{q-1}, -\frac{1}{k-q}, \dots, -\frac{k-p'}{k-q}$. If $q-1 \neq k-q$, then we must have

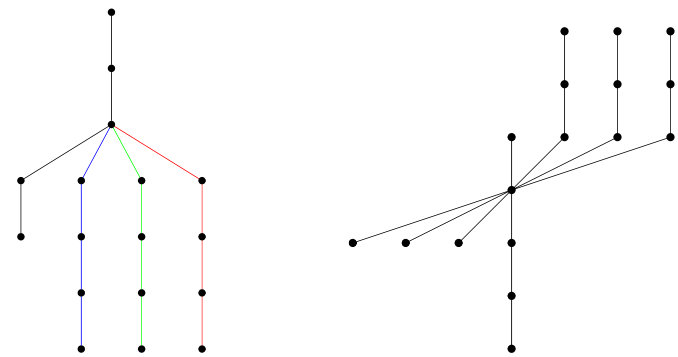

FIGURE 15. The unlabeled cluster C_n for a pattern π of length 5 and the Hasse diagram for the cluster poset P_{C_n} when $n = 3$, $i = 3$, $\pi(1) = 2$, and $\pi(3) = 4$.

that $p = p'$, corresponding to when it is not the case that $\pi(1) = m$ and $k = 2m - 1$. Otherwise, we must have that $p = p'$ or $p + p' = k + 1$, so the lemma follows. □

Note that when $\pi(1) = \frac{k+1}{2}$, the family of cluster posets we construct in the proof of Lemma 4.36 cannot distinguish between the two cases $\pi(i) = \pi'(i)$ or $\pi(i) + \pi'(i) = k+1$ because if we were to construct the same family for $\overline{\pi}$, then we would get an isomorphism between the corresponding cluster posets after reversing the relative order. Any proof of Theorem 1.7 must take into account the case that $\pi' = \overline{\pi}$. When $\pi(1) \neq \frac{k+1}{2}$, this is taken care of by the complementation we use at the beginning to assume $\pi(1) = \pi'(1)$, but this does not quite work when $\pi(1) = \frac{k+1}{2}$. Lemma 4.36 shows that each individual term $\pi'(i)$ of π' is as we expect (equal to $\pi(i)$ or $k + 1 - \pi(i)$) but cannot show that the choice between $\pi(i)$ and $k+1-\pi(i)$ is uniform over all i. To fix this, we add another "branch" to the family of clusters we constructed to introduce asymmetry between π and $\overline{\pi}$. By Lemma 4.36, we can take complements to fix at least one other term of the pattern between π and π'. We make a specific choice for the term fixed and modify our family of clusters to show that when $\pi \neq \pi'$ the cluster posets do not have the same number of linear extensions for every cluster. The proof of Lemma 4.37 uses a few somewhat tricky algebraic manipulations. It would be interesting to see a more combinatorial proof for Lemma 4.36 and especially Lemma 4.37.

Lemma 4.37. *If $\pi(1) = \pi'(1) = m$ and $\pi \overset{ssc}{\sim} \pi'$ with $k = 2m - 1$ and $\pi(\ell) = \pi'(\ell)$ for some $\ell > 1$, then $\pi = \pi'$.*

Proof. Let i be such that $\pi(i) = 1$. By Lemma 4.36, we know that $\pi'(i) = 1$ or $\pi'(i) = k$. By complementation, we may without loss of generality suppose that $\pi'(i) = 1$. Suppose that $\pi(j) = a$. We will show that it cannot be the case that $\pi'(j) = k + 1 - a$, so $\pi'(j) = a$. Then it will follow that π and π' must be either equal or complements, so if π and π' agree at two indices then they are equal. We already know this to be the case for $\pi(j) = 1, m, k$ by assumption, so we only need to consider $a \neq 1, m, k$.

Consider the unlabeled cluster C_n of order k defined as follows: we have one instance I beginning at the root of C, n instances $J_1, \ldots, J_n$ beginning at the ith vertex of I and sharing no other vertices with I or each other, and one instance K beginning at the jth vertex of I and sharing no other vertices with $I, J_1, \ldots, J_n$. We now describe the cluster posets P_{C_n} and P'_{C_n}, assuming that $\pi'(j) = k + 1 - a$. Both cluster posets consist of $n + 2$ chains of length k, corresponding to the $n + 2$ instances in C. Let the *central chain* be the chain corresponding to I with lowest element u, the *lower chains* be the chains corresponding to $J_1, \ldots, J_n$, and the *upper chain* be the chain corresponding to K, which intersects the central chain at the element v. Note that u is the common median and only pairwise common element of the lower chains and that v is the median of the upper chain as $\pi(1) = m$. In P_{C_n}, v is the ath smallest element of the central chain and in P'_{C_n}, v is the ath largest element of the central chain. Figure 16 gives an example of an unlabeled cluster C_n and its cluster poset P_{C_n}.

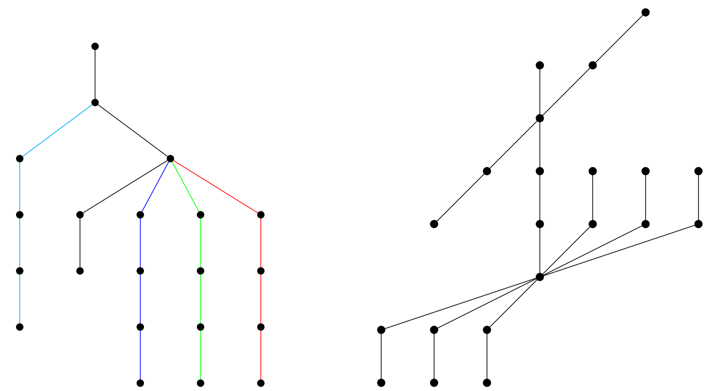

FIGURE 16. The unlabeled cluster C_n for a pattern π of length 5 and the Hasse diagram for the cluster poset P_{C_n} when $n = 3$, $i = 3$, $j = 2$, and $a = 4$.

Now, we will enumerate linear extensions of P_{C_n} and P'_{C_n}. We proceed using casework on the number s of elements of the upper chain with label less than the label of v. Here, we must have $s < m$ as there are only $m-1$ elements of the upper chain that are incomparable to v and all other elements are greater than v. Note that the label of u must be $(m-1)n + s + 1$. There are then $\binom{(m-1)n+s}{s}$ ways to choose which of the numbers in $[(m-1)n + s]$ are labels of elements in the upper chain. We next choose the labels of the remaining elements not in one of the lower chains. Note that there are $4(m-1) - s$ such elements, and the labels can be anything in $\{(m-1)n + s + 2, \ldots, 2(m-1)n + 4(m-1) + 1\}$, resulting in $\binom{(m-1)n+4(m-1)-s}{4(m-1)-s}$ ways. We then assign the specific labels to the elements not in the lower chains. The s smallest elements of the upper chain already have their labels determined. The remaining elements form a poset consisting of two chains that intersect at v. One of these chains has $m-1$ elements greater than v and $m-s-1$ elements less than v. For P_{C_n} the other chain has $2m-a-1$ elements greater than v and $a-2$ elements less than v, and for P'_{C_n} the other chain has $a-1$ elements greater than v and $2m-a-2$ elements less than v. Thus, there are $\binom{3m-a-2}{m-1}\binom{m-s+a-3}{m-s-1}$ ways to do so for P_{C_n} and $\binom{m+a-2}{m-1}\binom{3m-s-a-3}{m-s-1}$ ways to do so for P'_{C_n}. Finally, to decide the labels of the elements in the lower chains, note that there are $N = \binom{(m-1)n}{m-1,\ldots,m-1}$ ways to choose the labels of the elements smaller than u as those just consist of n disjoint chains of $m-1$ elements. Similarly, there are N ways to choose the labels of the elements larger than u. Thus, the number of linear extensions of P_{C_n} and P'_{C_n} are

$$N^2 \sum_{s=0}^{m-1} \binom{(m-1)n+s}{s}\binom{(m-1)n+4(m-1)-s}{4(m-1)-s}\binom{3m-a-2}{m-1}\binom{m-s+a-3}{m-s-1}$$

and

$$N^2 \sum_{s=0}^{m-1} \binom{(m-1)n+s}{s}\binom{(m-1)n+4(m-1)-s}{4(m-1)-s}\binom{m+a-2}{m-1}\binom{3m-s-a-3}{m-s-1},$$

respectively.

