Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 63 (2005)

Smoothed Analysis of Binary Search Trees*

Bodo Manthey†,[1] Rüdiger Reischuk[2]

[1] Yale University
Department of Computer Science
manthey@cs.yale.de

[2] Universität zu Lübeck
Institut für Theoretische Informatik
reischuk@tcs.uni-luebeck.de

Abstract

Binary search trees are one of the most fundamental data structures. While the height of such a tree may be linear in the worst case, the average height with respect to the uniform distribution is only logarithmic. The exact value is one of the best studied problems in average-case complexity.

We investigate what happens in between by analysing the smoothed height of binary search trees: Randomly perturb a given (adversarial) sequence and then take the expected height of the binary search tree generated by the resulting sequence. As perturbation models, we consider partial permutations, partial alterations, and partial deletions.

On the one hand, we prove tight lower and upper bounds of roughly $\Theta((1-p) \cdot \sqrt{n/p})$ for the expected height of binary search trees under partial permutations and partial alterations, where n is the number of elements and p is the smoothing parameter. This means that worst-case instances are rare and disappear under slight perturbations. On the other hand, we examine how much a perturbation can increase the height of a binary search tree, i.e. how much worse well balanced instances can become.

Keywords: Smoothed Analysis, Binary Search Trees, Discrete Perturbations, Permutations.

1 Introduction

To explain the discrepancy between average-case and worst-case behaviour of the simplex algorithm, Spielman and Teng introduced the notion of *smoothed analysis* [34]. Smoothed analysis interpolates between average-case and worst-case analysis: Instead of taking the worst-case instance or, as in average-case analysis, choosing an instance completely at random, we analyse the complexity of (worst-case) objects subject to slight random perturbations, i.e. the expected complexity in a small

*An extended abstract of a preliminary version of this work appeared in the Proceedings of the 16th Ann. Int. Symposium on Algorithms and Computation (ISAAC 2005) [20].

†Work done at the Institute for Theoretical Computer Science of the University of Lübeck supported by DFG research grant RE 672/3. On leave from the Department of Computer Science, Universität des Saarlandes.

ISSN 1433-8092

neighbourhood of (worst-case) instances. Smoothed analysis takes into account that on the one hand typical instances are not necessarily random instances and that on the other hand worst-case instances are often artificial and rarely occur in practice.

Let C be some complexity measure. The worst-case complexity is $\max_x C(x)$, and the average-case complexity is $\mathbb{E}_{x \sim \Delta} C(x)$, where $\mathbb{E}$ denotes expectation with respect to a probability distribution Δ. The smoothed complexity is defined as $\max_x \mathbb{E}_{y \sim \Delta(x,p)} C(y)$. Here, x is chosen by an adversary and y is randomly chosen according to some probability distribution $\Delta(x, p)$ that depends on the adversarial instance x and a smoothing parameter p. The distribution $\Delta(x, p)$ should favour instances in the vicinity of x. This means that $\Delta(x, p)$ should put almost all weight on the neighbourhood of x, where "neighbourhood" has to be defined appropriately depending on the problem considered. The smoothing parameter p denotes how strong x is perturbed, i.e. we can view it as a parameter for the size of the neighbourhood of x. Intuitively, for $p = 0$, smoothed complexity becomes worst-case complexity, while for large p, smoothed complexity becomes average-case complexity.

For continuous problems, Gaussian perturbations seem to be a natural perturbation model: they are concentrated around their mean, and the probability that a perturbed number deviates from its unperturbed counterpart by distance d decreases exponentially in d. Thus, such probability distributions favour instances in the neighbourhood of the adversarial instance. There are, however, only few results about smoothed analysis of discrete problems. For such problems, even the term "neighbourhood" is often not well defined. Thus, special care is needed when defining perturbation models for discrete problems. Perturbation models should reflect "natural" perturbations, and the probability distribution for an instance x should be concentrated around x, particularly for small values of the smoothing parameter p.

Here, we will conduct a smoothed analysis of an ordering problem, namely the *smoothed height of binary search trees*. Binary search trees are one of the most fundamental data structures and, as such, building blocks for many advanced data structures. The main criteria of the "quality" of a binary search tree is its height, i.e. the length of the longest path from the root to a leaf. Unfortunately, the height is equal to the number of elements in the worst case, i.e. for totally unbalanced trees generated by an ordered sequence of elements. On the other hand, if a binary search tree is chosen at random, then the expected height is only logarithmic in the number of elements (more details will be discussed in Section 1.1.2). Thus, there is a huge discrepancy between the worst-case and the average-case behaviour of binary search trees.

