Asymptotics of Minimal Deterministic Finite Automata Recognizing a Finite Binary Language

Andrew Elvey Price

Université de Bordeaux, Laboratoire Bordelais de Recherche en Informatique, UMR 5800, 351 Cours de la Libération, 33405 Talence Cedex, France
Université de Tours, Institut Denis Poisson, UMR 7013, Parc de Grandmont, 37200 Tours, France
https://www.idpoisson.fr/elveyprice/en/
andrew.elvey@univ-tours.fr

Wenjie Fang

Laboratoire d'Informatique Gaspard-Monge, UMR 8049, Université Gustave Eiffel, CNRS, ESIEE Paris, 77454 Marne-la-Vallée, France
http://igm.univ-mlv.fr/~wfang/
wenjie.fang@u-pem.fr

Michael Wallner

Université de Bordeaux, Laboratoire Bordelais de Recherche en Informatique, UMR 5800, 351 Cours de la Libération, 33405 Talence Cedex, France
TU Wien, Institute for Discrete Mathematics and Geometry, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria
https://dmg.tuwien.ac.at/mwallner/
michael.wallner@tuwien.ac.at

Abstract

We show that the number of minimal deterministic finite automata with $n+1$ states recognizing a finite binary language grows asymptotically for $n \to \infty$ like

$$\Theta\left(n!\, 8^n e^{3a_1 n^{1/3}} n^{7/8}\right),$$

where $a_1 \approx -2.338$ is the largest root of the Airy function. For this purpose, we use a new asymptotic enumeration method proposed by the same authors in a recent preprint (2019). We first derive a new two-parameter recurrence relation for the number of such automata up to a given size. Using this result, we prove by induction tight bounds that are sufficiently accurate for large n to determine the asymptotic form using adapted Netwon polygons.

2012 ACM Subject Classification Theory of computation → Regular languages; Mathematics of computing → Enumeration; Mathematics of computing → Generating functions

Keywords and phrases Airy function, asymptotics, directed acyclic graphs, Dyck paths, bijection, stretched exponential, compacted trees, minimal automata, finite languages

Digital Object Identifier 10.4230/LIPIcs.AofA.2020.11

Related Version The companion paper "Compacted binary trees admit a stretched exponential" by the same authors is available at arXiv:1908.11181.

Funding *Andrew Elvey Price*: European Research Council (ERC) in the European Union's Horizon 2020 research and innovation programme, under the Grant Agreement No. 759702.
Wenjie Fang: *Recherche en esprit libre, i.e.*, not supported by fundings with precise predefined goals.
Michael Wallner: Erwin Schrödinger Fellowship of the Austrian Science Fund (FWF): J 4162-N35.

Acknowledgements We would like to thank Cyril Banderier, Tony Guttmann, and Andrea Sportiello for interesting discussions on the presence of a stretched exponential. We also thank our referees for their careful reading and helpful comments.

31st International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2020).
Editors: Michael Drmota and Clemens Heuberger; Article No. 11; pp. 11:1–11:13
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

A *deterministic finite automaton* (DFA) A is a 5-tuple $(\Sigma, Q, \delta, q_0, F)$, where Σ is a finite set of letters called the *alphabet*, Q is a finite set of states, $\delta : Q \times \Sigma \to Q$ is the *transition function*, q_0 is the *initial state*, and $F \subseteq Q$ is the set of *final states* (sometimes called *accept states*). States not in F are called *non-final* or *reject states*. A DFA can be represented by a directed graph with one vertex v_s for each state $s \in Q$, with the vertices corresponding to final states being highlighted, and for every transition $\delta(s, w) = \hat{s}$, there is an edge from s to $\hat{s}$ labeled w (see Figure 1).

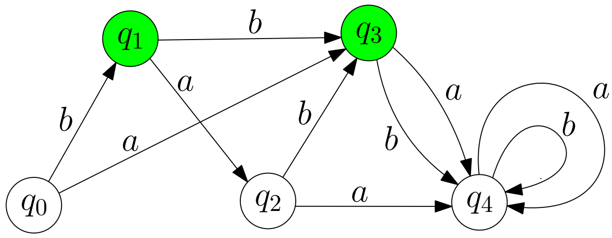


Figure 1 The unique minimal DFA for the language $\{a, b, bab, bb\}$. Here, q_0 is the initial state, q_1 and q_3 are the final states, and q_4 is the unique sink.

A word $w = w_1 w_2 \cdots w_\ell \in \Sigma^*$ is *accepted* by A if the sequence of states $(s_0, s_1, \ldots, s_\ell) \in Q^{\ell+1}$ defined by $s_0 = q_0$ and $s_{i+1} = \delta(s_i, w_i)$ for $i = 0, \ldots, \ell - 1$ ends with $s_\ell \in F$ a final state. The set of words accepted by A is called the *language* $\mathcal{L}(A)$ recognized by A. It is well-known that DFAs recognize exactly the set of regular languages. Note that every DFA recognizes a unique language, but a language can be recognized by several different DFAs. A DFA is called *minimal* if no DFA with fewer states recognizes the same language. The Myhill-Nerode Theorem states that every regular language is recognized by a unique minimal DFA (up to isomorphism) [8, Theorem 3.10]. For more details on automata see [8].

In this paper we show that the counting sequence $(m_{2,n})_{n\in\mathbb{N}}$ of minimal DFAs of size n recognizing a finite binary language admits a stretched exponential. Until now, the problem of counting these automata, even asymptotically, was widely open, see for example [4].

▶ **Theorem 1.** *The number $m_{2,n}$ of non-isomorphic minimal DFAs on a binary alphabet recognizing a finite language with $n + 1$ states satisfies for $n \to \infty$*

$$m_{2,n} = \Theta\left(n!\, 8^n e^{3a_1 n^{1/3}} n^{7/8}\right),$$

where $a_1 \approx -2.338$ is the largest root of the Airy function.

Since every regular language defines a unique minimal automaton, one may define the (space) complexity of the language to be the number of states in this corresponding automaton. Defining *space complexity* in this way, the number $m_{2,n}$ is simply the number of finite languages over a binary alphabet of space complexity $n + 1$.

