Asymptotics of Relaxed k-Ary Trees

Manosij Ghosh Dastidar 🔾
Institut für Diskrete Mathematik und Geometrie, TU Wien, Austria

Michael Wallner 🏠 🔾
Institut für Diskrete Mathematik und Geometrie, TU Wien, Austria

── **Abstract** ──────────────

A relaxed k-ary tree is an ordered directed acyclic graph with a unique source and sink in which every node has out-degree k. These objects arise in the compression of trees in which some repeated subtrees are factored and repeated appearances are replaced by pointers. We prove an asymptotic theta-result for the number of relaxed k-ary tree with n nodes for $n \to \infty$. This generalizes the previously proved binary case to arbitrary finite arity, and shows that the seldom observed phenomenon of a stretched exponential term $e^{cn^{1/3}}$ appears in all these cases. We also derive the recurrences for compacted k-ary trees in which all subtrees are unique and minimal deterministic finite automata accepting a finite language over a finite alphabet.

1 Introduction and Main Result

The enumeration of directed acyclic graphs (DAGs) is an important and timely topic in computer science [3], mathematics [2,13,16], and many related areas such as phylogenetics [15] and theoretical physics [11,12]. Several problems in this area have remained open for a long time, with bounds sometimes differing by an exponential factor. One of those problems is the enumeration of minimal deterministic finite automata (DFAs) with n states recognizing a finite language over a finite alphabet [14]. In [6] Elvey Price, Fang, and Wallner solved the corresponding asymptotic counting problem for a binary alphabet, and uncovered the remarkable phenomenon of a stretched exponential term $e^{c\,n^{1/3}}$. This term provides an explanation for the previously encountered difficulties. For example, the associated generating function cannot be algebraic, and it can only be D-finite (satisfy a linear differential equation with polynomial coefficients) if it has an irregular singularity.

This phenomenon was first observed by the above-mentioned authors in [7] in the asymptotic number of another class of DAGs: compacted binary trees of size n. These arise in the compression of XML documents [3], in the common subexpression problem in, e.g., compiling [5], and in data structures of, e.g., computer algebra software [8]. Since then, this phenomenon has also been shown in many classes of DAGs and related objects, such as phylogenetic networks [4], permutation patterns [18], and Young tableaux [1]. In this paper,

we show that the examples of compacted binary trees and minimal DFAs are just single cases of infinite families admitting a stretched exponential. Our main result is the following asymptotics of a super-class of the latter two.

▶ **Theorem 1.** *Let $k \geq 2$ be an integer. The number of relaxed k-ary trees satisfies for $n \to \infty$*

$$\Theta\left(n!^{k-1}\left(\frac{k^k}{(k-1)^{k-1}}\right)^n e^{3\left(\frac{k(k-1)}{2}\right)^{1/3}a_1 n^{1/3}} n^{\frac{2k-1}{3}}\right),$$

where $a_1 \approx -2.338$ is the largest root of the Airy function $\mathrm{Ai}(x)$ defined as the unique function satisfying $\mathrm{Ai}''(x) = x\mathrm{Ai}(x)$ and $\lim_{x\to\infty} \mathrm{Ai}(x) = 0$.
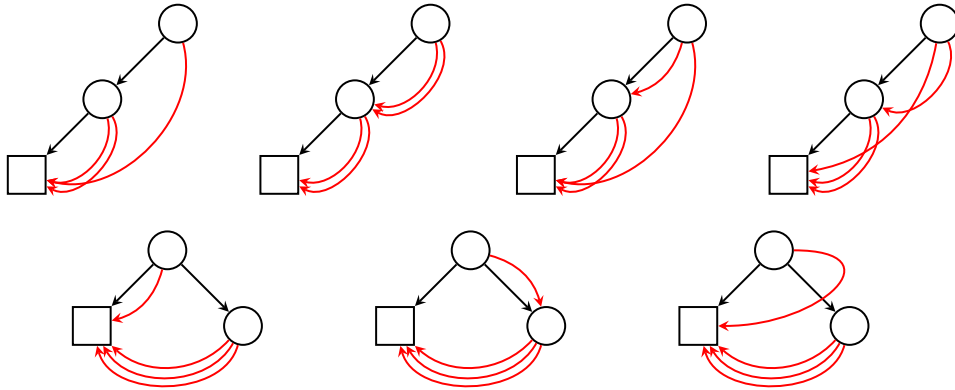
Figure 1 All 7 ternary relaxed ternary trees with 2 internal nodes.

2   Bijections and recurrences

In this paper, we consider a special class of DAGs, in which the outgoing edges (equivalently, the children) are equipped with an order.

▶ **Definition 2.** *An* ordered DAG *is a directed acyclic graph where there is a left to right ordering among the children.*

This brings us to the main object of this paper: the class of relaxed trees, a subclass of ordered DAGs, defined as follows; see Figure 1 and 2. The word relaxed signifies that they are a super-class of compacted trees, which are in bijection with trees after a compression procedure; see [7,8].