We analyse what happens in between: An adversarial sequence will be perturbed randomly and then the height of the binary search tree generated by the perturbed sequence is measured. Thus, our instances are neither adversarial nor completely random. As perturbation models, we consider *partial permutations*, *partial alterations*, and *partial deletions*. For all three, we show tight lower and upper bounds. As a by-product, we obtain tight bounds for the smoothed number of left-to-right maxima, which is the number of new maxima seen when scanning a sequence from the left to the right, thus improving a result by Banderier et al. [3].

In smoothed analysis one analyses how fragile worst-case instances are. We suggest examining also the dual property: Given a good (or best-case) instance, how much can the complexity increase by slightly perturbing the instance? In other words, how stable are best-case instances under perturbations? For binary search trees, we show that there are best-case instances that indeed are not stable, i.e. there are sequences that yield trees of logarithmic height, but slightly perturbing the sequences yields trees of polynomial height.

1.1 Existing Results

Since we are concerned with smoothed analysis and binary search trees, we briefly review both areas.

1.1.1 Smoothed Analysis

Santha and Vazirani introduced the semi-random model, in which an adversary adaptively chooses a sequence of bits, each of which is corrupted independently with some fixed probability [28]. They showed how to obtain sequences of quasi-random bits from such semi-random sources. Their work inspired research on semi-random graphs [7, 17], which can be viewed as a forerunner of the smoothed analysis of discrete problems.

Spielman and Teng introduced smoothed analysis as a hybrid of average-case and worst-case complexity [34]. They showed that the simplex algorithm for linear programming with the shadow vertex pivot rule has polynomial smoothed complexity. This means that the running time of the algorithm is expected to be polynomial in terms of the input size and the variance of the Gaussian perturbation. Since then, smoothed analysis has been applied to a variety of fields [31], for instance several variants of linear programming [8, 12, 33], online and other algorithms [5, 18, 29], discrete optimisation [6, 27], and other topics [4, 10, 11, 32].

Banderier, Beier, and Mehlhorn [3] applied the concept of smoothed analysis to ordering problems. In particular, they analysed the number of left-to-right maxima of a sequence. Here the worst case is the sequence $1, 2, \ldots, n$, which yields n left-to-right maxima. On average, we expect $\sum_{i=1}^{n} 1/i \approx \ln n$ left-to-right maxima. Banderier et al. used the perturbation model of *partial permutations*, where each element of the sequence is independently selected with a given probability of $p \in [0, 1]$ and then a random permutation on the selected elements is performed (see Section 3.1 for a precise definition). Banderier et al. proved that the number of left-to-right maxima under partial permutations is $O(\sqrt{(n/p)\log n})$ in expectation for $0 < p < 1$. Furthermore, they showed a lower bound of $\Omega(\sqrt{n/p})$ for $0 < p \leq 1/2$.

1.1.2 Binary Search Trees

Given a sequence $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ of n distinct elements from any ordered set, we obtain a binary search tree $T(\sigma)$ by iteratively inserting the elements $\sigma_1, \sigma_2, \ldots, \sigma_n$ into the initially empty tree (this is formally described in Section 2.2).

The study of binary search trees is one of the most fundamental problems in computer science since they are the building blocks for a large variety of data structures (see for instance Aho et al. [1,2] and Knuth [19]). Beyond being an important data structure, binary search trees play a central role in the analysis of algorithms.

For instance, the height of $T(\sigma)$ is equal to the number of levels of recursion required by Quicksort when sorting σ if the first element is always chosen as the pivot (see for instance Cormen et al. [9]). Somehow related to the smoothed analysis of binary search trees is Seidel and Aragon's analysis of a randomised strategy for balancing binary search trees [30]: Their strategy remains efficient even if only a few random bits are available.