In the recent paper [6] we showed lower and upper asymptotic bounds on $m_{2,n}$ by first establishing a connection between automata counted by $m_{2,n}$ and classes of directed acyclic graphs (DAGs) and then solving their asymptotic enumeration problem. In particular, we proved that

$$2^{n-1} c_n \leq m_{2,n} \leq 2^{n-1} r_n, \tag{1}$$

where c_n is the number of compacted and r_n the number of relaxed binary trees of size n. These appear naturally in the compression of XML documents [3, 7]. In the same paper, we showed that as $n \to \infty$,

$$c_n = \Theta\left(n!\, 4^n e^{3a_1 n^{1/3}} n^{3/4}\right) \qquad \text{and} \qquad r_n = \Theta\left(n!\, 4^n e^{3a_1 n^{1/3}} n\right),$$

leading to asymptotic lower and upper bounds on $m_{2,n}$. The results of the present work arise as a further application of the general method from [6] for proving the appearance of such stretched exponentials. They showcase the strength of our method, and we expect that our method may be applied to yet other combinatorial objects governed by similar recurrences.

The asymptotic proportion of general minimal DFAs (not necessarily recognizing a *finite* language) was solved by Bassino, Nicaud, and Sportiello in [1], building on enumeration results by Korshunov [9, 10] and Bassino and Nicaud [2]. The result in [1] also exploits an underlying tree structure of the related automata, but from a different traversal than what we use. In that case, no stretched exponential appears in the asymptotic enumeration, and the minimal automata account for a constant fraction of all automata.

2 Recurrence relation

To derive a recurrence for automata recognizing a finite language, we need the following lemma. In the following, we only consider automata on the binary alphabet $\{a, b\}$.

▶ **Lemma 2** ([11, Lemma 2.3], [8, Section 3.4]). *A DFA A is the minimal automaton for some finite language if and only if it has the following properties:*

(a) *There is a unique* sink s. *That is, a state which is not a final state such that all transitions from s end at s that is, $\delta(s, w) = s$.*

(b) *A is* acyclic*: the underlying directed graph has no cycles except for the loops at the sink.*

(c) *A is* initially connected*: for any state p there exists a word $w \in \Sigma^*$ such that A reaches the state p upon reading w.*

(d) *A is* reduced*: for any two different states q, q', the two automata with initial state q and q' recognize different languages.*

Next, we identify a property that can replace the one of being *reduced*.

▶ **Lemma 3.** *An acyclic, initially connected DFA A with a unique sink is reduced if and only if it satisfies the following condition:*

(d') State uniqueness: *there are no two distinct states q and q' with $\delta(q, a) = \delta(q', a)$ and $\delta(q, b) = \delta(q', b)$ such that both q and q', or neither q nor q', are accept states.*

Proof. By definition, being reduced implies state uniqueness. Now suppose that A is not reduced while state uniqueness holds. Then there are two states $q \neq q'$ in A such that the two automata with initial state q and q' recognize the same language L. As A is acyclic, L is finite. We define the weight of L to be $\sum_{w\in L}(|w| + 1)$, and we pick q, q' such that the weight of L is minimal.

Suppose that L is not empty. By the state uniqueness, we must have $\delta(q, a) \neq \delta(q', a)$ or $\delta(q, b) \neq \delta(q', b)$. Without loss of generality, suppose that $r = \delta(q, a) \neq \delta(q', a) = r'$. The two automata with initial state r and r' recognize the same language $a^{-1}L = \{w \mid aw \in L\}$. Since the weights of $a^{-1}L$ are strictly less that that of L, we have r and r' violating the minimality of the weight of L. Therefore, L must be empty.

Since L is empty, q and q' are both rejecting. They cannot both be the sink as the sink is unique. Suppose that q is not the sink. Then due to state uniqueness, among $\delta(q, a)$ and $\delta(q, b)$ there is at least one state q_1 that is not the sink. As L is empty, q_1 is also rejecting.

We can then replace q with q_1 and perform the same argument to q_1 and q', repeating *ad infinitum*. This creates an infinite sequence of states without repetition since A is acyclic. This is impossible as A is a DFA. Therefore, the existence of q and q' is impossible, meaning that A is reduced. We thus have the desired equivalence. ◀

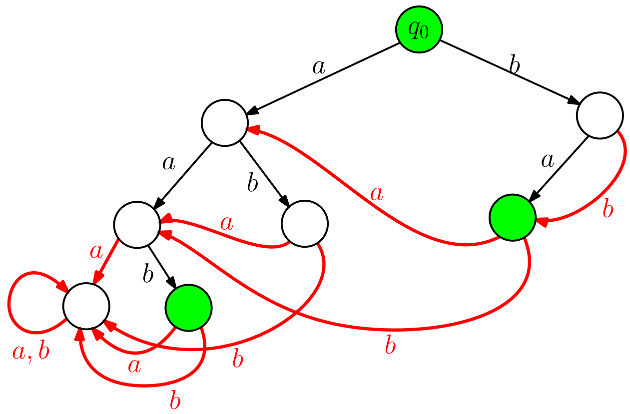


Figure 2 An acyclic DFA with its spanning subtree in black and all other edges in red. The initial state is q_0 and the finial states are colored green.

We now consider two sets of DFAs: the set $\mathfrak{F}$ of minimal DFAs recognizing finite languages, and the set $\mathfrak{G}$ of acyclic and initially connected DFAs with a unique sink. From Lemmas 2 and 3, $\mathfrak{F}$ consists of precisely the automata in $\mathfrak{G}$ that also possess the state uniqueness.

In order to derive our recurrence, we first transform DFAs in $\mathfrak{G}$ into decorated lattice paths that we call *B-paths*. For a given $A \in \mathfrak{G}$, our first step is to construct a spanning subtree of A (excluding the sink) using a depth-first search (*DFS* hereinafter) from the initial state q_0 as shown in Figure 2. This DFS is uniquely defined by taking edges marked by a before edges marked by b. Since A is initially connected, the tree obtained is a spanning tree.

Using the same DFS, we construct a path P starting at the point $(-1, 0)$ and illustrated by a blue line in Figure 3 as follows:

- Whenever the directed blue line around the tree in Figure 3 goes up we add a *vertical step* $V = (0, 1)$ to the path. We say that the state we just quit *corresponds* to this step.
- Whenever the directed blue line crosses an outgoing edge (including the edge leading to the sink), which is not part of the tree, we add a *horizontal step* $H = (1, 0)$.

The order of states corresponding to V-steps is called the *postorder* of states. It is clear that the first step of P is a H-step, and removing it from P gives a Dyck path under the main diagonal. We now decorate P with spots and crosses. Each step V is decorated by a green or white spot, according to whether the corresponding state is accepting or rejecting.

Since A is acyclic, during the DFS, for an edge e pointing from the current state q to an already visited state q', the state q' must not be an ancestor of q in the constructed tree, meaning that q' must either come before q in postorder or be the sink. In the former case, we put a cross in the cell at the intersection between the column of the H-step corresponding to e, and the row of the V-step corresponding to q', while in the latter case we put the cross in the row just below $y = 0$. Clearly the crosses are under P and above $y = -1$. We thus obtain a path B with decorations, and we say that B is the B-path of the automaton A.