▶ **Definition 3.** *A* relaxed k-ary tree *is an ordered DAG consisting of a unique source and a unique sink such that every node except the sink has out degree exactly k. Its* spine *is the spanning subtree created by the depth first search. The edges of the spine are called* internal edges*, while the other edges are called* external edges *or* pointers*. All nodes except the unique sink, are called* internal nodes*.*

▶ Remark 4. Note that in the binary case, the previous definition is equivalent to the one given in [7], in which relaxed trees are defined as trees enriched by pointers. Whereas here we characterize relaxed trees as a subclass of ordered DAGs.
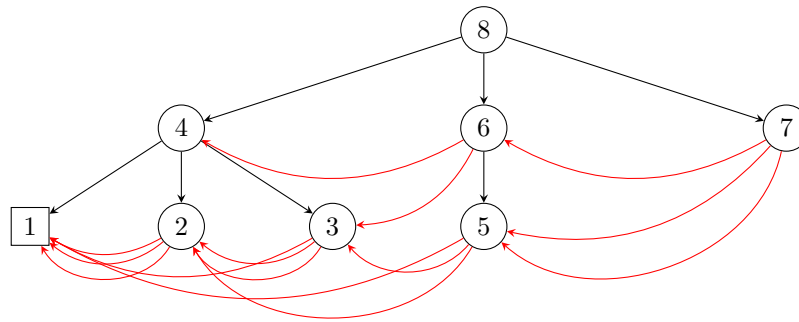
Figure 2 Example of relaxed ternary tree with 7 internal nodes (circles) labelled in postorder. The unique sink is depicted by a square. The black edges belong to the spine, the red ones are so-called pointers.

We start as in [7] by drawing a bijection between relaxed k-ary trees and Dyck paths with weights on their horizontal steps. This is done so that we can convert internal nodes into vertical steps and pointers into horizontal steps.

Every internal node u in the spine of a relaxed tree has a k-ary tree $T(u)$ associated with it where the nodes are traversed in postorder.

▶ **Definition 5.** *A* compacted k-ary tree *is a special case of a relaxed k-ary tree where for any arbitrary two nodes u, v in the spine, the k-ary trees associated with them $T(u), T(v)$ are not identical.*

In order to count relaxed k-ary trees, we will now describe a bijection to a class of paths, which are easier to enumerate. Let us first define the specific paths.

▶ **Definition 6.** *A horizontally k-decorated path P is defined as a lattice path consisting of up steps $U = (0, 1)$ and horizontal steps $H = (1, 0)$ from $(0, -1)$ with decorations such that:*
- *The first step is a U step, and its removal leaves a path never crossing the diagonal $y = \frac{x}{k-1}$.*
- *Below each H step, there is exactly one cross in one of the unit boxes below this H step and $y = -1$.*

The following lemma describes the bijection between relaxed k-ary trees and horizontally k-decorated paths. An example is shown in Figure 3.

▶ **Lemma 7.** *There is a bijection between horizontally k-decorated paths ending at $((k-1)n, n)$ and relaxed k-ary trees of size $n + 1$.*

Proof. Let R be a relaxed k-ary tree of size $n + 1$. We transform this tree uniquely into a k-decorated path ending at $((k-1)n, n)$. We traverse the spine of R in post-order and label nodes from 1 to $n + 1$. During the traversal when we move up (i.e., traverse a spine edge the second time) we add a U step and when crossing a pointer for the last time (i.e., circling around its parent) we add a H step. Thereby, we associate each U step with the node we leave, and each H step with the parent of its pointer. Moreover, below each H step we draw a cross in the unit box that intersects the column of this H step with the row of the U step that is associated with the target of the pointer of this H step.

Observe now, that the first step is always a U step, as the unique sink is processed first in post-order and has no pointers. Thus, after removing this step, we start at the origin $(0, 0)$. Furthermore, note that R consists of n spine edges, and $(k-1)n$ pointers. Hence, the
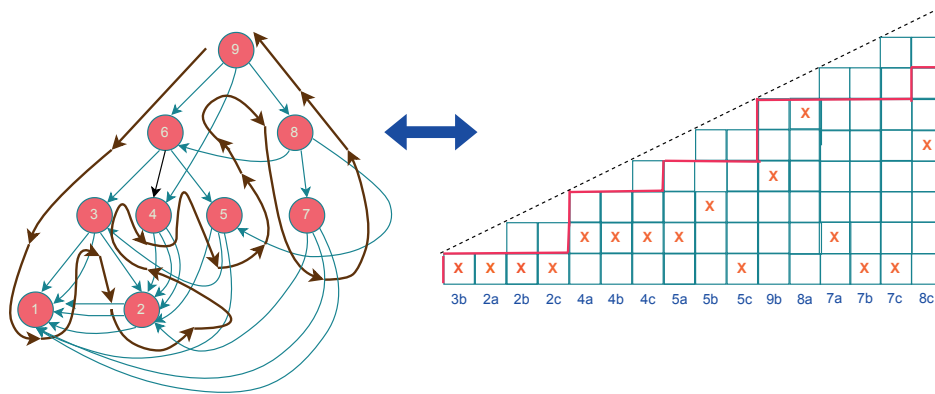
Figure 3 Example of the bijection described in Lemma 7 between relaxed ternary trees and horizontally 3-decorated paths.

path consists of n U steps and $(k-1)n$ H steps, and after attaching a final U step, the path ends at $((k-1)n, n)$. It remains to show that the path never crosses $y = \frac{x}{k-1}$. Note that before a U step is added, the relaxed subgraph of which the associated node is the root has been traversed. The subgraph consisting only of the node 1 is treated by the fact that we start at $(0, -1)$. Now we proceed by induction on the depth of the subgraph. We need to distinguish, whether the node is on the left-branch from the root or not. First, we assume it is not. The minimal cases is a single pointer, which leads to a step H and increases the distance to the diagonal $y = \frac{x}{k-1}$ by one unit. Now assume that a subgraph with $i > 0$ nodes is given. The root has k children, which all have size less than i. Thus by induction, each part does not cross $y = \frac{x}{k-1}$ and moves one unit further to the right from the diagonal. Therefore, processing the k children moves the path k steps to the right from $\frac{x}{k-1}$, while (after that) processing the root, moves the path one unit up. Thus the path stays below $y = \frac{x}{k-1}$ and is one unit further to the right. Second, we assume that the node is on the diagonal. Then, the first child does not move the path one unit to the right, but the other $k-1$ do. By induction, the path still does not cross the diagonal $\frac{x}{k-1}$ but the distance now also stays the same.