To have π super-strongly equivalent to π, these two quantities must be equal for all n. Dividing both sides by N^2, we have by Lemma 4.35 that

$$\sum_{s=0}^{m-1} \binom{(m-1)n+s}{s} \binom{(m-1)n+4(m-1)-s}{4(m-1)-s} \binom{3m-a-2}{m-1} \binom{m-s+a-3}{m-s-1}$$
$$= \sum_{s=0}^{m-1} \binom{(m-1)n+s}{s} \binom{(m-1)n+4(m-1)-s}{4(m-1)-s} \binom{m+a-2}{m-1} \binom{3m-s-a-3}{m-s-1}$$

for all n. Note that both sides are in fact polynomials in n. Thus, it follows by substituting $x = (m-1)n$ that

$$\sum_{s=0}^{m-1} \binom{x+s}{s} \binom{x+4(m-1)-s}{4(m-1)-s} \binom{3m-a-2}{m-1} \binom{m-s+a-3}{m-s-1}$$
$$= \sum_{s=0}^{m-1} \binom{x+s}{s} \binom{x+4(m-1)-s}{4(m-1)-s} \binom{m+a-2}{m-1} \binom{3m-s-a-3}{m-s-1}$$

is a polynomial identity in x. Dividing both sides by $x+1$ gives

$$\sum_{s=0}^{m-1} \binom{x+s}{s} \frac{\binom{x+4(m-1)-s}{4(m-1)-s}}{x+1} \binom{3m-a-2}{m-1} \binom{m-s+a-3}{m-s-1}$$
$$= \sum_{s=0}^{m-1} \binom{x+s}{s} \frac{\binom{x+4(m-1)-s}{4(m-1)-s}}{x+1} \binom{m+a-2}{m-1} \binom{3m-s-a-3}{m-s-1}$$

where $\frac{1}{x+1}\binom{x+4(m-1)-s}{4(m-1)-s} = \frac{1}{4(m-1)-s}\binom{x+4(m-1)-s}{4(m-1)-s-1}$ is a polynomial that does not vanish at $x = -1$. Plugging in $x = -1$, we note that $\binom{s-1}{s} = 0$ for all $s > 0$ while $\binom{-1}{0} = 1$. Hence, all terms from $s > 0$ will vanish, leaving the identity

$$\frac{1}{4(m-1)} \binom{3m-a-2}{m-1} \binom{m+a-3}{m-1} = \frac{1}{4(m-1)} \binom{m+a-2}{m-1} \binom{3m-a-3}{m-1}.$$

This is equivalent to $\frac{3m-a-2}{2m-a-1} = \frac{m+a-2}{a-1}$, which cannot happen unless $a = m$, a contradiction. Thus, we cannot have that $\pi'(j) = k+1-a$, and the lemma is proven. □

The combination of Lemmas 4.36 and 4.37 proves Theorem 1.7. By examining our proof closely, we also have the following corollary.

Corollary 4.38. *If patterns π and π' satisfy $\pi \overset{sc}{\sim} \pi'$ nontrivially, then by Theorem 4.3, for all m and n the number of m-clusters of size n with respect to π is equal to the number of m-clusters of size n with respect to π'. No bijection between clusters with respect to π and clusters with respect to π' can be structure-preserving. By this, we mean that it cannot be the case that such a bijection always maps a cluster C with respect to π to a cluster C' with respect to π' such that the underlying unlabeled rooted trees of C and C' are isomorphic.*

Proof. Note that in our proofs of Lemmas 4.36 and 4.37, the families of unlabeled clusters we construct have the following property: given the underlying rooted tree structure T of the unlabeled cluster and the number m of highlighted instances, there is only one way to highlight m paths of length k in T to yield an unlabeled cluster. This is as T has exactly m leaves, each of which must correspond to the endpoint of a highlighted instance. If $\pi \overset{sc}{\sim} \pi'$, then we have that $\pi(1) = \pi'(1)$ or $\pi(1) + \pi'(1) = k+1$ by Theorem 4.1. In that case, the proofs of Lemmas 4.36 and 4.37 tell us that the number of linear extensions of the corresponding cluster posets, which is precisely the number of clusters for π or π' with underlying rooted tree T up to a factor of automorphisms, cannot be equal unless $\pi = \pi'$ or $\pi' = \overline{\pi}$. $\square$

When $\pi' = \overline{\pi}$, complementation gives a structure-preserving bijection between forest clusters. In Remark 4.26, we briefly described a bijection between clusters of 125364 and 125463. This bijection is not structure-preserving, as it involves cutting and replanting branches of the tree in the cluster. A similar bijection can be given for all of the nontrivial strong c-forest-Wilf equivalences that we constructed in Subsection 4.1. Corollary 4.38 essentially tells us that doing some sort of "cutting and replanting" of the underlying tree is unavoidable if we are to prove a nontrivial strong equivalence via a bijection between forest clusters. This highlights a key difference between consecutive pattern avoidance in forests and permutations: the forest structure plays a fundamental role in determining strong c-forest-Wilf equivalence.

5. Forest Stanley-Wilf Limits

In this section, we discuss the asymptotics of $f_n(S)$ and $t_n(S)$, the number of forests and trees, respectively, on $[n]$ that avoid a set S of patterns. Our main focus will be on classical non-consecutive avoidance, though we will make a few remarks on how to modify our techniques and results to deal with consecutive avoidance and other types of pattern avoidance as well.

In Subsection 5.1, we prove Theorem 1.9 using analytic techniques and describe how our result can be applied to pattern avoidance in labeled forests in a very general sense. Subsection 5.2 discusses the problem of determining the forest Stanley-Wilf limit of a given set S of patterns, mostly those sets of patterns covered by Theorem 1.9.

We will be working closely with the exponential generating functions $F_S(x) = \sum_{n\geq 0} \frac{f_n(S)}{n!}x^n$ and $T_S(x) = \sum_{n\geq 0} \frac{t_n(S)}{n!}x^n$ of $f_n(S)$ and $t_n(S)$. We generally only consider one set S at a time, so we will often suppress S from the notation and write $f_n, t_n, F(x), T(x)$ instead. We make use of many basic properties of exponential generating functions of labeled combinatorial classes, such as the fact that $F(x) = e^{T(x)}$ since a forest that avoids S is a set of trees avoiding S. Formally, if $\mathcal{F}$ is the class of rooted labeled forests avoiding S and $\mathcal{T}$ is the class of rooted labeled trees avoiding S, then $\mathcal{F} = \mathsf{SET}(\mathcal{T})$. We refer readers unfamiliar with the theory of labeled combinatorial classes and analytic combinatorics to [12] for a comprehensive treatise.

Before we move on to our proofs, we record the following definition, which is made primarily for the sake of brevity in the later arguments.

Definition 5.1. A set S of patterns is *covered* if it contains two patterns $\pi = \pi(1)\cdots\pi(k)$ and $\sigma = \sigma(1)\cdots\sigma(\ell)$ with $\pi(1) = 1$ and $\sigma(1) = \ell$. Otherwise, S is said to be *uncovered*.

In other words, S is covered if it contains a pattern that begins with its lowest element and a pattern that begins with its highest element, which in particular means that singleton sets S are uncovered (we will ignore the pattern 1 as only the empty forest avoids it). Here, the word covered refers to the fact that the two "ends" of the possible values of the first element, the highest and lowest number, both appear. Note that Theorem 1.9 states that the forest Stanley-Wilf limit exists for uncovered S. By complementation, when working with uncovered S we may assume that no patterns in S begin with 1.

5.1. The forest Stanley-Wilf conjecture for uncovered sets.

In this subsection, we prove Theorem 1.9, which states that for any uncovered S the forest Stanley-Wilf limit $\lim_{n\to\infty} \frac{f_n^{1/n}}{n}$ exists and is finite. We first make a few remarks before giving our proof.

In contrast with the Stanley-Wilf conjecture for permutations, the main difficulty lies not in the finiteness but in the existence of the limit. Indeed, the total number of rooted labeled forests on $[n]$ is $(n+1)^{n-1}$ by Cayley's formula, which in particular implies that $\limsup_{n\to\infty} \frac{f_n^{1/n}}{n} \leq 1$. Under the assumption that no patterns in S begin with 1, note that if $21 \in S$, then the other patterns in S are superfluous as they contain all 21. In this case, $f_n = n!$, and we already know by Stirling's approximation that $\lim_{n\to\infty} \frac{(n!)^{1/n}}{n} = e^{-1}$. Henceforth, we will assume that no patterns in S begin with 1 and that $21 \notin S$. Another consequence of this limit is that it instead suffices to show the existence of $\lim_{n\to\infty} \left(\frac{f_n}{n!}\right)^{1/n} = e \lim_{n\to\infty} \frac{f_n^{1/n}}{n}$ instead, which we do using the exponential generating function of f_n.

The supermultiplicativity argument given by Arratia in [2] for the existence of the limit for permutations does not easily extend to forests. The analogous supermultiplicativity inequality is $\dfrac{f_{m+n}}{(m+n)!} \geq \dfrac{f_m}{m!} \cdot \dfrac{f_n}{n!}$, or $f_{m+n} \geq \binom{m+n}{m} f_m f_n$. We remark that for *ordered* rooted forests, this inequality follows from the observation that one can obtain an ordered forest F on $[m+n]$ avoiding S by choosing an m-element subset A of $[m+n]$ and merging an ordered forest F_A on A avoiding S with an ordered forest F_B on $B = [m+n] \setminus A$ avoiding S. We do so by placing the trees in F_A before the trees in F_B in the ordering of the trees in F. This does not work for unordered forests because the construction is not injective; the order of the trees in the forest no longer matters so we cannot say that every choice of A, F_A, F_B results in a different forest on $[m+n]$. We were unable to repair this argument for unordered forests, where data suggests that the inequality holds.