The worst-case height of a binary search tree is obviously n: just take $\sigma = (1, 2, \ldots, n)$. (We define the length of a path as the number of vertices.) The expected height of the binary search tree obtained from a random permutation (with all permutations being equally likely) has been the subject of a considerable amount of research in the past. Let the random variable $H(n)$ denote the height of a binary search tree obtained from a random permutation of n elements. Robson [23] proved that $\mathbb{E}H(n) \approx c\ln(n) + o(\ln(n))$ for some $c \in [3.63, 4.3112]$ and observed that $H(n)$ does not vary much from experiment to experiment [24]. Pittel [21] proved the existence of an α with $\alpha = \lim_{n\to\infty} \frac{\mathbb{E}H(n)}{\ln(n)}$. Devroye [13] then proved that $\alpha \approx 4.31107$ is the larger root of $\alpha\ln(2e/\alpha) = 1$. The variance of $H(n)$ was shown to be $O((\log\log n)^2)$ by Devroye and Reed [14] and by Drmota [15]. Robson [25] proved that the expectation of the absolute value of the difference between the height of two random trees is constant. Thus, the height of random trees is concentrated around the mean. Drmota [16] and Reed [22] discovered that the variance of $H(n)$ is actually $O(1)$. Furthermore, Reed [22] proved that the expectation of $H(n)$ is $\alpha\ln n + \beta\ln(\ln n) + O(1)$ with $\beta = \frac{3}{2\ln(\alpha/2)} \approx 1.953$. Finally, Robson [26] proved strong upper bounds on the probability of large deviations from the median. His results suggest that all moments of $H(n)$ are bounded from above by a constant.

Although the worst-case and average-case height of binary search trees are very well understood, nothing is known in between, i.e. when the sequences are not completely random, but the randomness is limited.

1.2 New Results

We will consider the height of binary search trees subject to slight random perturbations (*smoothed height*), i.e. the expected height under limited randomness.

1.2.1 Perturbation Models

The height of a binary search tree obtained from a sequence of elements depends only on the ordering of the elements. Therefore, we use perturbation models that slightly perturb the order of the elements of the sequence. We consider three perturbation models (formally defined in Section 3).

Partial permutations, introduced by Banderier et al. [3], rearrange some elements, i.e. they randomly permute a small subset of the elements.

The other two perturbation models are new. *Partial alterations* do not move elements, but replace some elements by new elements chosen at random. Thus, they change the rank of the elements. *Partial deletions* remove some of the elements of the sequence without replacement, i.e. they shorten the input. This model turns out to be useful for analysing the other two models.

1.2.2 Lower and Upper Bounds

We prove matching lower and upper bounds for the expected height of binary search trees under all three perturbation models (Section 6). More precisely: For all smoothing parameters p with $p \leq 1-\epsilon$ and $p \geq n^{\epsilon-1}$, for an arbitrary but fixed $\epsilon > 0$, and all sequences of length n, the expectation of the height of a binary search tree obtained via p-partial permutation is at most $6.7 \cdot (1-p) \cdot \sqrt{n/p}$ for all sufficiently large n. In particular, the bounds hold for all constant values of $p \in (0,1)$. On the other hand, the expected height of a binary search tree obtained from the sorted sequence via p-partial permutation is at least $0.8 \cdot (1-p) \cdot \sqrt{n/p}$, which matches the upper bound up to a constant factor.

For the number of left-to-right maxima under partial permutations, we are able to prove an even better upper bound of $3.6 \cdot (1-p) \cdot \sqrt{n/p}$ for all sufficiently large n and a lower bound of $0.6 \cdot (1-p) \cdot \sqrt{n/p}$ (Section 5).

All these bounds hold for partial alterations as well.

Thus, under limited randomness, the behaviour of binary search trees differs significantly from both the worst case and the average case.

1.2.3 Smoothed Analysis and Stability

In smoothed analysis one analyses how fragile worst case instances are. We suggest examining also the dual property: Given a good (or best-case) instance, how much can the complexity increase if the instance is perturbed slightly? In other words, how stable are best-case instances under perturbations?

The lower and upper bound for partial deletions are straightforward. The main reason for considering partial deletions is that we can bound the expected height under partial alterations and permutations by the expected height under partial deletions (Section 7). The converse holds as well, we only have to blow up the sequences quadratically.

We exploit this when considering the stability of the perturbation models in Section 8: We prove that partial deletions and, thus, partial permutations and partial alterations as well are quite unstable, i.e. they can cause best-case instances to become much worse. More precisely: There are sequences of length n that yield trees of height $O(\log n)$, but the expected height of the tree obtained after smoothing is $n^{\Omega(1)}$.

2 Preliminaries

2.1 Notation

We denote by log and ln the logarithm to base 2 and e, respectively, while exp denotes the exponential function to base e. For $n \in \mathbb{N}$, let $[n] = \{1, 2, \ldots, n\}$.