Now in the reverse direction, let us take a horizontally k-decorated path ending at $((k-1)n, n)$. We want to recover a relaxed k-ary tree from this path. We start by noting that the number of up steps in the path is n and therefore the number of nodes is $n + 1$. Our first step is a horizontal step.

We start by adding a H step which corresponds to the left most leaf labelled 1. Then along with the path P, we create the spine of the the relaxed k-ary tree. Thus along the up steps we create nodes of the spine as the ith up step creates the $(i+1)$th node in post-order. Along the horizontal steps we say that we add a pointer from the existing node that we are in to a node $j + 1$ which is indicated by the cross placed on the level j below the path (counting from the bottom up). Completing this process we get a relaxed k-ary tree with its nodes labelled in post-order.   ◀

This bijection allows directly to derive the following recurrence relations, following the step-by-step construction of the paths. In Table 1 we give the initial terms for $k = 2, \ldots, 5$.

▶ **Proposition 8.** *Let $r_{n,m}$ be the number of horizontally decorated paths ending at (n,m) and $\forall k \in \mathbb{Z}$, $k > 1$.*

$$r_{n,m} = r_{n,m-1} + (m+1)r_{n-1,m}, \qquad\qquad for\ 1 \le m \le \frac{n}{k-1},$$ (1)

$$r_{n,m} = 0 \qquad\qquad for\ m > \frac{n}{k-1},$$

$$r_{n,0} = 1 \qquad\qquad for\ n \ge 0.$$

Thus, by Lemma 7, the number of relaxed k-ary trees with n internal nodes is equal to $r_{(k-1)n,n}$.

■ **Table 1** Number of relaxed k-ary trees. Also the number of unlabeled acyclic single-source automata with n transient states on a k-letter input alphabet. The matrix consisting of these rows is given by A128249.

k	Relaxed k-ary trees (r_k)	OEIS
2	$(1, 3, 16, 127, 1363, 18628, 311250, 6173791, \ldots)$	A082161
3	$(1, 1, 7, 139, 5711, 408354, 45605881, 7390305396, \ldots)$	A082162
4	$(1, 1, 15, 1000, 189035, 79278446, 63263422646, 86493299281972, \ldots)$	A102102
5	$(1, 1, 31, 6631, 5470431, 12703473581, 68149976969707, \ldots)$	—

In a next step, we will also state the respective recurrences for compacted k-ary trees and minimal DFAs accepting a finite language on an alphabet of size k. These results generalize [7, Proposition 2.11] and [6, Proposition 5], respectively.

In a relaxed k-ary tree an internal node will be called a *cherry* if all its children are pointers.

▶ **Theorem 9.** *A relaxed k-ary tree C is a compacted k-ary tree if and only if no internal nodes u, v have the same children in the same order. Moreover, if C is not a compacted k-ary tree there exists a pair (u, v) with identical children and v is a cherry.*

Proof. Let us take C to be a compacted k-ary tree. Let two internal nodes u, v have the same children in the same order. This implies that the k-ary trees associated with u and v are isomorphic, which violates the definition of a compacted tree.

Conversely, let us assume that C is not a compacted tree. Therefore from the definition of a compacted tree it follows that we can find at least one pair of internal nodes (u, v) such that the k-ary trees associated with them are the same. At this point if v is a cherry then we are done. But if v is not a cherry then we take the first child of both nodes, say u_1 and v_1, respectively, and note that $T(u_1) = T(v_1)$. Now, if v_1 is not a cherry then we can continue this process. Therefore by infinite descent we see that there has to be two nodes u, v such that v is a cherry. ◀

In order to continue, we need the following concept. First, recall that in each horizontally k-decorated path, each step H is decorated by a marked box below. Further, each up step U corresponds to an internal node in the k-ary tree. Thus, we assign to each step U a k-tuple $(u_1, u_2, \ldots, u_k)$ corresponding to the nodes of its children. (This will help us to construct the compacted trees from the relaxed ones.) Let us define a C_k-*decorated path* as a horizontally decorated path with the restriction that for consecutive steps H^kU (k steps H followed by U) the k-tuple $(h_1, h_2, \ldots h_k)$ corresponding to the labels below the H steps, we have $(h_1, h_2, \ldots h_k) \neq (u_1, u_2, \ldots u_k)$ for all preceding steps U.

▶ **Proposition 10.** *There is a bijection between the number of compacted k-ary trees of size n and C_k-decorated paths of length kn.*

Proof. Note that the C_k-decorated paths are a subset of the horizontally k-decorated paths and the previous bijection sends the relaxed k-ary trees to horizontally k-decorated paths. The C_k-decorated paths have been constructed in this way to reinforce the condition that the relaxed trees corresponding to these decorated paths are such that the k-ary trees associated with any two arbitrary nodes can never be identical. ◀

This allows us, again by a direct step-by-step construction, to derive the following bivariate recurrence for C_k-decorated paths; see Table 2.

▶ **Proposition 11.** *Let $c_{n,m}$ denote the number of C_k-decorated paths ending at (n,m). Then*

$$c_{n,m} = c_{n,m-1} + (m+1)c_{n-1,m} - (m-1)c_{n-k,m-1} \qquad \text{for } 1 \leq m \leq \frac{n}{k-1},$$

$$c_{n,m} = 0 \qquad \text{for } m > \frac{n}{k-1},$$

$$c_{n,0} = 1 \qquad \text{for } n \geq 0.$$

The number of compacted k-ary trees with n internal nodes is $c_{(k-1)n,n}$.