To characterize B-paths obtained from DFAs in $\mathfrak{G}$, we propose the following definition. An *automatic B-path* P of size n is defined as a lattice path consisting of up steps and horizontal steps from $(-1, 0)$ to (n, n) with decorations such that

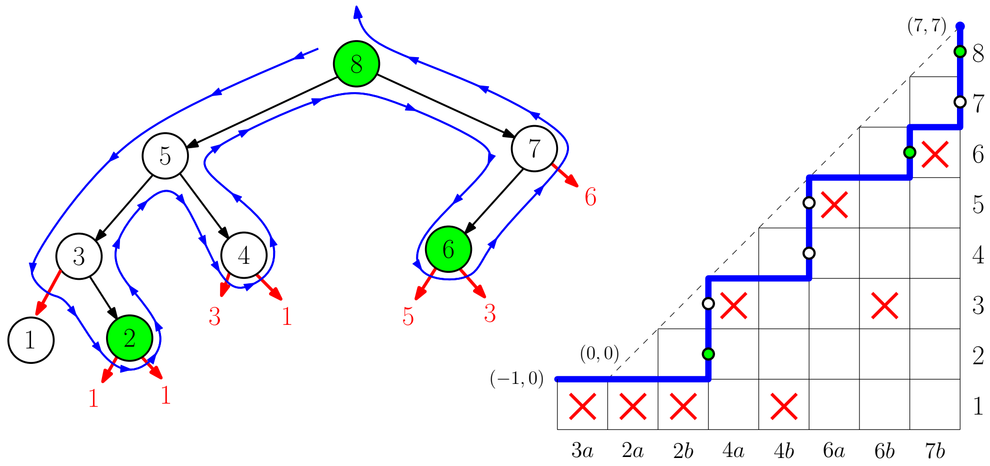


Figure 3 The transformation from an acyclic DFA to a B-path. In the DFA, the states are numbered in order of their corresponding up steps and we have labelled each outgoing edge not in the tree with the number of the state it points to.

- The first step is an H-step, and its removal leaves a Dyck path below the main diagonal;
- Each H-step has a cross in its column, under P and above $y = -1$.
- Every V-step has a white or green spot.

It is not difficult to see that automatic B-paths are in bijection with $\mathfrak{G}$, with the size preserved, since a B-path P obtained from a DFA $A \in \mathfrak{G}$ is clearly automatic, and the construction of B-paths can be easily reversed to obtain a DFA in $\mathfrak{G}$ from an automatic B-path.

Now we examine automatic B-paths corresponding to DFAs in $\mathfrak{F}$. By definition, we only need to take the state uniqueness into account. Given $A \in \mathfrak{G}$, let T be its depth-first search tree and B its corresponding automatic B-path. A state $q \in A$ is called a *cherry* if it is a leaf of T but not the sink. Seen on B, a cherry state corresponds to a sequence HHV of steps. We now propose a seemingly weaker notion of state uniqueness called *cherry-state uniqueness*, which is in fact equivalent in our case.

▶ **Lemma 4.** *Suppose that $A \in \mathfrak{G}$, then A has state uniqueness if and only if it has* cherry-state uniqueness*, i.e., any two states q, q' such that q comes before q' in postorder, and q' is a cherry state, satisfy the conditions in the definition of state uniqueness.*

Proof. State uniqueness clearly implies cherry-state uniqueness. For the other direction, let T be the DFS tree of A. Suppose that we have two states $q \neq q'$ such that $\delta(q, a) = \delta(q', a)$ and $\delta(q, b) = \delta(q', b)$. We suppose that q precedes q' in postorder. It is clear that q' is not an ancestor of q, but q is also not an ancestor of q, or else q would have a transition to itself or to one of its ancestors, which is impossible as A is acyclic. This implies that both $\delta(q, a)$ and $\delta(q, b)$ come before q in postorder, so neither $\delta(q, a)$ nor $\delta(q, b)$ can be a child of q'. Hence, q' is a cherry. Therefore, cherry-state uniqueness implies state uniqueness. ◀

We now try to construct step by step automatic B-paths corresponding to DFAs in $\mathfrak{F}$. We denote by $B_{n,m}$ the set of prefixes ending at (n, m) of such paths. We always start by an H-step from $(-1, 0)$, thus there is exactly one path in $B_{0,0}$. Suppose that we have constructed all automatic B-paths ending at $0 \leq m' \leq m$ and $m' \leq n' \leq n$ except for (n, m), and we now construct paths in $B_{n,m}$. First, from any path in $B_{n-1,m}$, we can construct a path $P \in B_{n,m}$ by adding an H-step at height m with a cross, and there are $(m + 1)$

possibilities for the cross. Second, from any path in $B_{n,m-1}$, we can construct a path P by adding a V-step with a spot that can be green or white. Such a path P ends in a V-step, thus it is different from paths in the first case. However, it may not be in $B_{n,m}$, because it may end in HHV with H-steps at height $m-1$. In such a case it corresponds to a cherry state that violates the cherry-state uniqueness. Such paths violating the condition for $\mathfrak{F}$ are all constructed by adding HHV at the end of paths in $B_{n-2,m-1}$, then adding crosses for the last two H-steps to make the corresponding cherry state "copy" one of the m states precedng it in postorder. Excluding such paths, we obtain all the paths in $B_{n,m}$. In this way, we construct all automatic B-paths corresponding to DFAs in $\mathfrak{F}$. This construction can be translated into the following recurrence.

▶ **Proposition 5.** *Let $b_{n,m}$ be the number of initial segments of automatic B-paths corresponding to DFAs in $\mathfrak{F}$ ending at (n,m). Then*

$$\begin{cases} b_{n,m} &= 2b_{n,m-1} + (m+1)b_{n-1,m} - mb_{n-2,m-1}, \quad \text{for } n \geq m \geq 1, \\ b_{n,m} &= 0, \quad \text{for } n < m, \\ b_{n,0} &= 1, \quad \text{for } n \geq -1. \end{cases}$$

The number $m_{2,n}$ of minimal binary DFAs of size n recognizing a finite language is equal to $b_{n,n}$.

This recurrence relation can be directly used to compute all elements of the sequence $(m_{2,n})_{n\geq 0}$ up to size $n = N$ with $\mathcal{O}(N^2)$ arithmetic operations. The first few numbers of this sequence read

$$(m_{2,n})_{n\geq 0} = (1, 1, 6, 60, 900, 18480, 487560, 15824880, 612504240, 27619664640, \ldots).$$

We have added it as sequence OEIS A331120 in the Online Encyclopedia of Integer Sequences[1]. Previously, the first 7 elements were computed in [5, Section 6].