Let $\sigma = (\sigma_1, \ldots, \sigma_n) \in S^n$ for some ordered set S. We call σ a **sequence**. Usually, we assume that all elements of σ are distinct, i.e. $\sigma_i \neq \sigma_j$ for all $i \neq j$. The length of σ is n. In most cases, σ will simply be a permutation of $[n]$. We denote the sorted sequence $(1, 2, \ldots, n)$ by σ^n_{sort}.

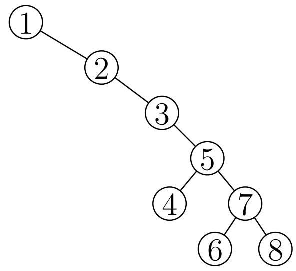


Figure 1: $T(\sigma)$ obtained from $\sigma = (1, 2, 3, 5, 7, 4, 6, 8)$. We have $\text{height}(\sigma) = 6$.

Let $\tau = (\tau_1, \ldots, \tau_t)$. We call τ a **subsequence of σ** if there are indexes $i_1 < i_2 < \ldots < i_t$ with $\tau_j = \sigma_{i_j}$ for all $j \in [t]$. Let $\mu = \{i_1, \ldots, i_t\} \subseteq [n]$. Then $\sigma_\mu = (\sigma_{i_1}, \ldots, \sigma_{i_t})$ denotes the subsequence consisting of all elements of σ at positions in μ. For instance, $\sigma_{[k]}$ denotes the prefix of length k of σ. When it is clear from the context, we sometimes use σ_μ to mean the set of elements at positions in μ, i.e. in this case $\sigma_\mu = \{\sigma_i \mid i \in \mu\}$. For $\mu \subseteq [n]$, we define $\overline{\mu} = [n] \setminus \mu$.

2.2 Binary Search Trees and Left-to-right Maxima

Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ be a sequence. We obtain a **binary search tree $T(\sigma)$** from σ by iteratively inserting the elements $\sigma_1, \sigma_2, \ldots, \sigma_n$ into the initially empty tree as follows:

- The root of $T(\sigma)$ is the first element σ_1 of σ.
- Let $\sigma_< = \sigma_{\{i \mid \sigma_i < \sigma_1\}}$ be σ restricted to elements smaller than σ_1. The left subtree of the root σ_1 of $T(\sigma)$ is obtained inductively from $\sigma_<$.

  Analogously, let $\sigma_> = \sigma_{\{i \mid \sigma_i > \sigma_1\}}$ be σ restricted to elements greater than σ_1. The right subtree of σ_1 of $T(\sigma)$ is obtained inductively from $\sigma_>$.

Figure 1 shows an example. We denote the height of $T(\sigma)$ by **height(σ)**, i.e. $\text{height}(\sigma)$ is the number of nodes on the longest path from the root to a leaf.

The element σ_i is called a **left-to-right maximum** of σ if $\sigma_i > \sigma_j$ for all $j \in [i-1]$. Let **ltrm(σ)** denote the number of left-to-right maxima of σ. We have $\text{ltrm}(\sigma) \leq \text{height}(\sigma)$ since the number of left-to-right maxima of a sequence is equal to the length of the right-most path in the tree $T(\sigma)$.

2.3 Probability Theory

We denote probabilities by $\mathbb{P}$ and expectations by $\mathbb{E}$. To bound large deviations from the mean of binomially distributed random variables, we will frequently use Chernoff bounds [35, Chapter B]. Let $p \in [0, 1]$, and let $X_1, X_2, \ldots, X_k$ be mutually independent random variables with $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$ and $X = \sum_{i=1}^{k} X_i$. Clearly, $\mathbb{E}(X) = pk$. The probability that X is δpk smaller or larger than its expectation is bounded from above by

$$\begin{aligned}\mathbb{P}\big(X < (1-\delta)\cdot pk\big) &< \exp\big(-\frac{pk\delta^2}{2}\big) \text{ and}\\ \mathbb{P}\big(X > (1+\delta)\cdot pk\big) &< \exp\big(-\frac{pk\delta^2}{4}\big) ,\end{aligned}$$

respectively. The first inequality holds for $\delta \in (0,1]$ while the second one holds for $\delta \in (0, 2e-1]$. The following lemma follows immediately from the inequalities above.