■ **Table 2** Number of compacted k-ary trees, which are defined as relaxed k-ary trees with the additional constraint that each fringe subtree is unique.

k	Compacted k-ary trees (c_k)	OEIS
2	$(1, 1, 3, 15, 111, 1119, 14487, 230943, 4395855, \ldots)$	A254789
3	$(1, 1, 7, 133, 5299, 371329, 40898599, 6561293893, \ldots)$	—
4	$(1, 1, 15, 975, 182175, 75961695, 60422966655, 82450320955455, \ldots)$	—
5	$(1, 1, 31, 6541, 5373571, 12458850121, 66790559866471, \ldots)$	—

We have a similar recursion for the minimal DFAs over a k-ary alphabet. We use the same scheme for the bijection with the H steps having a decoration. Since we have accepting and rejecting states, the U steps are colored either red or green. We call such paths 2-coloured. Again, we summarize the first values of Table 3.

▶ **Proposition 12.** *Let $b_{n,m}$ denote the number of 2-coloured horizontally decorated paths corresponding to DFAs ending in (n,m). Then*

$$b_{n,m} = 2b_{n,m-1} + (m+1)b_{n-1,m} - m\, b_{n-k,m-1} \qquad \text{for } 1 \leq m \leq \frac{n}{k-1},$$

$$b_{n,m} = 0 \qquad \text{for } m > \frac{n}{k-1},$$

$$b_{n,0} = 1 \qquad \text{for } n \geq 0.$$

The number of minimal DFAs accepting a accepting a finite language on an alphabet of size k with $n+1$ states is $b_{(k-1)n,n}$.

■ **Table 3** Number of minimal deterministic finite automata recognizing a finite k-letter language.

k	DFA (m_k)	OEIS
2	$(1, 1, 6, 60, 900, 18480, 487560, \ldots)$	A331120
3	$(1, 1, 14, 532, 42644, 6011320, 1330452032, \ldots)$	—
4	$(1, 1, 30, 3900, 1460700, 1220162880, \ldots)$	—
5	$(1, 1, 62, 26164, 43023908, 199596500056, \ldots)$	—

3 Asymptotics of relaxed k-ary trees

The goal of this section is to prove our main Theorem 1 on relaxed k-ary trees. As the proof is rather technical and complex, we will begin with an overview of the main steps.

First, in Section 3.1 we transform the recurrence (1) into a new recurrence $(d_{i,j})_{i,j \geq 0}$ by changing the used steps and extracting the super-exponential and certain exponential and polynomial contributions. Second, in Section 3.2 we perform a heuristic analysis of the asymptotics of $(d_{i,j})_{i,j \geq 0}$, and guess the shape for a rigorous proof. Using this shape, we build and prove in Section 3.3 explicit sequences $(\tilde{X}_{i,j})_{i,j \geq 0}$ and $(\hat{X}_{i,j})_{i,j \geq 0}$ that satisfy the same recurrence as $(d_{i,j})_{i,j \geq 0}$ but with the equality sign replaced by the inequality signs $\leq$ and $\geq$, respectively. This is the most technical part in which we heavily rely on computer algebra. Finally, in Section 3.4 we use these explicit sequences to prove inductively asymptotically tight upper and lower bounds, which yield our main theorem.

3.1 Transformation into a Dyck-like recurrence

We start from recurrence (1) and our goal is to determine the asymptotics of $r_{(k-1)n,n}$. First, we observe that the path closest to the diagonal given by $(H^{k-1}V)^n$ has a weight $(n!)^{k-1}$ as there are n H steps at height $0, 1, \ldots, n-1$. All other paths will get smaller weights. Therefore, after rescaling with this weight, all paths have a weight bounded by one. Note that it is technically easier to work with the rescaling $((k-1)n)!$, which has by Stirling's formula the same super-exponential growth. The difference in the exponential growth is a factor $(k-1)^{(k-1)n}$, which we will take into account later in (5). Thus, we set $\tilde{r}_{n,m} = \frac{r_{n,m}}{((k-1)n)!}$ and get the new recurrence

$$\tilde{r}_{n,m} = \frac{m+1}{n} \tilde{r}_{n-1,m} + \tilde{r}_{n,m-1}, \qquad\qquad m \leq \frac{n}{k-1}.$$

Next, as we are interested in the asymptotics of $\tilde{r}_{(k-1)n,n}$, we will transform this recurrence from North and East steps, to Dyck-like up and down steps. For this purpose, we define the new variables

$$\begin{cases} i = n + m, \\ j = n - (k-1)m, \end{cases} \qquad \text{equivalently} \qquad \begin{cases} n = \frac{(k-1)i+j}{k}, \\ m = \frac{i-j}{k}. \end{cases}$$

The idea behind this choice, is that i tracks the length of the walk, i.e., the number of steps, and j the distance to the diagonal. By this choice, we have $i, j \geq 0$. This gives the following generalized Dyck-like recurrence:

$$\tilde{d}_{i,j} = \tilde{U}(i,j)\tilde{d}_{i-1,j-1} + \tilde{d}_{i-1,j+k-1}, \qquad\qquad i > 0, j \geq 0, \qquad\qquad (2)$$

with the initial condition $\tilde{d}_{0,0} = 1$ and the weight $\tilde{U}(i,j) = \frac{i-j+k}{(k-1)i+j}$. Thus, we have the new steps $(1,1)$ and $(1,-k+1)$. This is a simple directed lattice path model with space-dependent

weights. Note that as the change in the x-direction is one unit per step, a path of length n consists of n steps, and it suffices therefore to track the current altitude. Therefore, from now on we consider only the changes in the y-direction, which we call jumps $+1$ and $-(k-1)$.