3 A stretched exponential appears

We now perform an asymptotic analysis of the numbers $m_{2,n}$ using the recurrence derived in the previous section. As a first step we define an auxiliary sequence, which simplifies the subsequent analysis by absorbing the leading exponential behaviour:

$$\tilde{b}_{n,m} = \frac{b_{n,m}}{2^{m-1}}, \qquad \text{for } m \geq 1,$$
$$\tilde{b}_{n,0} = b_{n,0} = 1.$$

This gives

$$\begin{cases} \tilde{b}_{n,m} &= \tilde{b}_{n,m-1} + (m+1)\tilde{b}_{n-1,m} - \frac{m}{2}\tilde{b}_{n-2,m-1}, \quad \text{for } n \geq m > 1, \\ \tilde{b}_{n,m} &= 0, \quad \text{for } n < m, \\ \tilde{b}_{n,0} &= 1, \quad \text{for } n \geq -1. \end{cases}$$

Next, we transform the sequence $(\tilde{b}_{n,m})_{0\leq m\leq n}$ into a sequence $(e_{n,m})_{\substack{0\leq m\leq n \\ n-m \text{ even}}}$ using

$$e_{n,m} = \frac{1}{((n+m)/2)!}\tilde{b}_{(n+m)/2,(n-m)/2},$$

[1] `https://oeis.org`

(note that $e_{n,m}$ is only defined when $n-m$ is even). Then, the terms $e_{n,m}$ are determined by the following recurrence for $n, m \geq 1$

$$\begin{cases} e_{n,m} &= \frac{n-m+2}{n+m} e_{n-1,m-1} + e_{n-1,m+1} - \frac{n-m}{(n+m)(n+m-2)} e_{n-3,m-1}, & \text{for } n \geq m \geq 0, \\ e_{0,0} &= 1, & \\ e_{n,m} &= 0, & \text{for } n < m, \\ e_{n,-1} &= 0, & \text{for } n \geq -1. \end{cases}$$

The number of minimal DFAs of size n is equal to $n!2^{n-1}e_{2n,0}$. Now, for some simple cases of $e_{n,m}$, elementary computations show that $e_{n,n} = \frac{1}{n!}$, $e_{n,n-2} = \frac{2^{n-1}-1}{(n-1)!}$, and $e_{n,n-4} = \frac{3^{n-2}-3\cdot 2^{n-3}}{(n-2)!}$. Comparing the recurrence above with the one of compacted binary trees given in [6, Section 5] for $e_{n,m}$, we notice only two differences:

1. a slightly different factor $\frac{2(n-m-2)}{(n+m)(n+m-2)}$ of $e_{n-3,m-1}$ and
2. no special cases for $n \geq m > n-3$.

Therefore, we are anticipating the same method to be applicable. The very basic idea is that we will prove lower and upper bounds which differ only in the constant term. This method requires that the recurrence involves only non-negative terms on the right-hand side. As in the case of compacted binary trees, we solve this problem by finding suitable upper and lower bounds given in the subsequent Lemma. We omit its technical proof as it follows exactly the same lines as [6, Lemma 5.1].

▶ **Lemma 6.** *For $n-3 \geq m \geq 2$, the term $e_{n,m}$ is bounded below by*

$$L_e = \frac{n-m+2}{n+m} e_{n-1,m-1} + \frac{n-m-1}{n-m} e_{n-1,m+1} + \frac{n-m-3}{n-m-2} \left(\frac{1}{n-m} e_{n-2,m+2} + \frac{1}{n+m} e_{n-3,m+1} \right)$$

and for $n \geq 5, n > m \geq 0$ bounded above by

$$U_e = \frac{n-m+2}{n+m} e_{n-1,m-1} + \frac{n-m-1}{n-m} e_{n-1,m+1} + \frac{1}{n-m} e_{n-2,m+2} + \frac{1}{n+m} e_{n-3,m+1}.$$

That is, $L_e(n,m) \leq e_{n,m} \leq U_e(n,m)$.

3.1 Lower bound

The following technical Lemma is at the heart of the following inductive proof of the lower bound. It links the recurrence of $e_{n,m}$ (or rather its lower bound L_e) with two explicit sequences $\tilde{s}_n$ and $\tilde{X}_{n,m}$ involving the Airy function, shifted to its right-most root a_1.

▶ **Lemma 7.** *For all $n, m \geq 0$ let*

$$\tilde{X}_{n,m} := \left(1 - \frac{2m^2}{3n} + \frac{3m}{8n}\right) \operatorname{Ai}\left(a_1 + \frac{2^{1/3}(m+1)}{n^{1/3}}\right) \qquad \text{and}$$

$$\tilde{s}_n := 2 + \frac{2^{2/3}a_1}{n^{2/3}} + \frac{29}{12n} - \frac{1}{n^{7/6}}.$$

Then, for any $\varepsilon > 0$, there exists a constant $\tilde{n}_0$ such that

$$\begin{aligned} \tilde{X}_{n,m}\tilde{s}_n\tilde{s}_{n-1}\tilde{s}_{n-2} \leq{}& \frac{n-m+2}{n+m}\tilde{X}_{n-1,m-1}\tilde{s}_{n-1}\tilde{s}_{n-2} + \frac{n-m-1}{n-m}\tilde{X}_{n-1,m+1}\tilde{s}_{n-1}\tilde{s}_{n-2} \\ &+ \frac{n-m-3}{n-m-2}\left(\frac{1}{n-m}\tilde{X}_{n-2,m+2}\tilde{s}_{n-2} + \frac{1}{n+m}\tilde{X}_{n-3,m+1}\right), \end{aligned}$$

for all $n \geq \tilde{n}_0$ and all $0 \leq m < n^{2/3-\varepsilon}$.