Lemma 1. *Let $k \in \mathbb{N}$, $\alpha \in (1,2)$, and $p \in [0,1]$. Let $X_1, \ldots, X_k$ be mutually independent random variables as above. Then*

$$\mathbb{P}\big((X > \alpha pk) \vee (X < \alpha^{-1} pk)\big) \leq 2\cdot \exp\big(-\frac{pk\cdot(1-\frac{1}{\alpha})^2}{4}\big) .$$

We will frequently use the following lemma to bound deviations from the means. We will need the cases $k \in O(\sqrt{n/p}\cdot \operatorname{polylog}(n)) \cap \Omega(\sqrt{n/p}/\operatorname{polylog}(n))$ as well as $k = n$, which are covered since $k \geq n^{-\epsilon/8}\cdot\sqrt{n/p}$ and $p \geq n^{\epsilon-1}$.

Lemma 2. *Fix $\epsilon > 0$. Let $n \in \mathbb{N}$, $p \geq n^{\epsilon-1}$, $k \geq n^{-\epsilon/8}\cdot\sqrt{n/p}$, and $\alpha = 1 + n^{-\epsilon/8}$. Let X be as in Lemma 1. Then*

$$\mathbb{P}\big((X > \alpha pk) \vee (X < \alpha^{-1} pk)\big) \leq 2\cdot \exp\left(-n^{\epsilon/8}/16\right) .$$

Proof. We have $1 - \frac{1}{\alpha} = \frac{n^{-\epsilon/8}}{1+n^{-\epsilon/8}} \geq \frac{n^{-\epsilon/8}}{2}$. By Lemma 1, we obtain

$$\begin{aligned}\mathbb{P}\big((X > \alpha pk) \vee (X < \alpha^{-1} pk)\big) &\leq 2\cdot\exp\big(-\frac{pk\cdot(1-\frac{1}{\alpha})^2}{4}\big)\\ \leq 2\cdot\exp\big(-\frac{n^{-\epsilon/8}\cdot\sqrt{np}\cdot n^{-\epsilon/4}}{16}\big) &\leq 2\cdot\exp\big(-\frac{n^{\epsilon/8}}{16}\big) .\end{aligned}$$

□

3 Perturbation Models for Permutations

Since we deal with ordering problems, we need perturbation models that slightly change a given permutation of elements. There seem to be two natural possibilities: Either *change the positions* of some elements, or *change the elements* themselves.

Partial permutations implement the first option: A subset of the elements is randomly chosen, and then these elements are randomly permuted. The second possibility is realised by partial alterations. Again, a subset of the elements is chosen randomly. These elements are then replaced by random elements. The third model, partial deletions, also starts by randomly choosing a subset of the elements. These elements are then removed without replacement.

For all three models, we obtain the random subset as follows. Let σ be a sequence of length n and $p \in [0,1]$ be a probability. Every element of σ is marked

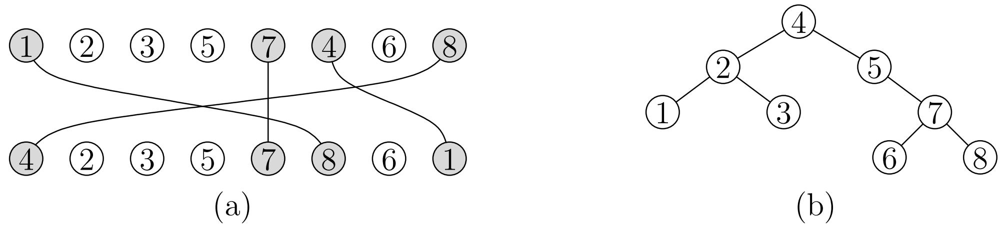


Figure 2: A partial permutation. (a) Top: The sequence $\sigma = (1, 2, 3, 5, 7, 4, 6, 8)$ with $\mu = M_p^n = \{1, 5, 6, 8\}$; Figure 1 shows $T(\sigma)$. Bottom: The marked elements are randomly permuted. The result is $\sigma' = \Pi(\sigma, \mu) = (4, 2, 3, 5, 7, 8, 6, 1)$. (b) $T(\sigma')$ with $\text{height}(\sigma') = 4$.

independently of the others with probability p. More formally: M_p^n is a random subset of $[n]$ with $\mathbb{P}(i \in M_p^n) = p$ for all $i \in [n]$. For any $\mu \subseteq [n]$ we have $\mathbb{P}(M_p^n = \mu) = p^{|\mu|} \cdot (1-p)^{|\overline{\mu}|}$.