Let us now look at the drift of this new model. For now, we assume that i is large and that $j = o(i)$. The *drift* $\tilde{\delta}(i,j)$ at a point (i,j) is defined as the expected next jump size when leaving (i,j). Therefore, at (i,j) the jump $+1$ gets weight $\tilde{U}(i+1,j+1)$ and the jump $-(k-1)$ weight 1 for $j \geq k-1$ and 0 otherwise. Hence, we get that $\tilde{\delta}(i,j) = 1$ for $0 \leq j \leq k-2$ and

$$\tilde{\delta}(i,j) = \frac{i-j+k}{(k-1)(i+1)+j+1} - (k-1) = -\frac{k(k-2)}{k-1} + \mathcal{O}\left(\frac{1}{i}\right), \qquad \text{for } j > k-2.$$

Note that in the binary case ($k=2$), the first term is zero and therefore the drift is converging to zero for large i. However, in the general case the drift is negative for large i. But we can define the following transformation to achieve the same behavior: $d_{i,j} = (k-1)^{2n}\tilde{d}_{i,j}$ where $n = \frac{(k-1)i+j}{k}$. This gives the final Dyck-like recurrence

$$d_{i,j} = U(i,j)d_{i-1,j-1} + d_{i-1,j+k-1}, \qquad\qquad i > 0, j \geq 0, \qquad\qquad (3)$$

with the initial condition $d_{0,0} = 1$ and the following weight for the up step

$$U(i,j) = \frac{(k-1)^2(i-j+k)}{(k-1)i+j} = (k-1)\left(1 - \frac{k(j-k+1)}{(k-1)i+j}\right). \qquad\qquad (4)$$

The drift $\delta(i,j)$ in this model is again $\delta(i,j) = 1$ for $0 \leq j \leq k-2$ but now we have

$$\delta(i,j) = -\frac{k(k-1)(j-k+2)}{k(i+1)-i+j} = -\frac{k(j-k+2)}{i} + \mathcal{O}\left(\frac{1}{i^2}\right), \qquad \text{for } j > k-2,$$

and therefore converges to zero. In this final model (3), we are interested in $d_{kn,0}$ since

$$r_{(k-1)n,n} = ((k-1)n)!\tilde{r}_{(k-1)n,n} = ((k-1)n)!\tilde{d}_{kn,0} = \frac{((k-1)n)!}{(k-1)^{2(k-1)n}}d_{kn,0}$$

$$\sim \sqrt{k-1}(2\pi)^{1-k/2}\frac{(n!)^{k-1}}{(k-1)^{(k-1)n}}n^{1-k/2}d_{kn,0}, \qquad\qquad (5)$$

where the last equality follows by Stirling's formula.

3.2   Heuristic analysis

In the next step, we will now heuristically analyze the recurrence (3). Inspired by the binary case and numerical experiments, we use the ansatz

$$d_{i,j} = h(i)f\left(\frac{j+1}{i^{1/3}}\right). \qquad\qquad (6)$$

The function $h(i)$ captures a macroscopic amplitude that is independent of j, while the function $f(x)$ captures the local behavior around the origin. The rescaling by $i^{1/3}$ is motivated by an analogy to pushed Dyck path and the binary case [7].

Substituting this ansatz into the recurrence (3) and reordering we get

$$\frac{h(i)}{h(i-1)} = \frac{U(i,j)f\left(\frac{j}{(i-1)^{1/3}}\right) + f\left(\frac{j+k}{(i-1)^{1/3}}\right)}{f\left(\frac{j+1}{i^{1/3}}\right)}.$$

As the left-hand side is independent of j, for large i both sides should have an expansion in i with whose coefficients are independent of j. Let us now zoom into the region $i^{1/3}$ by binding the variables i and j using the transformation $j = x\, i^{1/3} + 1$. This gives

$$\frac{h(i)}{h(i-1)} = \frac{U(i, x\, i^{1/3} + 1) f\left(\frac{x\, i^{1/3}+1}{(i-1)^{1/3}}\right) + f\left(\frac{x\, i^{1/3}+k+1}{(i-1)^{1/3}}\right)}{f(x)}. \tag{7}$$

Assuming now that $f(x)$ has a convergent Taylor series expansion around x, we get

$$\frac{h(i)}{h(i-1)} = k + \frac{k}{2}\frac{(k-1)f''(x) - 2xf(x)}{f(x)} i^{-2/3} + \mathcal{O}(i^{-1}).$$

Hence, in order to be consistent, the left-hand side needs to have an expansion in decreasing powers of $i^{1/3}$. In particular, it starts like

$$\frac{h(i)}{h(i-1)} = k + c\, i^{-2/3} + \mathcal{O}(i^{-1}),$$

for a constant $c \in \mathbb{R}$. Comparing coefficients, we see that $f(x)$ has to satisfy the differential equation

$$f''(x) = \frac{2(kx+c)}{k(k-1)} f(x).$$

This differential equation is, as in the binary case, solved by the Airy functions of the first and second kind. Due to the combinatorial nature of the problem, we require $f(x) \geq 0$ for $x \geq 0$ as well as $\lim_{x \to \infty} f(x) = 0$. Therefore, we get up to a multiplicative constant the following solution

$$f(x) = \mathrm{Ai}\left(\frac{2^{1/3}(kx+c)}{k(k-1)^{1/3}}\right).$$

Now, note that due the boundary conditions $d_{n,-1} = 0$ we must have $f(0) = 0$. Together, with $f(x) \geq 0$ for $x \geq 0$, this means that the argument of the Airy function must evaluate to the largest zero $a_1 \approx -2.338$ of $\mathrm{Ai}(x)$. Therefore

$$c = \frac{k(k-1)^{1/3}}{2^{1/3}} a_1$$

and

$$f(x) = \mathrm{Ai}\left(a_1 + B\, x\right), \qquad \text{where} \qquad B = \left(\frac{2}{k-1}\right)^{1/3}.$$