We now make some definitions relevant to our proof. As before, let $T(x) = \sum_{k\geq 0} \frac{t_k x^k}{k!}$ and $F(x) = \sum_{k\geq 0} \frac{f_k x^k}{k!}$ be the exponential generating functions of t_n and f_n. Let

$$A(x) = T'(x) - F(x) = \frac{F'(x)}{F(x)} - F(x) = \sum_{k\geq 0} \frac{(t_{k+1} - f_k)x^k}{k!},$$

$$B(x) = \int_0^x A(t)dt,$$

$$C(x) = e^{B(x)},$$

$$D(x) = \int_0^x C(t)dt.$$

For positive integers n, let

$$A_n(x) = \sum_{0\leq k\leq n} \frac{(t_{k+1} - f_k)x^k}{k!},$$

$$B_n(x) = \int_0^x A_n(t)dt,$$

$$C_n(x) = e^{B_n(x)},$$

$$D_n(x) = \int_0^x C_n(t)dt,$$

$$F_n(x) = \frac{C_n(x)}{1 - D_n(x)}.$$

Given nonnegative integers $0 = a_0, a_1, \ldots, a_n$, where a_k represents the number of objects in a labeled combinatorial class $\mathcal{C}$ of size k, let $E(a_0, \ldots, a_n)$ denote the number of objects in $\mathsf{SET}(\mathcal{C})$ of size n. Note that $\frac{E(a_1,\ldots,a_n)}{n!}$ is the coefficient of x^n in $\exp\left(\sum_{k=1}^n \frac{a_k}{k!} x^n\right)$ and that $f_n = E(t_1, \ldots, t_n)$ by construction.

The rough idea in our proof is that $\limsup_{n\to\infty} \left(\frac{f_n}{n!}\right)^{1/n}$ is the reciprocal of the radius of convergence of F. To control $\liminf_{n\to\infty} \left(\frac{f_n}{n!}\right)^{1/n}$, we approximate F from below by a series of functions F_m with the property that $F_m(z)$, viewed as a function in a complex variable z of a sufficiently small magnitude, admits a meromorphic continuation to the entire complex plane. By [12, Theorem

IV.10], the corresponding limit for the coefficients of F_m exist, and this limit is a lower bound for $\liminf_{n\to\infty}\left(\frac{f_n}{n!}\right)^{1/n}$. The theorem then follows by showing that F_m tends to F in an appropriate sense. The first step is to make a combinatorial observation about the coefficients of F and T, using the condition that S is uncovered. The remainder of the proof after Proposition 5.2 is essentially purely analytic.

Proposition 5.2. *We have that $t_{k+1} \geq f_k$ for all k. Thus, all of the coefficients of A are nonnegative.*

Proof. Because none of the patterns in S start with 1, any forest on $[k]$ avoiding S can be turned into a tree on $[k+1]$ avoiding S by increasing all labels by 1 and attaching the root of each tree in the forest to a new root vertex labeled 1. This operation is clearly injective, so the proposition follows. □

Lemma 5.3. *There exist unique positive real numbers r_n such that $D_n(r_n) = 1$. Furthermore, the sequence $r_1, r_2, \ldots$ is nonincreasing with limit $r = \sup\{t : D(t) \leq 1\}$.*

Proof. Since all of the coefficients of A are nonnegative, so are all of the coefficients of B, C, D as they are constructed from A using integration and exponentiation. The same is true for A_n, B_n, C_n, D_n. By our assumption that S does not contain 21, we know that $t_2 = 2$. Since $f_1 = 1$, the coefficient of x in $A(x)$ is equal to 1. In particular, this means that A_n, B_n, C_n, D_n each have a strictly positive non-constant coefficient, so they are strictly increasing functions in x that tend to infinity since A_n is a polynomial. We note here that $A(x), B(x), C(x), D(x)$ are defined on $x \in [0, R)$ where R is their common radius of convergence. By Cayley's formula, $t_{k+1} - f_k \leq t_{k+1} \leq (k+1)^k$, so $R^{-1} \leq \lim_{k\to\infty}\left(\frac{(k+1)^k}{k!}\right)^{1/k} = e$ and R is positive.

Since A_n, B_n, C_n, D_n are finite on $[0,\infty)$, are strictly increasing, and tend to infinity, by the fact that $D_n(0) = 0$ we know that the r_n exist and are unique. Furthermore, since $D_1, D_2, \ldots$ is pointwise nondecreasing, $r_1, r_2, \ldots$ is nonincreasing. Note that A is the pointwise increasing limit of A_n on $[0,\infty)$, so B is the pointwise increasing limit of B_n on $[0,\infty)$ by the integral monotone convergence theorem. It then follows that C is the pointwise increasing limit of C_n on $[0,\infty)$, so D is the pointwise increasing limit of D_n on $[0,\infty)$ as well, again by the integral monotone convergence theorem. If $R = \infty$, then $D(x)$ is defined for all $x \geq 0$ and tends to infinity. Hence, $D(r) = 1$, r_n approaches r from above, and D_n approaches D from below pointwise.

Suppose that $R < \infty$. Note that $r \leq R$. We now split into cases of whether $r < R$. If $r = R$, then it suffices to show that r_n has limit R. Since $D_n \uparrow D$ and $R = \sup\{t : D(t) \leq 1\}$, we have that $D_n(R) < 1$ for all n. On the other hand, for all $\epsilon > 0$ and sufficiently large n we have that $D_n(R+\epsilon) > 1$ since $D(R+\epsilon) = \infty$. Thus, for sufficiently large n we have that $r_n < R+\epsilon$ while $r_n > R$, so $\lim_{n\to\infty} r_n = R = r$, as desired. If $r < R$, then we know that $D(r) = 1$ and for some $\epsilon > 0$ we have that $D(R-\epsilon) > 1$. Since $D_n \uparrow D$, for sufficiently large n we have that $D_n(R-\epsilon) > 1$ and $r_n < R-\epsilon$, and by restricting to the interval $[0, R-\epsilon]$ the result is clear. □

Proposition 5.4. *The differential equation $\frac{G'(x)}{G(x)} - G(x) = A_n(x)$ with initial condition $G(0) = 1$ has $F_n(x)$ as a unique solution.*

Proof. Rewriting the differential equation as $G'(x) = G(x)^2 + G(x)A_n(x)$, we obtain a Bernoulli differential equation which has a unique solution with the initial condition $G(0) = 1$. It is easy to verify that $F_n(0) = 1$ and that F_n satisfies this differential equation (in fact our construction of F_n follows the solution of the Bernoulli differential equation). □

Lemma 5.5. *For some $\epsilon_n > 0$, $F_n(z)$ as a function of a complex variable z is meromorphic on $\{z : |z| < r_n + \epsilon_n\}$, with its only pole in this disk at $z = r_n$.*

Proof. By construction, A_n, B_n, C_n, D_n are all entire, and their series expansions around $z = 0$ have all nonnegative real coefficients with at least one positive coefficient, so $F_n = \frac{C_n}{1-D_n}$ is meromorphic and nonconstant. Since $C_n(r_n) > 0$ and $|D_n(z)| < 1$ for $|z| < r_n$ due to the nonnegative real coefficients, F_n has a pole at r_n and no other poles in $|z| \le r_n$ by the triangle inequality. The only poles are at roots of $D_n(z) = 1$, of which there are only finitely many in the compact set $|z| \le 2r_n$. It follows that for some $\epsilon_n > 0$, r_n is the only pole in $|z| < r_n + \epsilon_n$, as desired. □

Proof of Theorem 1.9. We will show that $\limsup_{k\to\infty} \left(\frac{f_k}{k!}\right)^{1/k} \le \frac{1}{r}$ and $\liminf_{k\to\infty} \left(\frac{f_k}{k!}\right)^{1/k} \ge \frac{1}{r_n}$ for all n. Then, we have that $\lim_{k\to\infty} \left(\frac{f_k}{k!}\right)^{1/k} = \frac{1}{r}$, so Stirling's approximation implies that $\lim_{n\to\infty} \frac{f_n^{1/n}}{n} = \frac{1}{re}$.

Note that $F(x)$ solves the differential equation $G'(x) = G(x)^2 + G(x)A(x)$ with initial condition $G(0) = 1$, which is a Bernoulli differential equation with unique solution $\frac{C(x)}{1-D(x)}$ by construction. Thus, $F(x) = \frac{C(x)}{1-D(x)}$.

Now, we show that $F(x)$ converges for $x \in [0, r)$. Recall that $r \le R$, where R is the common radius of convergence of A, B, C, D. Thus, since $F(x) = \frac{C(x)}{1-D(x)}$, $F(x)$ converges as long as $x \in [0, R)$ and $D(x) < 1$. By definition, $F(x)$ converges for $x \in [0, r)$. It thus follows that the radius of convergence of F is at least r, so $\limsup_{k\to\infty} \left(\frac{f_k}{k!}\right)^{1/k} \le \frac{1}{r}$, as desired.