Let $\mu \subseteq [n]$ be the set of marked positions. If $i \in \mu$, then we say that position i and element σ_i are marked. Thus, σ_μ is the sequence (or set) of all marked elements.

In the following, σ is always a permutation of $[n]$.

We denote by **height-perm**$_{\boldsymbol{p}}\boldsymbol{(\sigma)}$, **height-alter**$_{\boldsymbol{p}}\boldsymbol{(\sigma)}$, and **height-del**$_{\boldsymbol{p}}\boldsymbol{(\sigma)}$ the random variable of the height of the tree $T(\sigma')$, where σ' is obtained from σ by p-partial permutation, alteration, and deletion, respectively (all three models will be defined formally in the following). Analogously, **ltrm-perm**$_{\boldsymbol{p}}\boldsymbol{(\sigma)}$, **ltrm-alter**$_{\boldsymbol{p}}\boldsymbol{(\sigma)}$, and **ltrm-del**$_{\boldsymbol{p}}\boldsymbol{(\sigma)}$ denote the random variables of the number of left-to-right maxima of the sequence σ' obtained from σ via p-partial permutation, alteration, and deletion, respectively.

3.1 Partial Permutations

The notion of ***p*-partial permutations** was introduced by Banderier et al. [3]. Given a random subset M_p^n of $[n]$, the elements at positions in M_p^n are permuted according to a permutation drawn uniformly at random: Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ and $\mu \subseteq [n]$. Then the sequence $\sigma' = \Pi(\sigma, \mu)$ is a random sequence with the following properties:

- Π chooses a permutation π of μ uniformly at random and
- sets $\sigma'_{\pi(i)} = \sigma_i$ for all $i \in \mu$ and $\sigma'_i = \sigma_i$ for all $i \notin \mu$.

Figure 2 illustrates partial permutations.

By varying p, we can interpolate between the average and the worst case: for $p = 0$, no element is marked and $\sigma' = \sigma$, while for $p = 1$, all elements are marked and σ' is a random permutation of the elements of σ with all permutations being equally likely.

Let us show that partial permutations are indeed a suitable perturbation model by proving that the distribution of $\Pi(\sigma, M_p^n)$ favours sequences close to σ. To do this, we have to introduce a metric on sequences. Let σ and τ be two sequences of length n. We define the distance $d(\sigma, \tau)$ between σ and τ as $d(\sigma, \tau) = |\{i \mid \sigma_i \neq \tau_i\}|$, thus d

is a metric. Note that $d(\sigma,\tau)=1$ is impossible since there are no two permutations that differ in exactly one position.

The distribution of $\Pi(\sigma, M_p^n)$ is symmetric around σ with respect to d, i.e. the probability that $\Pi(\sigma, M_p^n)=\tau$ depends only on $d(\sigma,\tau)$.

Lemma 3. *Let $p \in (0,1)$, and let σ and τ be permutations of $[n]$ with $d = d(\sigma,\tau)$. Then*

$$\mathbb{P}\big(\Pi(\sigma, M_p^n)=\tau\big) = \sum_{k=0}^{n-d} p^{k+d}\cdot(1-p)^{n-d-k}\cdot\binom{n-d}{k}\cdot\frac{1}{(k+d)!}\,.$$

Proof. All d positions where σ and τ differ must be marked. This happens with probability p^d. The probability that k of the remaining positions are marked is $\binom{n-d}{k}\cdot p^k\cdot(1-p)^{n-d-k}$. Thus, the probability that $k+d$ positions are marked, d of which are positions where σ and τ differ, is $\binom{n-d}{k}\cdot p^{k+d}\cdot(1-p)^{n-d-k}$. There is only one permutation that maps σ to τ, which is chosen with probability $\frac{1}{(k+d)!}$. □

Let $\mathbb{P}_d = \sum_{k=0}^{n-d} p^{k+d}\cdot(1-p)^{n-d-k}\cdot\binom{n-d}{k}\cdot\frac{1}{(k+d)!}$ be the probability that $\Pi(\sigma, M_p^n)=\tau$ for a fixed sequence τ with distance d to σ. Then $\mathbb{P}_d$ tends exponentially to zero with increasing d. Thus, the distribution of $\Pi(\sigma, M_p^n)$ is highly concentrated around σ.