In order to capture the polynomial term, we need to use a more general ansatz than (6) that includes a second function $g(x)$ capturing the scale of order $i^{-1/3}$. In particular, we use

$$d_{i,j} = h(i)\left(f\left(\frac{j+1}{i^{1/3}}\right) + \frac{g\left(\frac{j+1}{i^{1/3}}\right)}{i^{1/3}}\right). \tag{8}$$

Note that the function $g(x)$ will influence the terms starting from order i^{-1} in (7). Analogous computations as performed above for $f(x)$ lead us to the ansatz

$$g(x) = -\left(\frac{(k+2)x^2}{6(k-1)} + \frac{2^{2/3}a_1(k-2)x}{6(k-1)^{2/3}}\right) f(x).$$

3.3   Explicit bounds

All these heuristic arguments above guide us to the following two results. These generalize [7, Lemmas 4.2 and 4.4], whose results are recovered by setting $k = 2$. The proofs are analogous to [4, 7]; for the details we refer to the accompanying Maple worksheet. In particular, this shows that the method developed in [7] is powerful enough to analyze bivariate recurrences that include parameters.

▶ **Lemma 13.** *Let $k \geq 2$ be an integer and $B = (\frac{2}{k-1})^{1/3}$. For all $i, j \geq 0$ let*

$$
\begin{aligned}
\tilde{X}_{i,j} := &\left(1 - \frac{2^{2/3} a_1(k-2)}{6(k-1)^{2/3}} \frac{j}{i^{2/3}} - \frac{k+2}{6(k-1)} \frac{j^2}{i} + \frac{7k-11}{6(k-1)} \frac{j}{i} \right. \\
&\left. + \frac{a_1^2(k-2)^2 B^4}{72} \frac{j^2}{i^{4/3}} + \frac{a_1(k^2-4)^2 B^5}{72} \frac{j^3}{i^{5/3}} \right) \mathrm{Ai}\left(a_1 + \frac{B(j+1)}{i^{1/3}} \right), \\
\tilde{s}_i := &\, k\left(1 + \frac{a_1}{Bi^{2/3}} + \frac{7k-6}{6i} - \frac{1}{i^{7/6}} \right).
\end{aligned}
$$

Then, for any $\varepsilon > 0$, there exists an $\tilde{i}_0$ such that

$$
\tilde{X}_{i,j}\tilde{s}_i \leq U(i,j)\tilde{X}_{i-1,j+1} + \tilde{X}_{i-1,j-1}
$$

for all $i \geq \tilde{i}_0$ and for all $0 \leq j < i^{2/3-\varepsilon}$, where $U(i,j)$ is defined in (4).

▶ **Lemma 14.** *Let $k \geq 2$ be an integer and $B = (\frac{2}{k-1})^{1/3}$. Choose $\eta > \frac{(k+2)^2}{72(k-1)^2}$ fixed and for all $i, j \geq 0$ let*

$$
\begin{aligned}
\hat{X}_{i,j} := &\left(1 - \frac{2^{2/3} a_1(k-2)}{6(k-1)^{2/3}} \frac{j}{i^{2/3}} - \frac{k+2}{6(k-1)} \frac{j^2}{i} + \frac{7k-11}{6(k-1)} \frac{j}{i} \right. \\
&\left. + \frac{a_1^2(k-2)^2 B^4}{72} \frac{j^2}{i^{4/3}} + \frac{a_1(k^2-4)^2 B^5}{72} \frac{j^3}{i^{5/3}} + \eta\frac{m^4}{n^2} \right) \mathrm{Ai}\left(a_1 + \frac{B(j+1)}{i^{1/3}} \right), \\
\hat{s}_i := &\, k\left(1 + \frac{a_1}{Bi^{2/3}} + \frac{7k-6}{6i} + \frac{1}{i^{7/6}} \right).
\end{aligned}
$$

Then, for any $\varepsilon > 0$, there exists an $\hat{i}_0$ such that

$$
\hat{X}_{i,j}\hat{s}_i \geq U(i,j)\hat{X}_{i-1,j+1} + \hat{X}_{i-1,j-1}
$$

for all $i \geq \hat{i}_0$ and for all $0 \leq j < i^{1-\varepsilon}$, where $U(i,j)$ is defined in (4).

Note that in the binary case $k = 2$, many terms in the previous Lemmas 13 and 14 are zero. These terms do not affect the final asymptotics, but they are needed for the proof using generalized Newton polygons. In particular, they allow to set certain points on the convex hull to zero, which then leads to the same behavior as in the binary case. For this reason, the technical proofs follow nearly verbatim the binary ones and we omit them in this extended abstract and refer to [7] and the accompanying Maple worksheet.

These explicit recurrences that satisfy the recurrence (3) with = replaced by $\leq$ and $\geq$, are the key ingredient to prove our main theorem.

3.4 Proof of Theorem 1 on relaxed k-ary trees

We start with the lower bound. First, we define a sequence $X_{i,j} := \max\{\tilde{X}_{i,j}, 0\}$ which satisfies the inequality of Lemma 13 for *all* $j \leq \frac{i}{k-1}$. Note that the factor of the Airy function Ai becomes negative for large j. Then, we define an explicit sequence $\tilde{h}_i := \tilde{s}_i \tilde{h}_{i-1}$ for $i > 0$ and $\tilde{h}_0 = \tilde{s}_0$. Using this we can prove by induction that $d_{i,j} \geq C_0 \tilde{h}_i X_{i,j}$ for some constant $C_0 > 0$ and all $i \geq \tilde{i}_0$ and all $0 \leq j \leq \frac{i}{k-1}$. Therefore,

$$d_{kn,0} \geq C_0 \tilde{h}_{kn} X_{kn,0}$$

$$\geq C_0 \prod_{i=1}^{kn} k \left(1 + \frac{a_1}{Bi^{2/3}} + \frac{7k-6}{6i} - \frac{1}{i^{7/6}}\right) \text{Ai}\left(a_1 + \frac{B}{(kn)^{1/3}}\right)$$

$$\geq C_1 k^{kn} e^{3a_1(kn)^{1/3}/B} n^{\frac{7k-8}{6}}. \tag{9}$$

Finally, combining this with (5) we get the lower bound.