Finally, let $F_n(x) = \sum_{k\ge 0} \frac{a_k}{k!} x^k$. Note that since $F_n(z)$ is meromorphic on $|z| < r_n + \epsilon_n$ with its only pole at r_n, we have that $\lim_{k\to\infty} \left(\frac{a_k}{k!}\right)^{1/k} = r_n$ by [12, Theorem IV.10]. Thus, it suffices to show that $f_k \ge a_k$ for all k. Let $T_n(x) = \log F_n(x)$, so $F_n(x) = e^{T_n(x)}$ and $T_n(0) = 0$ since $F_n(0) = 1$. The differential equation $\frac{F_n'(x)}{F_n(x)} - F_n(x) = A_n(x)$ can then be rewritten as $T_n'(x) - e^{T_n(x)} = A_n(x)$. Suppose that $T_n(x) = \sum_{k\ge 0} \frac{b_k}{k!} x^k$. Equating coefficients in the differential equation, we have that $b_{k+1} - E(b_0, \ldots, b_k) = t_{k+1} - f_k$ for $k \le n$ and $b_{k+1} - E(b_0, \ldots, b_k) = 0$ for $k > n$. But we know that $t_{k+1} - E(t_0, \ldots, t_k) = t_{k+1} - f_k$ for all k and $t_0 = b_0 = 0$, so $t_k = b_k$ and thus $f_k = E(t_0, \ldots, t_k) = E(b_0, \ldots, b_k) = a_k$ for all $k \le n+1$ by strong induction. For $k > n+1$, we proceed using strong induction to show that $t_k \ge b_k$ and $f_k \ge a_k$, with the base case of $k \le n+1$ already shown. For the inductive step, we have that $b_{k+1} = E(b_0, \ldots, b_k) = a_k \le f_k \le t_{k+1}$ by Proposition 5.2 and $a_{k+1} = E(b_0, \ldots, b_{k+1}) \le E(t_0, \ldots, t_{k+1}) = f_{k+1}$, where we are using the monotonicity of E for nonnegative inputs, so we are done.

To see that $\lim_{n\to\infty} \frac{f_n^{1/n}}{n} = \lim_{n\to\infty} \frac{t_n^{1/n}}{n}$, we again make use of the inequality $t_{k+1} \ge f_k$. We have that $f_{k-1} \le t_k \le f_k$. Taking kth roots, dividing by k, and taking the limit, we have that $\lim_{k\to\infty} \frac{f_{k-1}^{1/k}}{k} \le \lim_{k\to\infty} \frac{t_k^{1/k}}{k} \le \lim_{k\to\infty} \frac{f_k^{1/k}}{k}$. But $\lim_{k\to\infty} f_{k-1}^{1/(k(k-1))} = 1$ as $f_{k-1} \le (k-2)^k$. Thus, $\lim_{k\to\infty} \frac{f_{k-1}^{1/k}}{k} = \lim_{k\to\infty} \frac{f_{k-1}^{1/(k-1)}}{k-1} = \lim_{k\to\infty} \frac{f_k^{1/k}}{k}$, and the result follows. □

This proof allows us to compute convergent lower bounds for the Stanley-Wilf limit for S, which we discuss in Subsection 5.2.

Remark 5.6. The condition $t_{k+1} \ge f_k$ that we used may be replaced by $t_{k+1} \ge cf_k$ for any $c > 0$, and the proof is essentially the same with minor modifications. However, this condition seems to be difficult to show for covered sets S even for $c < 1$, and the easy proof of Proposition 5.2 does not carry over. As long as $f_n = O(t_{n+1})$, the forest Stanley-Wilf limit exists. We believe that the limit also exists when $t_{n+1} = o(f_n)$ but that there are fundamental differences between sets S satisfying $f_n = O(t_{n+1})$ and sets S satisfying $t_{n+1} = o(f_n)$. We will remark more on these differences in Subsection 5.2 and Section 6.

Remark 5.7. The series F_n we used to approximate F from below has a combinatorial interpretation. One viewpoint, essentially given in the proof of Theorem 1.9, is that we initially force equality to hold in $t_{k+1} \geq f_k$ for all k, and then we iteratively replace t_n with its true value for all n (note that f_n is determined by $t_1, \ldots, t_n$). In this way, the coefficients of F_n agree with the coefficients of F up to x^n, and as $n \to \infty$, F_n converges coefficientwise to F. However, we can also view F_n as the exponential generating function of the combinatorial class $\mathscr{F}_n$ of forests that avoid S along with the stronger condition that every vertex with more than n descendants has the smallest label among all of its descendants. The asymptotics for such forests in $\mathscr{F}_n$ are given by r_n, which converge to r as $n \to \infty$ by our proof. Heuristically, $\mathscr{F}_n$ forms a good approximation for $\mathscr{F}$ because in a typical forest, we expect most vertices to not have too many descendants. Furthermore, if a vertex has many descendants, than in order to avoid S it is intuitively more efficient for S to have a small label since the patterns in S do not start with 1, especially if S contains many patterns. This relates to the rate at which r_n converges to r, which our proof gives no insight into.

Remark 5.8. Throughout our proof, the pattern avoidance condition is only relevant for Proposition 5.2 to establish $t_{k+1} \geq f_k$, and after that the proof relies on the analytic interpretation of the relation between the trees and forests in a combinatorial class of rooted labeled forests. Consequently, the proof is quite robust and immediately generalizes to give forest Stanley-Wilf limits for avoiding consecutive patterns, generalized patterns, any type of pattern in which the smallest element does not come first, and arbitrary combinations thereof. The limit's existence is not driven by the pattern avoidance, but rather by the tree-forest structure in the combinatorial class. Thus, we believe that our techniques may also be useful in asymptotically enumerating other types of rooted labeled forests that may be unrelated to pattern avoidance.

5.2. **Determining forest Stanley-Wilf limits.**

We now turn to the problem of finding the value of the forest Stanley-Wilf limit for a given set S of patterns. Much of our work in this subsection also applies to asymptotics for consecutive- or generalized-pattern-avoiding forests, and we leave such computations to the interested reader.

For a set S of patterns, let $L_S = \lim_{n\to\infty} \frac{f_n^{1/n}}{n}$ denote the forest Stanley-Wilf limit for S. By Theorem 1.9, L_S exists for all uncovered sets S. We will show the existence of L_S for a few other sets in this subsection. We will also drop braces in the subscript in L_S, so for example we will write $L_{123,231}$ instead of $L_{\{123,231\}}$.

The proof of Theorem 1.9 given in Subsection 5.1 allows us to compute convergent lower bounds for L_S. Indeed, note that $\dfrac{1}{er_n}$ increases to L_S, where r_n is the unique positive root of $D_n(x) = 1$, as previously defined. The functions A_n, B_n, C_n, D_n are determined by t_{k+1} and f_k for $k \leq n$, so we are able to estimate r_n by computing the sequences t_k and f_k up to $n+1$. Anders and Archer provide many explicit formulas for f_n for certain sets in [1], and Garg and Peng give many recursions for f_n for some other sets in [17]. Using these, we are able to find lower bounds for L_S for certain S displayed in Figure 17.

Here, n denotes the amount of terms we computed, and the lower bound in the proven column corresponds to the one found with solving $D_n(x) = 1$. The conjectured column contains five conjectured values of L_S given by Garg and Peng in [17, Conjecture 7.2] and five conjectured upper bounds based on our computations. In all of the cases we computed, the sequence $\frac{f_k^{1/k}}{k}$ was decreasing for $k \leq n$, and our five conjectured bounds correspond to the value of $\frac{f_n^{1/n}}{n}$. We have also added in any nontrivial forest-Wilf equivalences in the S column. We did not include results of complementation in this column, but clearly those sets also have the same forest Stanley-Wilf limit.