Lemma 4. *Let $p \in (0,1)$ be fixed. There exists a positive constant $c<1$ such that for all sufficiently large n, we have $\mathbb{P}_2 \le c\cdot\mathbb{P}_0$ and $\mathbb{P}_{d+1}\le c\cdot\mathbb{P}_d$ for all d with $2\le d<n$.*

Proof. By omitting the last summand, we obtain

$$\mathbb{P}_d \ge \sum_{k=0}^{n-d-1} p^{k+d}\cdot(1-p)^{n-d-k}\cdot\binom{n-d}{k}\cdot\frac{1}{(k+d)!}\,.$$

Thus,

$$\begin{aligned}
\frac{\mathbb{P}_{d+1}}{\mathbb{P}_d} &\le \frac{\sum_{k=0}^{n-d-1} p^{k+d+1}\cdot(1-p)^{n-(d+1)-k}\cdot\binom{n-(d+1)}{k}\cdot\frac{1}{(k+d+1)!}}{\sum_{k=0}^{n-d-1} p^{k+d}\cdot(1-p)^{n-d-k}\cdot\binom{n-d}{k}\cdot\frac{1}{(k+d)!}}\\
&\le \max_{0\le k\le n-d-1}\left(\frac{p^{k+d+1}\cdot(1-p)^{n-d-1-k}\cdot\binom{n-d-1}{k}\cdot\frac{1}{(k+d+1)!}}{p^{k+d}\cdot(1-p)^{n-d-k}\cdot\binom{n-d}{k}\cdot\frac{1}{(k+d)!}}\right)\\
&\le \frac{p}{1-p}\cdot\max_{0\le k\le n-d-1}\left(\frac{n-d-k}{(n-d)\cdot(k+d+1)}\right)\le\frac{p}{1-p}\cdot\frac{1}{d+1}\,.
\end{aligned}$$

The second inequality holds because $\sum_{i\in I} a_i/\sum_{i\in I} b_i \le \max_{i\in I} a_i/b_i$ for any set I and nonnegative numbers a_i and b_i ($i\in I$). This proves the lemma for all d with $d+1>\frac{p}{1-p}$.

What remains to be considered is $d\le\frac{p}{1-p}-1$. Fix $\beta>1$ arbitrarily with $\beta p<1$. Then $\mathbb{P}_{d+1}=\sum_{k=0}^{n-d-1} p^{k+d+1}\cdot(1-p)^{n-d-1-k}\cdot\binom{n-d-1}{k}\cdot\frac{1}{(k+d+1)!}$ is dominated by the summands with $k < \beta pn$: According to Chernoff bounds, we have

$$\begin{aligned} & \sum_{0 \le k \le n-d-1} p^{k+d+1} \cdot (1-p)^{n-d-k-1} \cdot \binom{n-d-1}{k} \\ \le \quad & (1+o(1)) \cdot \sum_{0 \le k < \beta pn} p^{k+d+1} \cdot (1-p)^{n-d-k-1} \cdot \binom{n-d-1}{k} . \end{aligned}$$

The additional factor of $\frac{1}{(k+d+1)!}$ in each summand of $\mathbb{P}_{d+1}$ and $\mathbb{P}'_{d+1}$ strengthens the dominance of the terms for $k < \beta pn$ since $\frac{1}{(k+d)!}$ is monotonically decreasing in k. Overall, we have $\mathbb{P}_{d+1} \le (1+o(1)) \cdot \mathbb{P}'_{d+1}$.

Furthermore, we define

$$\mathbb{P}'_d = \sum_{0 \le k < \beta pn} p^{k+1+d} \cdot (1-p)^{n-d-k-1} \cdot \binom{n-d}{k+1} \cdot \frac{1}{(k+1+d)!} \le \mathbb{P}_d \, .$$

Now we have $\frac{\mathbb{P}_{d+1}}{\mathbb{P}_d} \le (1+o(1)) \cdot \frac{\mathbb{P}'_{d+1}}{\mathbb{P}_d} \le (1+o(1)) \cdot \frac{\mathbb{P}'_{d+1}}{\mathbb{P}'_d}$ and

$$\begin{aligned} \frac{\mathbb{P}'_{d+1}}{\mathbb{P}'_d} & \le \max_{0 \le k < \beta pn} \left(\frac{p^{k+d+1} \cdot (1-p)^{n-d-1-k} \cdot \binom{n-d-1}{k} \cdot \frac{1}{(k+d+1)!}}{p^{k+1+d} \cdot (1-p)^{n-d-k-1} \cdot \binom{n-d}{k+1} \cdot \frac{1}{(k+1+d)!}} \right) \\ & \le \max_{0 \le k < \beta pn} \left(\frac{k+1}{n-d} \right) = \frac{\beta pn}{n-d} \le \beta p + o(1) \end{aligned}$$