We continue with the upper bound, whose proof is similar, yet more technical. The starting point is Proposition 14 and, as in the lower bound, a function $X_{i,j}$ that is valid for *all* $0 \leq j \leq \frac{i}{k-1}$. For this purpose we define as in the binary case [7] a sequence $\tilde{d}_{i,j}$ depending on some large integer parameter $I > 0$ such that

$$\tilde{d}_{i,j} := \begin{cases} d_{i,j} & \text{for } 0 \leq j \leq i^{3/4} \text{ and } i > I, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

The missing key step is now to show that $d_{kn,0} = \mathcal{O}(\tilde{d}_{kn,0})$. Combining this with the analogous computations performed for the lower bound above we get

$$\tilde{d}_{kn,0} \leq \hat{C}_1 k^{kn} e^{3a_1(kn)^{1/3}/B} n^{\frac{7k-8}{6}}.$$

To complete the prove we show $d_{kn,0} \leq 2\tilde{d}_{kn,0}$ using lattice path theory and computer algebra. We start from Equation (3) of $d_{i,j}$, which we interpret as a recurrence counting lattice paths. They are composed of steps $(1,1)$ weighted by $U(i,j)$ and $(1,-k+1)$ weighted by 1 when the respective step ends at (i,j). The total weight of a path is the product of its weights. Now, we are interested in the paths never crossing $y = 0$ and ending at $(kn, 0)$. Let now $p_{\ell,k,kn}$ be the number of such paths starting at (r,s) and ending at $(kn, 0)$. From (3) we directly get

$$p_{r,s,kn} = \frac{(k-1)^2(r-s+k)}{(k-1)(r+1)+s+1} p_{r+1,s+1,kn} + p_{r+1,s-k+1,kn},$$

with $p_{\ell,-1,kn} = 0$ and $p_{kn,s,kn} = \delta_{s,0}$, where the factor is $U(r+1, s+1)$.

Now, we are able to show that

$$\frac{p_{r,s_1,kn}}{s_1 + 1} \geq \frac{p_{r,s_2,kn}}{s_2 + 1}, \tag{11}$$

for integers $0 \leq s_1 < s_2 \leq r \leq kn$ such that $k \mid s_2 - s_1$. The proof is completely elementary and uses reverse induction on r starting from $r = kn$. We refer to our accompanying Maple worksheet.

Finally, from (11) we directly get

$$p_{kx,ky,kn} \leq (ky+1) p_{kx,0,kn}. \tag{12}$$

This allows to prove the following generalization of [7, Lemma 4.6] whose proof follows exactly the same lines and we omit it here. For the following statement, recall that $d_{i,j}$ is the weighted number of paths ending at (i,j). Let $\tilde{d}_{i,j}$ be the number of these paths such that no intermediate point (kx, ky) on the path satisfies $x > I_\varepsilon$ and $y > x^{3/4}$.

▶ **Lemma 15.** *For all $\varepsilon > 0$ there exists a constant $I_\varepsilon > 0$ that acts as the parameter I in the definition of $\tilde{d}_{i,j}$ in (10), such that $d_{kn,0} \leq (1 + \varepsilon)\tilde{d}_{kn,0}$ for all $n > 0$.*

Proof. As in the binary case, we first rewrite the claimed inequality into $1 - \frac{\tilde{d}_{kn,0}}{d_{kn,0}} \leq \frac{\varepsilon}{1-\varepsilon}$. The left-hand side here represents the proportion of walks that pass through at least one point (kx, ky) such that $x > I_\varepsilon$ and $y > x^{3/4}$. Let $s_{x,y,n}$ be the proportion of such walks that pass through one fixed such point:

$$s_{x,y,n} = \frac{d_{kx,ky}p_{kx,ky,kn}}{d_{kn,0}}.$$

The idea is that the sum of these values over all violating points x, y is of course an upper bound for our claim. So we want to prove that this sum is very small for large x.

By (12) combined with the fact that $s_{x,y,n} \leq 1$, we get

$$s_{x,y,n} \leq \frac{(ky+1)p_{kx,0,kn}d_{kx,ky}}{d_{kn,0}s_{x,0,n}} = (ky+1)\frac{d_{kx,ky}}{d_{kx,0}}$$

$$\leq \frac{ky+1}{C_1 k^{kx} e^{3a_1(kx)^{1/3}/B} x^{\frac{7k-8}{6}}} (k-1)^{(k-1)x+y}\binom{kx}{x-y}$$

where we also used the lower bound (9) for $d_{kx,0}$ and the crude bound $d_{kx,ky} \leq (k-1)^{(k-1)x+y}\binom{kx}{x-y}$. The latter holds, as we may bound non-negative paths ending at (kx, ky) by unconstrained paths with weights $(k-1)$ for the up step $(1, 1)$ and 1 for down step $(1, -k+1)$, since $U(i, j) \leq k - 1$ for all $i, j \geq 0$.

Now this last expression is completely explicit, and for large x one can see that it is of order $\Theta(e^{-x^{1/2}})$. Now, the proof of the binary case follows verbatim and the claim follows. ◀

Finally, this proves $d_{kn,0} \leq 2\tilde{d}_{kn,0}$ and ends the proof of our main Theorem 1.