While our proven lower bounds on L_S are relatively close to the conjectured approximate values and upper bounds, in order to compute L_S to arbitrary precision, one would need a method to prove

S	n	Proven	Conjectured
123 132	350	≥ 0.6766	≈ 0.6801
213	2500	≥ 0.65493	≈ 0.65521
$123, 213$ $132, 213$	1700	≥ 0.555617	≈ 0.555843
$123, 231$	800	≥ 0.5402	≈ 0.5530
$132, 231$	1000	≥ 0.58145	≤ 0.58421
$213, 231$	2500	≥ 0.557725	≈ 0.557864
$123, 132, 213$	1650	≥ 0.51781	≤ 0.51939
$123, 132, 231$	2500	≥ 0.53057	≤ 0.53169
$132, 213, 231$	2500	≥ 0.48241	≤ 0.48317
$123, 2413, 3412$	1800	≥ 0.62765	≤ 0.62939

FIGURE 17. Computed lower bounds for L_S.

convergent upper bounds on L_S as well. Unfortunately, we were not able to adapt our methods from the proof of Theorem 1.9 to obtain upper bounds from the first few terms of f_n. A natural step would be to replace the inequality $t_{k+1} \geq f_k$ with the inequality $t_k \leq kf_{k-1}$. This inequality follows from the observation that a tree on $[k]$ that avoids S consists of a root vertex with label a and a forest on $[k-1] \setminus \{a\}$ that avoids S. There are k choices for a and for each choice of a, there are at most f_{k-1} forests on the remaining $k-1$ vertices that work, yielding the claimed bound of kf_{k-1}. Note that equality holds in the inequalities $t_{k+1} \geq f_k$ and $t_k \leq kf_{k-1}$ when $S = \{21\}$ and $S = \varnothing$, respectively. The method in the proof of Theorem 1.9 can be viewed as starting with forests avoiding 21, i.e. increasing forests, and iteratively adding in more forests that avoid S corresponding to using higher truncations of $A(x) = \sum_{k\geq 0} \frac{t_{k+1}-f_k}{k!}x^k$. We can try to take a similar approach with the upper bound, starting with all forests and iteratively removing more forests that do not avoid S corresponding to higher truncations of $P(x) = \sum_{k\geq 1} \frac{kf_{k-1}-t_k}{k!}x^k$. Instead of a differential equation, we get the equation $xe^{T(x)} - T(x) = P(x)$ for $T(x)$, which we can attempt to approximate with $T_n(x)$ satisfying $xe^{T_n(x)} - T_n(x) = P_n(x) = \sum_{1\leq k\leq n} \frac{kf_{k-1}-t_k}{k!}x^k$. Note that when $P(x) = 0$, we recover the equation $xe^{T(x)} = T(x)$, the functional equation for the Cayley tree function (see [12, Section II.5.1]). We would like for the growth rate of the coefficients of the T_n that solves $xe^{T_n} - T_n = P(x)$, or $T_n = xe^{T_n} - P_n(x)$, to be in the smooth implicit-function schema defined in [12, Section VII.4.1], in which case we can recover an upper bound for L_s. However, the presence of negative coefficients in the $-P_n(x)$ on the right-hand side makes this impossible. The example given at the end of [12, Section VII.4.1] shows that such negative coefficients can lead to pathological situations. It would be interesting to somehow repair this method or find a different way to compute upper bounds on L_S.

While Theorem 1.9 only shows the existence of L_S for uncovered S, it is possible to show that L_S exists in other cases as well. For example, we can classify all of the sets S of patterns satisfying $L_S = 0$.

Proposition 5.9. *We have that* $\lim_{n\to\infty} \frac{f_n^{1/n}}{n} = 0$ *if and only if* S *contains the patterns* $1\cdots k$ *and* $\ell\cdots 1$ *for some* k *and* ℓ.

Proof. Note that forests avoiding $1\cdots k$ and $\ell\cdots 1$ must have depth at most $k\ell$ by the Erdős-Szekeres Theorem. We will show that if $f_{m,n}$ and $t_{m,n}$ are respectively the number of forests and trees on $[n]$ of depth at most m, then $\lim_{n\to\infty} \frac{f_{m,n}^{1/n}}{n} = \lim_{n\to\infty} \frac{f_{m,n}^{1/n}}{n} = 0$. Let $F_m(x) = \sum_{k=0}^{\infty} \frac{f_{m,k}}{k!}x^k$ and $T_m(x) = \sum_{k=0}^{\infty} \frac{t_{m,k}}{k!}x^k$ denote the exponential generating functions of the sequences $\{f_{m,n}\}$

and $\{t_{m,n}\}$. We have by standard manipulations of labeled combinatorial classes and exponential generating functions that $F_m = e^{T_m}$ and $T_{m+1} = xF_m$ for all m. As $T_1 = x$, it follows by induction on m that $F_m(z)$ and $T_m(z)$ are entire functions in $z \in \mathbb{C}$ for all m. Thus, we have that $\lim_{n\to\infty}\left(\frac{f_{m,n}}{n!}\right)^{1/n} = \lim_{n\to\infty}\left(\frac{t_{m,n}}{n!}\right)^{1/n} = 0$, so by Stirling's approximation,

$$0 \le \lim_{n\to\infty}\frac{f_n^{1/n}}{n} = e\lim_{n\to\infty}\left(\frac{f_n}{n!}\right)^{1/n} \le e\lim_{n\to\infty}\left(\frac{f_{m,n}}{n!}\right)^{1/n} = 0.$$

On the other hand, note that if all increasing forests avoid S, then $f_n \ge n!$ so we have that $\liminf_{n\to\infty}\frac{f_n^{1/n}}{n} \ge e^{-1}$. Thus, S must contain a pattern of the form $1\cdots k$. The same holds for decreasing forests, so S must also contain a pattern of the form $\ell\cdots 1$, as desired. ☐

Corollary 5.10. *When L_S exists, it lies in $\{0\}\cup[e^{-1},1]$.*

By examining our proof of Theorem 1.9, we can also determine when an uncovered set S satisfies $L_S = e^{-1}$.

Proposition 5.11. *If S is an uncovered set of patterns, then $L_S = e^{-1}$ if and only if S contains* 12 *or* 21*.*

Proof. If $21 \in S$, then no patterns in S can start with 1, so all other patterns contain 21 and are superfluous. It then follows that $f_n = n!$ so $L_S = e^{-1}$. The same exact argument works for if $12 \in S$. In the other direction, note that $\frac{1}{er_1}$ is a lower bound for L_S. If S does not contain 12 or 21, then we have that $t_2 = 2$ while $f_1 = 1$, so $A_1(x) = x$ and $D_1(x) = \int_0^x e^{t^2/2}dt$. It is then clear that $D_1(1) > 1$ so $r_1 < 1$ and $L_S > \frac{1}{e}$, as desired. ☐

This proposition shows that a small change to the number of S-avoiding trees and forests for a small number of vertices already results in a strictly larger forest Stanley-Wilf limit. The asymptotics of f_n seem to be quite sensitive to changes in t_k and f_k for small k, at least for uncovered sets S. This is in sharp contrast with the situation for permutations. For example, there is only one permutation of $[n]$ that avoids 21, namely $1,\dots,n$. We then consider permutations of $[n]$ that avoid $\{213,231,312,321\}$, the set of all patterns of length at least 3 that do not start with 1. For $n>1$ there are only two such permutations, given by $1\cdots n$ and $1\cdots(n-2)n(n-1)$. The discrepancy between the number of permutations of $[n]$ that avoid 21 and $\{213,231,312,321\}$ for $n=2$ is not magnified for larger n. For forests, however, there are exponentially many more forests on $[n]$ avoiding $\{213,231,312,321\}$ than there are forests on $[n]$ avoiding 21. This can intuitively be explained by the observation that there are generally many ways to perturb an increasing forest into another forest that still avoids $\{213,231,312,321\}$. Any vertex whose children are all leaves can swap labels with one of its children, and the resulting forest will still avoid $\{213,231,312,321\}$ (see Figure 18 for an example). In contrast, when one tries to apply this to the increasing path, corresponding to the permutation $1,\dots,n$, there is only one way to do so which results in the one other permutation avoiding $\{213,231,312,321\}$. Heuristically, discrepancies between f_k for small k manifest close to the leaves of the forest. There are generally relatively many vertices close to leaves, so the discrepancy is magnified into a strictly larger limit.

For $S=\{213,231,312,321\}$, we can give more explicit properties of f_n and t_n.

Proposition 5.12. *For $S=\{213,231,312,321\}$, the exponential generating function $T(x)$ of t_n satisfies the differential equation $T' = T + e^T$ with initial condition $T(0)=0$.*

Proof. By definition, $T(0)=0$, so it suffices to show that $T(x)$ satisfies $T' = T+e^T$. In terms of the coefficients, this reduces to showing the identity $t_{k+1} = f_k + t_k$.