Let us show how this Lemma is used before stating its actual proof. First, we define the sequence $X_{n,m} := \max\{\tilde{X}_{n,m}, 0\}$ (note that the factor $1 - \frac{2m^2}{3n} + \frac{3m}{8n}$ is negative for large m). Then, using Lemma 7 we have

$$\begin{aligned} X_{n,m}\tilde{s}_n\tilde{s}_{n-1}\tilde{s}_{n-2} \le {} & \frac{n-m+2}{n+m} X_{n-1,m-1}\tilde{s}_{n-1}\tilde{s}_{n-2} + \frac{n-m-1}{n-m} X_{n-1,m+1}\tilde{s}_{n-1}\tilde{s}_{n-2} \\ & + \frac{n-m-3}{n-m-2}\left(\frac{1}{n-m} X_{n-2,m+2}\tilde{s}_{n-2} + \frac{1}{n+m} X_{n-3,m+1}\right), \end{aligned}$$

for n large enough and all $m \le n$. Finally, we define the sequence $\tilde{h}_n$ such that $\tilde{h}_n = \tilde{s}_n\tilde{h}_{n-1}$ for $n > 0$ and set $\tilde{h}_0 = \tilde{s}_0$. Then we deduce by induction that $e_{n,m} \ge b_0\tilde{h}_n X_{n,m}$ for some constant $b_0 > 0$, all sufficiently large n, and all $m \in [0, n]$:

$$\begin{aligned} b_0 X_{n,m}\tilde{h}_n &\le \frac{n-m+2}{n+m} e_{n-1,m-1} + \frac{n-m-1}{n-m} e_{n-1,m+1} + \frac{n-m-3}{n-m-2}\left(\frac{e_{n-2,m+2}}{n-m} + \frac{e_{n-3,m+1}}{n+m}\right) \\ &\le e_{n,m}, \end{aligned}$$

where the first inequality follows by induction and the second one by Lemma 6 for $m \le n-3$. For $m > n-3$ and n large enough the inequality holds trivially as $X_{n,m} = 0$. Therefore,

$$\begin{aligned} m_{2,n} &= n!2^{n-1} e_{2n,0} \\ &\ge b_0 n! 2^{n-1}\tilde{h}_{2n} X_{2n,0} \\ &\ge b_0 n! 2^{n-1} \prod_{i=1}^{2n}\left(2 + \frac{2^{2/3}a_1}{i^{2/3}} + \frac{29}{12i} - \frac{1}{i^{7/6}}\right) \mathrm{Ai}\left(a_1 + \frac{1}{n^{1/3}}\right) \qquad (2) \\ &\ge \gamma_{\mathrm{L}} n! 8^n e^{3a_1 n^{1/3}} n^{7/8}, \end{aligned}$$

for some constant $\gamma_{\mathrm{L}} > 0$.

▶ Remark 8. Let us compare the result of Lemma 7 to the respective results for compacted and relaxed binary trees to which this method was applied first. Recall the lower and upper bounds (1) which are tight up to the constant and the polynomial term. Indeed, the corresponding results [6, Lemmas 4.2 and 5.2] possess a very similar structure: First, in $\tilde{X}_{n,m}$ the only difference is in the factor $\frac{3m}{8n}$ which is $\frac{m}{2n}$ for relaxed trees and $\frac{m}{4n}$ for compacted trees. The purpose of this term is of technical nature as it simplifies the Newton polygon method, yet it has no influence on the final asymptotics; compare Figure 5. Second, in $\tilde{s}_n$ the only difference is in the term $\frac{29}{12n}$ which is $\frac{8}{3n}$ for relaxed trees and $\frac{13}{6n}$ for compacted trees. Now this term influences the polynomial factor in the asymptotics (compare with [6, Section 3.3]). More generally, whenever the third term in the expansion of $\tilde{s}_n$ has the form $\frac{\alpha}{n}$, we get in the enumeration a polynomial factor with exponent $\frac{\alpha}{2} - \frac{1}{3}$. Finally, the similarity in all other terms of the expansion for $\tilde{s}_n$ and $\tilde{X}_{n,m}$ is responsible for the fact that $m_{2,n}$ and the families of trees enumerated in [6] have the same exponential growth, as well as the same stretched-exponential behaviour.

Proof (Lemma 7). The proof follows nearly verbatim [6, Lemma 4.2], so we will only introduce the main idea, omitting the technical details. Note that all (often tedious) computations are available in the accompanying Maple worksheet [12].

We start by defining the following sequence

$$\begin{aligned} P_{n,m} := {} & -Z_{n,m}s_n s_{n-1}s_{n-2} \\ & + \frac{n-m+2}{n+m} Z_{n-1,m-1}s_{n-1}s_{n-2} + \frac{n-m-1}{n-m} Z_{n-1,m+1}s_{n-1}s_{n-2} \\ & + \frac{n-m-3}{n-m-2}\left(\frac{1}{n-m} Z_{n-2,m+2}s_{n-2} + \frac{1}{n+m} Z_{n-3,m+1}\right), \end{aligned}$$

where

$$s_n := \sigma_0 + \frac{\sigma_1}{n^{1/3}} + \frac{\sigma_2}{n^{2/3}} + \frac{\sigma_3}{n} + \frac{\sigma_4}{n^{7/6}},$$
$$Z_{n,m} := \left(1 + \frac{\tau_2 m^2 + \tau_1 m}{n}\right) \mathrm{Ai}\left(a_1 + \frac{2^{1/3}(m+1)}{n^{1/3}}\right),$$

with $\sigma_i, \tau_j \in \mathbb{R}$. Then the inequality is equivalent to $P_{n,m} \geq 0$ with $\sigma_0 = 2$, $\sigma_1 = 0$, $\sigma_2 = 2^{2/3}a_1$, $\sigma_3 = 29/12$, and $\sigma_4 = -1$ as well as $\tau_1 = 3/8$ and $\tau_2 = -2/3$. Next, we expand $\mathrm{Ai}(z)$ in a neighborhood of

$$\alpha = a_1 + \frac{2^{1/3} m}{n^{1/3}}, \tag{3}$$

and we get

$$P_{n,m} = p_{n,m}\mathrm{Ai}(\alpha) + p'_{n,m}\mathrm{Ai}'(\alpha),$$

where $p_{n,m}$ and $p'_{n,m}$ are functions of m and n^{-1} and may be expanded as power series in $n^{-1/6}$ with coefficients polynomial in m. We will see that, as long as $n > 1$ and $n > m$, this series converges absolutely because the Airy function is entire and so all the functions for which we need to perform a bivariate expansion (in n and m) are indeed are analytic in the region defined by $|n| > 1$ and $|m| < |n|^{2/3-\varepsilon}$.