4 Conclusion and Outlook

The aim of this paper was to show that the method from [7] developed for the asymptotics of compacted binary trees can be applied to more general recurrences. Previously, in [4] we showed how to handle more general weights including an integer parameter, while in this paper we generalized the used steps. Instead of steps $(1, 1)$ and $(1, -1)$ of Dyck type we studied a recurrence with larger steps given by $(1, k - 1)$ and $(1, -1)$, for an arbitrary fixed positive integer k. In particular, we proved in Theorem 1 that in the class of relaxed k-ary trees the phenomenon of a stretched exponential appears for any integer $k \geq 2$.

In the long version, we will also prove the generalizations from the binary to the k-ary case for compacted trees and minimal deterministic finite automata accepting a finite language on an alphabet of size k for which we showed in Propositions 11 and 12 that their respective recurrences have similar shapes. We expect that the phenomenon of a stretched exponential persists also in these cases. In general, the presented method can be applied to more general cases including several parameters in the weights and also recurrences with more than two steps.

Another research direction is to generalize the previously studied relaxed binary trees of bounded right height [9] to the k-ary case. In the binary case, this class has been proved to be D-finite using exponential generating function methods. In the k-ary case, a similar rescaling, as we used in Section 3.1 by $n!^{k-1}$ should allow to analyze this class using generalized exponential functions of the type $\sum_{n\geq 0} a_n \frac{z^n}{(n!)^{k-1}}$. This class is also interesting, since the binary case has rich combinatorial properties, such as bijection and closed-form enumeration formulas [17], which are also worth to investigate in the k-ary case.

──── **References** ────

1 Cyril Banderier and Michael Wallner. Young Tableaux with Periodic Walls: Counting with the Density Method. *Sém. Lothar. Combin.*, 85B, Article 85B.47:12 pp., 2021. URL: `https://www.mat.univie.ac.at/~slc/wpapers/FPSAC2021/47.html`.

2 Béla Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001. `doi:10.1017/CBO9780511814068`.

3 Mireille Bousquet-Mélou, Markus Lohrey, Sebastian Maneth, and Eric Noeth. XML compression via directed acyclic graphs. *Theory Comput. Syst.*, 57(4):1322–1371, 2015. `doi:10.1007/s00224-014-9544-x`.

4 Yu-Sheng Chang, Michael Fuchs, Hexuan Liu, Michael Wallner, and Guan-Ru Yu. Enumerative and distributional results for *d*-combining tree-child networks. *Adv. in Appl. Math.*, 157:Paper No. 102704, 2024. `doi:10.1016/j.aam.2024.102704`.

5 Peter J. Downey, Ravi Sethi, and Robert Endre Tarjan. Variations on the common subexpression problem. *J. Assoc. Comput. Mach.*, 27(4):758–771, 1980. `doi:10.1145/322217.322228`.

6 Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Asymptotics of Minimal Deterministic Finite Automata Recognizing a Finite Binary Language. In *AofA 2020*, volume 159 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 11:1–11:13, 2020. `doi:10.4230/LIPIcs.AofA.2020.11`.

7 Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Compacted binary trees admit a stretched exponential. *J. Combin. Theory Ser. A*, 177:105306, 40, 2021. `doi:10.1016/j.jcta.2020.105306`.

8 Philippe Flajolet, Paolo Sipala, and Jean-Marc Steyaert. Analytic variations on the common subexpression problem. In *Automata, languages and programming*, pages 220–234. Springer, 1990. `doi:10.1007/BFb0032034`.

9 Antoine Genitrini, Bernhard Gittenberger, Manuel Kauers, and Michael Wallner. Asymptotic enumeration of compacted binary trees of bounded right height. *J. Combin. Theory Ser. A*, 172:105177, 2020. `doi:10.1016/j.jcta.2019.105177`.

10 Manosij Ghosh Dastidar and Michael Wallner. Maple worksheet - asymptotics of relaxed k-ary trees. Software, FWF P 34142, swhId: `swh:1:dir:f9b3ed0c7c098f24ee8860808959b84e402dc276` (visited on 2024-07-09). URL: `https://gitlab.tuwien.ac.at/michael.wallner/relaxed-k-ary-trees`.

11 A. J. Guttmann and S. G. Whittington. Two-dimensional lattice embeddings of connected graphs of cyclomatic index two. *J. Phys. A*, 11(4):721–729, 1978. URL: `http://stacks.iop.org/0305-4470/11/721`.

12 Anthony J. Guttmann. Analysis of series expansions for non-algebraic singularities. *Journal of Physics A: Mathematical and Theoretical*, 48(4):045209, 2015. `doi:10.1088/1751-8113/48/4/045209`.

13 Svante Janson, Tomasz Łuczak, and Andrzej Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000. `doi:10.1002/9781118032718`.

14 Valery A. Liskovets. Exact enumeration of acyclic deterministic automata. *Discrete Applied Mathematics*, 154(3):537–551, 2006. `doi:10.1016/j.dam.2005.06.009`.

15 Colin McDiarmid, Charles Semple, and Dominic Welsh. Counting phylogenetic networks. *Ann. Comb.*, 19(1):205–224, 2015. `doi:10.1007/s00026-015-0260-2`.

16 Remco van der Hofstad. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017. `doi:10.1017/9781316779422`.

17 Michael Wallner. A bijection of plane increasing trees with relaxed binary trees of right height at most one. *Theoretical Computer Science*, 755:1–12, 2019. `doi:10.1016/j.tcs.2018.06.053`.

18 Michael Wallner. Dyck paths and inversion tables. *Permutation Patterns 2023*, pages 142–144, 2023. URL: `https://2023.permutationpatterns.com/booklet/bookletco.pdf#page=142`.