We prove that $t_{k+1} = f_k + t_k$ by casework on where the label 1 is in a tree on $[k+1]$ avoiding S. If 1 is at the root, then the rest of the tree must be a forest on $\{2,\dots,k+1\}$ that avoids S,

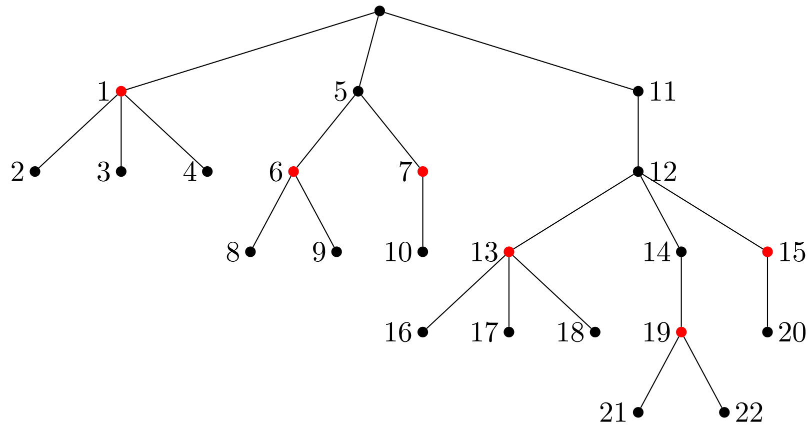

FIGURE 18. Any of the labels of the red vertices, which have labels $1, 6, 7, 13, 15, 19$, of this increasing forest can be swapped with one of its children, and the resulting forest will still avoid the set $\{213, 231, 312, 321\}$.

and any such forest will work, resulting in f_k such trees. If 1 is not at the root, then it cannot have any children. We also cannot have the vertex labeled 1 be at depth more than 2. Thus, the vertex labeled 1 must be a child of the root of the tree. Deleting this vertex results in a tree on $\{2, \ldots, k+1\}$, and any such tree can be turned into a tree on $[k+1]$ avoiding S by adding a vertex labeled 1 as a child of the root. The identity $t_{k+1} = t_k + f_k$ then follows, and the proposition is proven. □

This proposition tells us that the number of trees on $[n]$ avoiding S is $\left((x + e^x)\frac{d}{dx}\right)^n x$ evaluated at $x = 0$. By numerically approximating the singularity of the solution to this differential equation, we can obtain an approximation of L_S.

Corollary 5.13. *We have that* $L_{213,231,312,321} \approx 0.4562$.

This is indeed greater than $e^{-1} \approx 0.3679$. Notably, we are able to give an approximation of L_S here instead of just a lower bound because we have an explicit differential equation that $T(x)$ satisfies. Even if the differential equation is not explicitly solveable, we can numerically approximate L_S. It seems to be very rare that this is possible, and none of the other uncovered sets S other than the ones containing 21 seem to satisfy any simple differential equation.

With all of the limits computed so far, one might conjecture that having the same forest Stanley-Wilf limit implies forest-Wilf equivalence. While this may be the case for uncovered sets of patterns, it is not true in general.

Proposition 5.14. *For* $S = \{132, 231, 321\}$, *we have that* $t_n = n!$, $T(x) = \frac{x}{1-x}$, $F(x) = e^{\frac{x}{1-x}}$, *and* $L_{132,231,321} = e^{-1}$.

Proof. We first show the following characterization of trees on $[n]$ avoiding S. They are the trees that have an arbitrary root label but are otherwise increasing. Indeed, to avoid the patterns in S, we cannot have any instances of 21 not including the root. But as long as no such instances exist, we avoid S. There are n ways to select a label for the root and $(n-1)!$ ways to choose the increasing forest underneath the root, for a total of $n!$ ways, as desired.

Consequently, the exponential generating function of t_n is $T(x) = \frac{x}{1-x}$. We then have that the exponential generating function of f_n is $F(x) = e^{\frac{x}{1-x}}$. The radius of convergence of $T(x)$ and $F(x)$ is 1, so we have that $\limsup_{n\to\infty} \frac{f_n^{1/n}}{n} \leq e^{-1}$ by Stirling's approximation. But $f_n \geq n!$ as all increasing forests avoid S, so $\liminf_{n\to\infty} \frac{f_n^{1/n}}{n} \geq e^{-1}$, and we obtain the result that $L_{132,231,321} = e^{-1}$. □

Note that Proposition 5.11 does not apply here because S is not uncovered. We know that the inequality $t_{k+1} \geq f_k$ cannot hold for all k, or the same proof for Theorem 1.9 and Proposition 5.11 would apply. Indeed, we have that $t_9 = 362880$ while $f_8 = 394353$. We do not even have that $f_n = O(t_{n+1})$ here. Vaclav Kotesovec gives the asymptotic $f_n \sim \frac{1}{\sqrt{2e}} n^{n-1/4} e^{2\sqrt{n}-n}$ on the OEIS for f_n [28]. While we were able to show that the limit exists in this case, the fact that $t_{n+1} = o(f_n)$ for this covered set S suggests that we will not be able to modify our proof of Theorem 1.9 to work in general.

6. Future Work

We conclude this paper by discussing several conjectures, open questions, and potential directions for future research.

6.1. **Further (c-)forest-Wilf equivalences.**

We first ask whether there are any more forest-Wilf equivalences among the cases that we examined.

Question 6.1. *Does there exist a nontrivial forest-Wilf equivalence between two patterns that is not given by Theorem 1.1?*

Question 6.2. *Does there exist a nontrivial forest-Wilf equivalence between two sets consisting of a pattern of length* 3 *and a pattern of length* k *for* $k > 4$ *that is not given by Theorem 1.3?*

Garg and Peng have done a computer check up to patterns of length 5, finding no such unexpected patterns [18]. Unfortunately, the computation is rather time-consuming for longer patterns.

We ask a similar question about c-forest-Wilf equivalences.

Question 6.3. *Does there exist a nontrivial c-forest-Wilf equivalence between two patterns of length greater than* 4 *that is not given by Lemma 4.21?*

It is possible that the construction of Garg and Peng (Theorem 1.4) generalizes, though we were unable to find such a generalization. Based on the currently known c-forest-Wilf equivalences, we make the following conjecture.

Conjecture 6.4. *Suppose that* π *and* π' *are patterns of length* k *that are nontrivially c-forest-Wilf equivalent. Then* $\pi(1), \pi'(1) \in \{1, k\}$.

Nakamura conjectured in [27] that the notions of strong c-Wilf equivalence and c-Wilf equivalence are the same for permutations. We make the same conjecture for forests.

Conjecture 6.5. *If* π *and* π' *are patterns with* $\pi \overset{c}{\sim} \pi'$, *then* $\pi \overset{sc}{\sim} \pi'$.

Finally, we conjecture that a stronger necessary condition for c-forest-Wilf equivalence between grounded permutations can be shown.

Conjecture 6.6. *If* π *and* π' *are grounded patterns of length* k *that are strongly c-forest-Wilf equivalent, then for all* $1 < i \leq k$, *the height of* $\pi(i)$ *in* π *is equal to the height of* $\pi'(i)$ *in* π'.

6.2. **Asymptotics and forest Stanley-Wilf limits.**

Conjecture 1.8 is still unproven for covered sets of patterns. In the case of uncovered sets S, there remains the problem of finding the value of L_S to arbitrary precision, as it does not seem possible in general to find differential equations for the exponential generating functions.

Question 6.7. *Is there an algorithm that computes convergent upper bounds on* L_S *for uncovered sets* S?

Beyond this, we believe that L_S should satisfy certain "monotonicity" properties.

Conjecture 6.8. *If* $L_S = 1$, *then* $S = \varnothing$.

Conjecture 6.9. *If π and σ are different patterns such that π contains σ, then $L_\pi > L_\sigma$.*

One possible way to resolve Conjecture 6.8 is to find an algorithm that answers Question 6.7 and analyze when the upper bounds it gives are always 1. Note that Proposition 5.11 shows Conjecture 6.9 when $\sigma \in \{12, 21\}$. The main difficulty in generalizing our proof seems to be obtaining a comparison between $t_{k+1}(\pi) - f_k(\pi)$ and $t_{k+1}(\sigma) - f_k(\sigma)$. All we currently know is that these are nonnegative and equal to 0 for 12 and 21, which is only sufficient to prove the connjecture for $\sigma \in \{12, 21\}$.

We also have the following conjecture about sharper asymptotics for f_n.

Conjecture 6.10. *For an uncovered set S of patterns, there exist constants a_S and b_S such that $\frac{f_n}{n!} \sim a_S n^{b_S} (eL_s)^n$.*

Based on limited data, it seems that $b_S = 0$ for nonempty S, while by Cayley's formula for $S = \varnothing$ we have that $\frac{f_n}{n!} \sim \frac{e}{\sqrt{2\pi}} n^{-3/2} e^n$. The case that $S = \varnothing$ seems to be fundamentally different. The asymptotics for covered sets also seem to be very different. For example, for $S = \{1 \cdots k, \ell \cdots 1\}$, we have that $L_S = 0$, but clearly $\frac{f_n}{n!} \not\to 0$. Taking $k = 3$ and $\ell = 2$, forests avoiding S become increasing forests of depth at most 2. Such forests are in bijection with partitions of the label set $[n]$, so f_n is given by the nth Bell number B_n. The asymptotics of B_n are much more complicated than the behavior predicted by Conjecture 6.10 for uncovered sets. Yet another example is given by $S = \{132, 231, 321\}$ from Proposition 5.14, where $\frac{f_n}{n!} \sim \frac{1}{\sqrt{4\pi e}} n^{-3/2} e^{2\sqrt{n}}$.

Our heuristic for Conjecture 6.10 is that for uncovered sets, F is reasonably approximated by series F_m that have a meromorphic continuation to $\mathbb{C}$. The coefficients of these series all satisfy the type of asymptotic described in the statement of the conjecture, so we believe that F satisfies a similar estimate. This extends to any sets S satisfying $f_n = O(t_{n+1})$ as well.