Now we proceed with the technical analysis, which is only performed on a superficial level here. The first step is to show that $[m^i n^j]P_{n,m} = 0$ for $i + j > 1$, $i, j \in \mathbb{Q}$. Then, as a second step, we strengthen this result by choosing suitable values σ_i for $0 \leq i \leq 4$ in the definition of s_n in order to eliminate more terms. The results are summarized in Figure 4 where the initial non-zero coefficients are shown. A diamond at (i, j) is drawn if and only if the coefficient $[m^i n^j]P_{n,m}$ is non-zero for generic values of σ and τ. It is an empty diamond if the given choice of σ_i and τ_j makes it vanish, whereas it is a solid diamond if it remains non-zero. The convex hull is formed by the following three lines

$$L_1 : j = -\frac{7}{6} - \frac{7i}{18}, \qquad L_2 : j = -\frac{1}{3} - \frac{2i}{3}, \qquad L_3 : j = 1 - i.$$

From now on, we distinguish between the contributions arising from $p_{n,m}$ and $p'_{n,m}$. The non-zero coefficients are shown in Figure 5. For technical reasons we choose at this point $\tau_1 = 8/3$ and thereby reduce the slope of the convex hull of the non-zero coefficients of $p'_{n,m}$. The expansions for n tending to infinity start as follows, where the elements on the convex hull are written in color:

$$\begin{aligned} P_{n,m} = \mathrm{Ai}(\alpha)\Bigg(&-\frac{4\sigma_4}{n^{7/6}} - \frac{2^{11/3}a_1 m}{3n^{5/3}} - \frac{164m^2}{9n^2} - \frac{2^{14/3}a_1 m^3}{3n^{8/3}} - \frac{136m^4}{9n^3} - \frac{248m^5}{135n^4} + \dots\Bigg) + \\ \mathrm{Ai}'(\alpha)\Bigg(&\frac{2^{1/3}(8\tau_1 - 3)}{n^{4/3}} + \frac{2^{7/3}}{n^{3/2}} - \frac{32a_1 m}{9n^2} + \frac{2^{4/3}m^2(48\tau_1 - 65)}{9n^{7/3}} - \frac{2^{19/3}m^3}{9n^{7/3}} \\ &-5\frac{2^{10/3}m^4}{9n^{10/3}} - 89\frac{2^{10/3}m^5}{135n^{13/3}} + \dots\Bigg). \end{aligned}$$

We now choose $\sigma_4 = -1$ which leads to a positive term $\mathrm{Ai}(\alpha)n^{-7/6}$. Next, for fixed (large) n we prove that for all m the dominant contributions in $P_{n,m}$ are positive. Motivated by Figures 4 and 5, we consider three different regimes: $m \leq Cn^{1/3}$, $Cn^{1/3} < m \leq n^{7/18}$, and $n^{7/18} < m < n^{2/3-\epsilon}$ for a suitable constant $C > 0$. We end the proof by showing that there exists an $N > 0$ such that all terms are positive for $n > N$ and all $m < n^{2/3}$. ◀

In the next section we will show an upper bound with the same asymptotic form, but with a different constant γ_U.

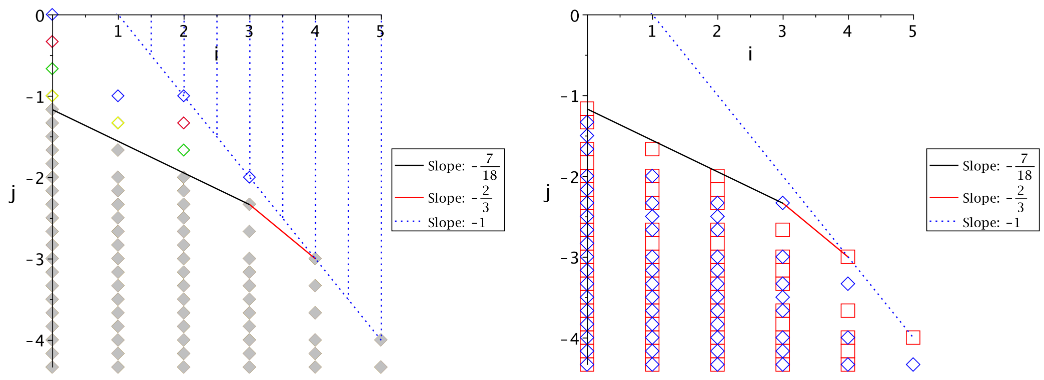


Figure 4 (Left) Non-zero coefficients of $P_{n,m} = \sum a_{i,j} m^i n^j$ shown by diamonds for $s_n := \sigma_0 + \frac{\sigma_1}{n^{1/3}} + \frac{\sigma_2}{n^{2/3}} + \frac{\sigma_3}{n} + \frac{\sigma_4}{n^{7/6}}$ and $Z_{n,m} := \left(1 + \frac{\tau_2 m^2 + \tau_1 m}{n}\right) \mathrm{Ai}\left(a_1 + \frac{2^{1/3}(m+1)}{n^{1/3}}\right)$. There are no terms in the blue dashed area. The blue terms vanish for $\sigma_0 = 2$, the red terms vanish for $\sigma_1 = 0$, the green terms vanish for $\sigma_2 = 2^{2/3} a_1$, and the yellow terms vanish for $\sigma_3 = 29/12$ and $\tau_2 = -2/3$. The black and red lines represent the two parts L_1 and L_2, respectively, of the convex hull. (Right) The solid gray diamonds are decomposed into the coefficients $p_{n,m}$ of $\mathrm{Ai}(\alpha)$ (red boxes) and $p'_{n,m}$ of $\mathrm{Ai}'(\alpha)$ (blue diamonds).

3.2 Upper bound

The following lemma links as in the case of the lower bound $e_{n,m}$ (and its upper bound U_e) with two explicit sequences $\hat{s}_n$ and $\hat{X}_{n,m}$ involving again the Airy function.

▶ **Lemma 9.** *Choose $\eta > 2/9$ fixed and for all $n, m \geq 0$ let*

$$\hat{X}_{n,m} := \left(1 - \frac{2m^2}{3n} + \frac{3m}{8n} + \eta \frac{m^4}{n^2}\right) \mathrm{Ai}\left(a_1 + \frac{2^{1/3}(m+1)}{n^{1/3}}\right) \qquad \text{and}$$

$$\hat{s}_n := 2 + \frac{2^{2/3} a_1}{n^{2/3}} + \frac{29}{12n} + \frac{1}{n^{7/6}}.$$

Then, for any $\varepsilon > 0$, there exists a constant $\hat{n}_0$ such that

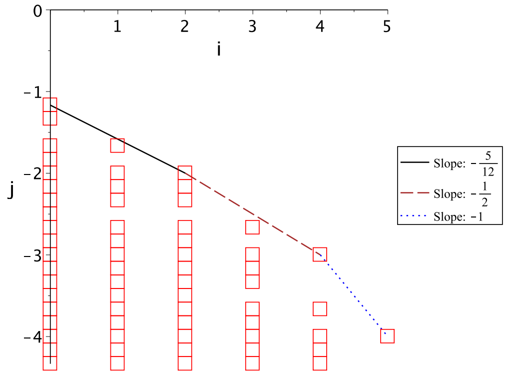
$$\begin{aligned} \hat{X}_{n,m} \hat{s}_n \hat{s}_{n-1} \hat{s}_{n-2} \geq{}& \frac{n-m+2}{n+m} \hat{X}_{n-1,m-1} \tilde{s}_{n-1} \tilde{s}_{n-2} + \frac{n-m-1}{n-m} \hat{X}_{n-1,m+1} \tilde{s}_{n-1} \tilde{s}_{n-2} \\ &+ \frac{1}{n-m} \hat{X}_{n-2,m+2} \tilde{s}_{n-2} + \frac{1}{n+m} \hat{X}_{n-3,m+1}, \end{aligned} \tag{4}$$

for all $n \geq \hat{n}_0$ and all $0 \leq m < n^{1-\varepsilon}$.