We in fact predict that the condition $f_n = O(t_{n+1})$ is what distinguishes uncovered sets and covered sets.

Conjecture 6.11. *A set S of patterns is uncovered if and only if it satisfies $f_n = O(t_{n+1})$.*

Given a forest on $[n]$, there are $n + 1$ ways we can extend this to a tree on $[n + 1]$. We choose a root label a for the tree in $[n + 1]$ and the rest of the tree is the given forest, relabeled with $[n + 1] \setminus \{a\}$. The quantity $\frac{t_{n+1}}{f_n}$ can be interpreted as the expected number of root labels we can choose for a uniform random forest on $[n]$ avoiding S such that the resulting tree on $[n + 1]$ also avoids S. For uncovered S, 1 or $n + 1$ is always a valid choice, so this expected value is always at least 1. We predict that this expected value tends to 0 for covered sets S. Small roots are unlikely to be possible because of the pattern in S starting with 1, and large roots are unlikely to be possible because of the pattern in S starting with its largest element. While it may be possible that moderately sized roots can keep the expected value high, we conjecture that this is not the case.

One way to lower bound L_S for a covered set $S = \{\pi_1, \ldots, \pi_m\}$ is to consider the limit L'_S for $S' = \{\pi'_1, \ldots, \pi'_m\}$, where π'_i is a subpattern of π_i and S' is an uncovered set. We conjecture that this is also how L_S is achieved, i.e. that there cannot be exponentially more ways to avoid S than there are to avoid S' for the best choice of S'.

Conjecture 6.12. *Define the* reduction $\widehat{\pi}$ *of a pattern $\pi = \pi(1) \cdots \pi(k)$ to be the pattern of length $k - 1$ in the same relative order as $\pi(2) \cdots \pi(k)$. Let $S = \{\pi_1, \ldots, \pi_m\}$ be a covered set of patterns, and let $S_i = S \setminus \{\pi_i\} \cup \{\widehat{\pi}_i\}$. Then $L_S = \max_{1 \le i \le m} L_{S_i}$.*

We can repeatedly replace patterns in S with their reductions until S is an uncovered set, and this yields a lower bound on L_S. The conjecture is that L_S is equal to the maximum lower bound achieved in this way. For example, this conjecture predicts that $L_{132,4213} = L_{132,213}$. Note that if this conjecture were true, it would provide an answer to the following question.

Question 6.13. *What are the possible values of L_S?*

The answer would then be the values of L_S over all uncovered sets S, which we are able to estimate. Conjecture 6.12 also provides a path to proving Conjecture 1.8, as it then remains to show that the values of f_n do not exceed the natural lower bounds from reducing the patterns in S by an exponential factor.

It is possible that Conjecture 6.12 is false even in simple cases such as $S = \{132, 312\}$. However, it is consistent with Propositions 5.9 and 5.14. In those cases, the values of f_n exceed the corresponding natural lower bounds by a subexponential factor, on the order of the Bell numbers or $\exp(O(\sqrt{n}))$. These can be interpreted as a result of more wildly behaved singularities of the exponential generating function $F(x)$ in the neighborhood of $\frac{1}{eL_S}$. Indeed, for a covered set of patterns, we no longer have the same approximation by meromorphic functions as in the uncovered case, which heuristically suggests more erratic behavior at the singularity.

We make one last generalization of forest Stanley-Wilf limits. Say that a rooted labeled forest F_1 *contains* another rooted labeled forest F_2 if there exists a graph minor of F_1 that is isomorphic to F_2 and whose corresponding labels are in the same relative order as F_2. For example, the type of pattern avoidance we have been studying in this paper can be viewed as forests avoiding a rooted labeled path. Similar to the closed permutation classes considered in [23], we can define a *closed forest class* Π to be a collection of rooted labeled forests such that if a forest $F_1 \in \Pi$ contains a forest F_2, then $F_2 \in \Pi$. We can make the following general conjecture about the growth rates of closed forest classes.

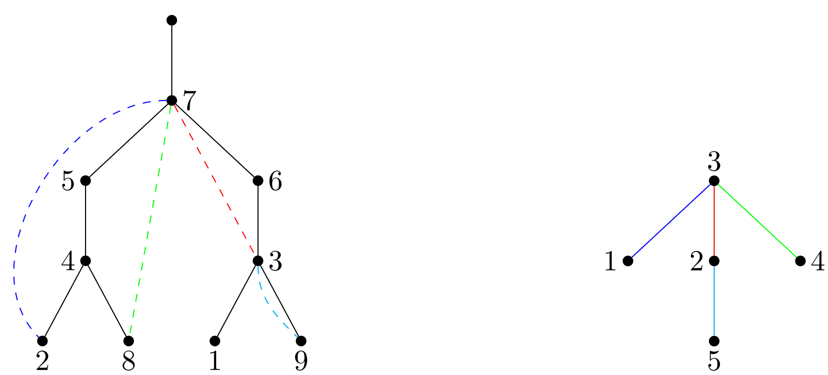

FIGURE 19. The forest on the left contains the forest pattern on the right. Note that our forests are unordered, so the branches of the pattern can appear in a different order in the forest.

Conjecture 6.14. *Let Π be a closed forest class, and let Π_n denote the set of forests on $[n]$ in Π. Then $\lim_{n\to\infty} \frac{|\Pi_n|^{1/n}}{n}$ exists.*

Note that forests avoiding a set S of patterns form a closed forest class, so this can be seen as a generalization of Conjecture 1.8. It would also be interesting to study forests that avoid a set S of *forest patterns* consisting of rooted labeled forests, where the avoidance and containment is in the sense described above for forests. Figure 19 gives an example of containment of a forest pattern. This is somewhat reminiscent of the poset pattern avoidance studied by Hopkins and Weiler in [20]. However, restricting to the setting of forest patterns allows us to carry over results we have shown in this paper. For example, the proof for Theorem 1.9 automatically gives the existence of the forest Stanley-Wilf limit for certain sets of forest patterns.

6.3. Limiting distributions for forest statistics.

Finally, we make some conjectures about how certain forest statistics are distributed in the limit $n \to \infty$.

Certain results in permutation pattern avoidance look at how permutation statistics such as inversions and ascents are distributed when we look at permutations avoiding certain patterns rather than the whole symmetric group (see, for example, [7]). Forests come with their own set of interesting statistics that seem to yield interesting limit distributions. We make a few conjectures about the root of a tree on $[n]$ avoiding S and the number of trees in a forest on $[n]$ avoiding S.

For a set S of patterns, let $R_{S,n}$ denote the label of the root of a uniform random tree on $[n]$ avoiding S, let $T_{S,n}$ denote the number of trees in a uniform random forest on $[n]$ avoiding S, and let $T_{S,n,k}$ denote the number of trees with k vertices in a uniform random forest on $[n]$ avoiding S.

Conjecture 6.15. *For any set S of patterns, there exists a random variable R_S such that $\frac{R_{S,n}}{n}$ converges in law to R_S as $n \to \infty$.*

Note that the limiting distribution can be continuous, such as a uniform distribution when $S = \{132, 231, 321\}$ by Proposition 5.14, or discrete, such as a convergence to 0 when $S = \{21\}$. When S is uncovered, we expect most of the trees to have root labels that are very small or very large. Heuristically, the "easiest" way to avoid S when S is uncovered is to have the root have label close to 1 or n. In the case that S contains a pattern starting with 1, this is no longer true if our root label is 1, but we can still have a root label close to n, and vice versa if S contains a pattern starting with its largest element. We have the following stronger conjecture that formalizes this.

Conjecture 6.16. *For any uncovered set S of patterns, $\frac{R_{S,n}}{n}$ converges in distribution to a Bernoulli random variable $Ber(p)$ for some $p \in [0,1]$. If S contains a pattern starting with 1, then $p = 1$, and if S contains a pattern starting with its largest element, then $p = 0$. Furthermore, there exist limiting probabilities $p_1, p_2, \ldots, q_1, q_2, \ldots$ summing to 1 such that $\mathbb{P}(R_{S,n} = k) \to p_k$ and $\mathbb{P}(R_{S,n} = n+1-k) \to q_k$ as $n \to \infty$. If S contains a pattern starting with 1, then $p_1 = p_2 = \cdots = 0$, and if S contains a pattern starting with its largest element, then $q_1 = q_2 = \cdots = 0$.*

Some data computed for $S = \{213\}$ and $S = \{123\}$ supports this conjecture, but we do not have any data for covered sets S.

We now turn to the distribution of $T_{S,n}$ as $n \to \infty$. Our main motivation comes from the fact that for $S = \{21\}$, i.e. for increasing forests, there exists a bijection between forests on $[n] = P_1 \sqcup \cdots \sqcup P_m$ with m components such that the labels in the components are $P_1, \ldots, P_m$ and permutations of $[n]$ with m cycles such that the elements in the cycles are $P_1, \ldots, P_m$. A classical result of Goncharov in [14, 15] states that in a uniform random permutation π of $[n]$, the number of cycles C_n in π is asymptotically normal: $\frac{C_n - \mathbb{E}[C_n]}{\mathrm{Var}(C_n)}$ converges in distribution to a standard Gaussian. Furthermore, $\mathbb{E}[C_n], \mathrm{Var}(C_n) \sim \log n$. Another result in this area, due to Arratia and Tavar in [3], is that if $C_{n,k}$ is the number of cycles in π of length k, then $(C_{n,1}, C_{n,2}, \ldots)$ converges in distribution to $(Z_1, Z_2, \ldots)$ as $n \to \infty$, where $Z_1, Z_2, \ldots$ are independent Poisson random variables with $\mathbb{E}[Z_k] = k^{-1}$. The correspondence between trees in increasing forests and cycles in permutations immediately gives us these results but for $T_{S,n}$ instead of C_n for $S = \{21\}$. For example, Goncharov's theorem implies that $\frac{T_{S,n} - \mathbb{E}[T_{S,n}]}{\mathrm{Var}(T_{S,n})}$ converges in distribution to a standard Gaussian as $n \to \infty$. We conjecture that these results also hold for other sets of patterns.

Conjecture 6.17. *For all nonempty sets S of patterns, the random variable $T_{S,n}$ is asymptotically normal. We have that $\frac{T_{S,n} - \mathbb{E}[T_{S,n}]}{Var(T_{S,n})}$ converges in distribution to a standard Gaussian as $n \to \infty$.*

Furthermore, if S is uncovered, then $\mathbb{E}[T_{S,n}]$, $Var(T_{S,n}) = \Theta(\log n)$, and $(T_{S,n,1}, T_{S,n,2}, \ldots)$ converges in distribution to $(Z_1, Z_2, \ldots)$, where $Z_1, Z_2, \ldots$ are independent Poisson random variables with $\mathbb{E}[Z_k] = \Theta(k^{-1})$.

Note that the hypothesis on S being nonempty is necessary. When $S = \varnothing$, the total number of rooted forests on $[n]$ is $(n+1)^{n-1}$ and the total number of rooted trees on $[n]$ is n^{n-1}, so $\mathbb{P}(T_{S,n} = 1) \to \frac{1}{e}$ as $n \to \infty$ and $T_{S,n}$ cannot be asymptotically normal in this case. The behavior of the limiting distribution is related to the behavior of the exponential generating function $F(x)$ around its singularity $\frac{1}{eL_S}$ by [12, Section IX.4]. Indeed, as mentioned previously, for uncovered S we expect $F(x)$ to be well-behaved because of the approximation by $F_m(x)$, which has a meromorphic continuation to $\mathbb{C}$. However, this shows a shortcoming of our method, which does not distinguish between when S is empty and when S is nonempty. More sophisticated analysis of the singularity of $F(x)$ is needed if we are to prove the conjecture using this approach. Data computed for all of the uncovered sets we considered in this section support this conjecture. On the other hand, $T_{S,n}$ does seem to be asymptotically normal for covered sets S as well, but the point of concentration is different, most likely due to the different behavior of $F(x)$ around its singularity. In the case of $S = \{132, 231, 321\}$, it appears that $\mathbb{E}[T_{S,n}] \sim \sqrt{n}$. The case of $S = \{123, 21\}$ is equivalent to the distribution of Stirling numbers of the second kind. This problem was considered by Harper in [19], and Harper's result translates to the asymptotic normality of $T_{S,n}$. The mean, however, is of a different order than $\sqrt{n}$ and $\log n$. It appears that a variety of asymptotics can occur for the mean of $T_{S,n}$ for covered sets, in contrast to uncovered sets.

It would also be interesting to examine other forest statistics as well. Some that we did not consider include the depth of the forest, the number of leaves in the forest, and the degree of the root of a random tree in the forest.

Acknowledgments

This research was funded by NSF-DMS grant 1949884 and NSA grant H98230-20-1-0009. The author thanks Amanda Burcroff, Swapnil Garg, and Alan Peng for fruitful discussions about this research and for reading drafts of the paper, as well as Noah Kravitz, Ashwin Sah, and Fan Zhou for helpful suggestions. The author also thanks Professor Joe Gallian for suggesting the problem and running the Duluth REU in which this research was conducted and Amanda Burcroff, Colin Defant, and Yelena Mandelshtam for fostering a productive virtual research environment through their mentorship.

References

[1] K. Anders and K. Archer. Rooted forests that avoid sets of permutations. *European J. Combin.*, 77:1–16, 2018.
[2] R. Arratia. On the StanleyWilf conjecture for the number of permutations avoiding a given pattern. *Electron. J. Combin.*, 6(1):N1, 1999.
[3] R. Arratia and S. Tavar. The cycle structure of random permutations. *Ann. Probab.*, 20:1567–1591, 1992.
[4] J. Backelin, J. West, and G. Xin. Wilf-equivalence for singleton classes. *Adv. Appl. Math.*, 38:133–148, 2007.
[5] M. Bna, Permutations avoiding certain patterns: the case of length 4 and some generalizations. *Discrete Math.*, 175:55–67, 1997.
[6] T. Dwyer and S. Elizalde. Wilf equivalence relations for consecutive patterns. *Adv. Appl. Math.*, 99:134–157, 2018.
[7] S. Elizalde. Statistics on pattern-avoiding permutations, PhD thesis, Massachusetts Institute of Technology, 2004.
[8] S. Elizalde. A survey of consecutive patterns in permutations, Recent trends in combinatorics, pp. 601–618. Springer, 2016.
[9] R. Ehrenborg, S. Kitaev, P. Perry. A spectral approach to consecutive pattern-avoiding permutations. *J. Comb.*, 2:305–353, 2011.
[10] S. Elizalde and M. Noy. Consecutive patterns in permutations. *Adv. Appl. Math.*, 30(12):110–125, 2003.
[11] S. Elizalde and M. Noy. Clusters, generating functions and asymptotics for consecutive patterns in permutations. *Adv. Appl. Math.*, 49:351–374, 2012.
[12] P. Flajolet and R. Sedgewick. Analytic Combinatorics. Cambridge University Press, 2009.
[13] J. Fox. Stanley-Wilf limits are typically exponential. `arXiv:1310.8378 [math.CO]`, 2013.
[14] V. L. Goncharov. On the field of combinatory analysis. *Soviet Math. Izv., Ser. Math*, 8:3–48, 1944.
[15] V. L. Goncharov. Some facts from combinatorics. *Izvestia Akad. Nauk. SSSR, Ser. Mat*, 8:3–48, 1944.
[16] I. P. Goulden and D. M. Jackson. An inversion theorem for cluster decompositions of sequences with distinguished subsequences. *J. London Math. Soc.*, 2(20):567–576, 1979.

[17] S. Garg and A. Peng. Classical and consecutive pattern avoidance in rooted forests. arXiv:2005.08889 [math.CO], 2020.

[18] S. Garg and A. Peng. Private communication, June 2020.

[19] L. H. Harper. Stirling behaviour is asymptotically normal. *Ann. Math. Statist.*, 38:410–414, 1967.

[20] S. Hopkins and M. Weiler. Pattern avoidance in poset permutations. *Order*, 33(2):299–310, 2016.

[21] S. Kitaev. Patterns in Permutations and Words. Monographs in Theoretical Computer Science. Springer-Verlag, 2011.

[22] M. Klazar, The Füredi-Hajnal conjecture implies the Stanley-Wilf conjecture, Formal Power Series and Algebraic Combinatorics, pp. 250-255. Springer, 2000.

[23] T. Kaiser and M. Klazar. On growth rates of closed permutation classes. *Electron. J. Combin.*, 9(2):R10, 2002.

[24] D. E. Knuth. The Art of Computer Programming, Volume 1. Addison-Wesley, 1968.

[25] M. Lee and A. Sah. Constraining strong c-Wilf equivalence using cluster poset asymptotics. *Adv. Appl. Math.*, 103:43–57, 2019.

[26] A. Marcus and G. Tardos, Excluded permutation matrices and the Stanley-Wilf conjecture. *J. Combin. Theory Ser. A*, 107:153–160, 2004.

[27] B. Nakamura. Computational approaches to consecutive pattern avoidance in permutations. *Pure Math. Appl. (PU.M.A.)*, 22(2):253–268, 2011.

[28] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences. http://oeis.org/A000262.

[29] E. S. Rowland. Pattern avoidance in binary trees. *J. Combin. Theory, Ser. A*, 117(6), 2010.

[30] R. Simion and F. W. Schmidt. Restricted permutations. *European J. Combin.*, 6(4):383–406, 1985.

[31] Z. Stankova and J. West. A new class of Wilf-equivalent permutations. *J. Algebraic Combin.*, 15(3):271–290, 2002.

[32] E. Steingrmsson. Generalized permutation patternsa short survey. Permutation patterns, *London Math. Soc. Lecture Note Ser.*, 376:137–152, 2010.

Massachusetts Institute of Technology, Cambridge, MA 02139, USA
E-mail address: mren36@mit.edu