Proof (Sketch). The proof follows the same lines as that of Lemma 7, so we will only elucidate the required modifications. As a first step we define the following sequence

$$\begin{aligned} Q_{n,m} := {}& \hat{X}_{n,m} \hat{s}_n \hat{s}_{n-1} \hat{s}_{n-2} - \frac{n-m+2}{n+m} \hat{X}_{n-1,m-1} \tilde{s}_{n-1} \tilde{s}_{n-2} - \frac{n-m-1}{n-m} \hat{X}_{n-1,m+1} \tilde{s}_{n-1} \tilde{s}_{n-2} \\ &- \frac{1}{n-m} \hat{X}_{n-2,m+2} \tilde{s}_{n-2} - \frac{1}{n+m} \hat{X}_{n-3,m+1}. \end{aligned}$$

Then the inequality is equivalent to $Q_{n,m} \geq 0$. Again, we expand $\mathrm{Ai}(z)$ in a neighborhood of $\alpha = a_1 + \frac{2^{1/3} m}{n^{1/3}}$, and we get the following expansion (see the accompanying Maple worksheet [12] for full details). As before, the elements on the convex hull are written in color.



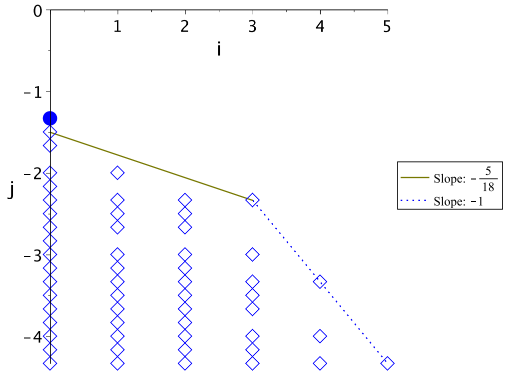


Figure 5 Non-zero coefficients $p_{n,m} = \sum \tilde{a}_{i,j} m^i n^j$ (red) and $p'_{n,m} = \sum \tilde{a}'_{i,j} m^i n^j$ (blue) of the expansion (3) for $P_{n,m}$. The coefficient of $n^{-4/3}$ in the right picture depicted as a solid blue circle disappears for $\tau_1 = 3/8$.

$$
\begin{aligned}
Q_{n,m} = \operatorname{Ai}(\alpha) \Bigg(& \frac{4}{n^{7/6}} + \frac{2^{11/3} a_1 m}{3n^{5/3}} + \frac{4m^2(41-108\eta)}{9n^2} + \frac{2^{14/3} a_1 m^3 (1-6\eta)}{3n^{8/3}} \\
& + \frac{8m^4(17-132\eta)}{9n^3} - \frac{2^{11/3} a_1 m^5 \eta}{n^{11/3}} - \frac{68m^6\eta}{3n^4} - \frac{124m^7\eta}{45n^5} + \dots \Bigg) + \\
\operatorname{Ai}'(\alpha) \Bigg(& \frac{2^{7/3}}{n^{3/2}} + \frac{32a_1 m}{9n^2} + \frac{2^{4/3} m^2 (47-216\eta)}{9n^{7/3}} + \frac{2^{16/3} m^3 (2-9\eta)}{9n^{7/3}} \\
& + \frac{2^{1/3} m^4 (40-549\eta)}{9n^{10/3}} - \frac{2^{16/3} m^5 \eta}{3n^{10/3}} - \frac{5m^6 2^{7/3} \eta}{3n^{13/3}} - \frac{89m^7 2^{7/3}\eta}{45n^{16/3}} + \dots \Bigg).
\end{aligned}
$$

Then we can finish in the same way as in the proof of Lemma 7. For the full details we refer to the proofs of [6, Lemma 4.4 and 5.3] which explains how to deal with the new cases required in the treatment of the upper bound (that happen to be analogous for the sequence at hand here, and the ones in that paper). Note that even the final convex hull in the Newton polygons is the same. ◀

The idea is now similar to the lower bound, yet a bit more intricate: We want to find an auxiliary sequence $(\tilde{e}_{n,m})_{n,m\geq 0}$ satisfying $e_{n,m} \leq C\tilde{e}_{n,m}$ for some constant $C > 0$, all n large, and all $m \leq n$ such that

$$\tilde{e}_{n,m} \leq \kappa_1 \hat{h}_n \hat{X}_{n,m}, \tag{5}$$

where the sequence $(\hat{h}_n)_{n\geq 1}$ is defined by $\hat{h}_n = \hat{s}_n \hat{h}_{n-1}$. As shown in (2), this implies that there is a constant $\gamma_{\mathrm{U}} > 0$ such that

$$\tilde{e}_{2n,0} \leq \gamma_{\mathrm{U}} 4^n e^{3a_1 n^{1/3}} n^{7/8}.$$

Now, in order to find such a sequence we use Lemma 6 and state the following definition for $(\tilde{e}_{n,m})_{n,m\geq 0}$:

$$
\begin{cases}
\tilde{e}_{n,m} &= \frac{n-m+2}{n+m}\tilde{e}_{n-1,m-1} + \frac{n-m-1}{n-m}\tilde{e}_{n-1,m+1} \\
&\quad + \frac{1}{n-m}\tilde{e}_{n-2,m+2} + \frac{1}{n+m}\tilde{e}_{n-3,m+1}, \qquad \text{for } n \geq 5, n^{3/4} > m \geq 0, \\
\tilde{e}_{n,m} &= e_{n,m}, \qquad \text{for } n < 5, n \geq m \geq 0, \\
\tilde{e}_{n,m} &= 0, \qquad \text{otherwise.}
\end{cases} \tag{6}
$$

There are several ideas in the choice of the sequence (6) which we want to explain now. Firstly, in order to prove (5), the sequence has to be zero for large values of m. We achieve this by cutting off the values for $m > n^{3/4}$. Secondly, it has to have positive coefficients, because then we can prove (5) by induction as it was done in the lower bound. Thirdly, it has to be an upper bound of $e_{n,m}$, i.e., $e_{n,m} \leq C\hat{e}_{n,m}$ for all n, m. Due to the cut off for $m > n^{3/4}$ this is, of course, impossible, so we introduce a second auxiliary sequence $(\hat{e}_{n,m})_{n,m\geq 0}$ with the same rules as (6) yet with no cut off, i.e., the recurrence holds for $n \geq 5$ and $n > m \geq 0$. Then, by Lemma 6 we have $e_{n,m} \leq \hat{e}_{n,m}$ for all n, m.

Hence, it remains to prove that there is a choice of N and a constant $C > 0$ such that

$$\hat{e}_{2n,0} \leq C\tilde{e}_{2n,0}$$

for all $n > N$. As a first step, we define a class $\mathcal{C}$ of weighted paths with the step set $\mathcal{S} := \{(1,1), (1,-1), (2,-2), (3,-1)\}$ and weights corresponding to the recurrence defining $\hat{e}_{n,m}$. Then $\hat{e}_{n,m}$ can be interpreted as the weighted enumeration of paths $p_0 p_1 \dots p_k \in \mathcal{C}$ ($p_i \in \mathbb{Z}^2$) from p_0 to $p_k = (n,m)$ such that $p_{i+1} - p_i \in \mathcal{S}$ for $0 \leq i \leq k-1$, with the additional initial condition that $p_0 = (u_0, v_0)$ and $p_1 = (u_1, v_1)$ satisfy $v_0 \leq u_0 < 5 \leq u_1$. In other words, the first jump $p_1 - p_0$ has to exit $\mathcal{I} := \{(i,j) : i < 5\}$. The weight given to each path in this enumeration is e_{u_0,v_0}

▶ **Lemma 10.** *Let $q_{\ell,m,2n}$ denote the weighted number of paths $p \in \mathcal{C}$ from (ℓ, m) to $(2n, 0)$. Then the numbers $q_{\ell,m,2n}$ satisfy the inequality*

$$\frac{q_{\ell,j,2n}}{j+1} \geq \frac{q_{\ell,k,2n}}{k+1},$$

for integers $0 \leq j < k \leq \ell \leq 2n$ satisfying $2|k-j$ and $n \geq 10$.

Proof (Sketch). Reversing the steps in (6) we see that q satisfies the following recurrence for $\ell < 2n$:

$$\begin{cases} q_{\ell,m,2n} &= 0, & \text{for} \quad m < 0, \\ q_{\ell,m,2n} &= \frac{\ell-m+1}{\ell-m+2} q_{\ell+1,m-1,2n} + \frac{\ell-m+2}{\ell+m+2} q_{\ell+1,m+1,2n} & \\ & \quad + \frac{1}{\ell-m+4} q_{\ell+2,m-2,2n} + \frac{1}{\ell+m+2} q_{\ell+3,m-1,2n} & \text{for} \quad m \geq 0. \end{cases}$$

Then we follow nearly verbatim the lines of the proof of [6, Lemma 5.4]. For more details we refer to the accompanying Maple worksheet [12]. ◀

The last ingredient we will need is that

$$\hat{e}_{2x,2y} \leq d_{2x,2y} \leq \binom{2x}{x+y},$$

where the sequence $d_{x,y}$ corresponds to the weighted number of Dyck meanders of length x ending at y; see [6, Proposition 3.2]. The first inequality is proved by induction using the recurrence relations of $\hat{e}_{x,y}$ and $d_{x,y}$. The second inequality is proved in [6], yet simply a consequence of the fact that $\binom{2x}{x+y}$ is the (unweighted) number of Dyck meanders from $(0,0)$ to $(2x, 2y)$, while the weights of weighted Dyck meanders are always smaller than 1.

Finally, among the $\hat{e}_{2n,0}$ weighted paths ending at $(2n, 0)$, the proportion of those passing through some point $(2x, 2y)$ is

$$\frac{\hat{e}_{2x,2y} q_{2x,2y,2n}}{\hat{e}_{2n,0}} \leq \frac{\hat{e}_{2x,2y} q_{2x,2y,2n}}{\hat{e}_{2x,0} q_{2x,0,2n}} \leq (2y+1)\frac{\hat{e}_{2x,2y}}{\hat{e}_{2x,0}} \leq \frac{2y+1}{\gamma_{\mathrm{L}} 4^x e^{3a_1 x^{1/3}} x^{3/4}} \binom{2x}{x+y}.$$

In the last inequality we used Lemma 10 as well as $e_{n,m} \leq \hat{e}_{n,m}$ and the lower bound (2) for $\hat{e}_{2x,0}$. Hence, we can use the same ideas as used in [6, Lemma 4.6] to show that there is some choice for N such that $\hat{e}_{2n,0} \leq 2\tilde{e}_{2n,0}$ for all n.

This proves the missing link and ends the proof of Theorem 1.

To conclude, we observe that all arguments in Section 2 can be extended to any finite alphabet of any size at least 2. Our analysis may also be extended to this more general case, but this remains a work in progress.

References

1. Frédérique Bassino, Julien David, and Andrea Sportiello. Asymptotic enumeration of minimal automata. In *29th International Symposium on Theoretical Aspects of Computer Science*, volume 14 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 88–99. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2012. `doi:10.4230/LIPIcs.STACS.2012.88`.
2. Frédérique Bassino and Cyril Nicaud. Enumeration and random generation of accessible automata. *Theoret. Comput. Sci.*, 381(1-3):86–104, 2007. `doi:10.1016/j.tcs.2007.04.001`.
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. Michael Domaratzki. Enumeration of formal languages. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS*, 89:117–133, 2006. URL: `http://eatcs.org/images/bulletin/beatcs89.pdf`.
5. Michael Domaratzki, Derek Kisman, and Jeffrey Shallit. On the number of distinct languages accepted by finite automata with n states. *J. Autom. Lang. Comb.*, 7(4):469–486, 2002. URL: `https://cs.uwaterloo.ca/~shallit/Papers/enum.ps`.
6. Andrew Elvey Price, Wenjie Fang, and Michael Wallner. Compacted binary trees admit a stretched exponential, 2019. `arXiv:1908.11181`.
7. 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`.
8. John E. Hopcroft and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation.* Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.
9. Aleksej D. Korshunov. Enumeration of finite automata. *Problemy Kibernet.*, 34:5–82, 1978. (In Russian).
10. Aleksej D. Korshunov. On the number of nonisomorphic strongly connected finite automata. *Elektron. Informationsverarb. Kybernet.*, 22(9):459–462, 1986.
11. Valery A Liskovets. Exact enumeration of acyclic deterministic automata. *Discrete Appl. Math.*, 154(3):537–551, 2006. `doi:10.1016/j.dam.2005.06.009`.
12. Michael Wallner. Personal website, 2020. URL: `http://dmg.tuwien.ac.at/mwallner`.