for sufficiently large n. The last inequality holds because $d \le \frac{p}{1-p} - 1 \in O(1)$. Thus, there exists a $c < 1$ with $\mathbb{P}_{d+1}/\mathbb{P}_d \le \beta p + o(1) \le c$ for sufficiently large n. Finally, the proof above yields $\mathbb{P}_2/\mathbb{P}_0 = \frac{\mathbb{P}_2 \cdot \mathbb{P}_1}{\mathbb{P}_1 \cdot \mathbb{P}_0} \le c^2 \le c < 1$. □

3.2 Partial Alterations

Let us now introduce ***p*-partial alterations**. Every element at a position in M_p^n is replaced by a real number drawn uniformly and independently at random from $[\frac{1}{2}, n + \frac{1}{2})$ to obtain a sequence σ'. All elements in σ' are distinct with probability one. (We could also draw the random numbers from $[0, n+1)$. The results would be the same though more technical.)

Like partial permutations, partial alterations interpolate between the worst case ($p = 0$) and the average case ($p = 1$). Partial alterations are somewhat easier to analyse: The majority of results on the average-case height of binary search trees is actually not obtained by considering random permutations (cf. e.g. [13, 14, 22]). Instead, the binary search trees are grown from a sequence of n random variables that are uniformly and independently drawn from $[0, 1)$. This corresponds to partial alterations for $p = 1$. There is no difference between partial permutations and partial alterations for $p = 1$. This appears to hold for all p in the sense that the lower and upper bounds obtained for partial permutations and partial alterations are equal for all p (see Conjecture 32).

3.3 Partial Deletions

As the third perturbation model, we introduce **p-partial deletions**: Again, we have a random marking M_p^n as in Section 3.1. Then we delete all marked elements to obtain the sequence $\sigma_{\overline{M_p^n}}$.

Partial deletions do not really perturb a sequence: any ordered sequence remains ordered even if elements are deleted. The main reason for considering partial deletions is that they are easy to analyse when considering the stability of perturbation models (Section 8). The results for partial deletions then carry over to partial permutations and alterations since the expected heights with respect to these three models are closely related (Section 7).

4 Basic Properties

4.1 Properties of Binary Search Trees

We start by introducing a new measure for the height of binary search trees. Let $\mu \subseteq [n]$ and let σ be a sequence of length n. The **μ-restricted height of $T(\sigma)$**, denoted by **height(σ, μ)**, is the maximum number of elements of σ_μ on a root-to-leaf path in $T(\sigma)$.

Lemma 5. *For all sequences σ of length n and $\mu \subseteq [n]$,*

$$\begin{aligned} \text{height}(\sigma) &\leq \text{height}(\sigma,\mu) + \text{height}(\sigma,\overline{\mu}) \quad \textit{and} \\ \text{height}(\sigma,\mu) &\leq \text{height}(\sigma_\mu)\,. \end{aligned}$$

Proof. Consider any path of maximum length from the root to a leaf in $T(\sigma)$. This path consists of at most height(σ, μ) elements of σ_μ and at most height$(\sigma, \overline{\mu})$ elements of $\sigma_{\overline{\mu}}$, which proves the first part.

For the second part, let a and b be elements of σ_μ that do not lie on the same root-to-leaf path in $T(\sigma_\mu)$. Assume that $a < b$. Then there exists a c prior to a and b in σ_μ with $a < c < b$. Thus, a and b do not lie on the same root-to-leaf path in the tree $T(\sigma)$ either, which implies the lemma. □

Of course we have height$(\sigma, \mu) \leq$ height(σ) for all σ and μ. But height$(\sigma_\mu) \leq$ height(σ), which would imply height-del$_p(\sigma) \leq$ height(σ), does not hold in general. This will be investigated further in Section 8, when we consider the stability of the perturbation models.

To bound the smoothed height from above, we will use the following lemma, which is an immediate consequence of Lemma 5.

Lemma 6. *For all sequences σ of length n and $\mu \subseteq [n]$, we have*

$$\text{height}(\sigma) \leq \text{height}(\sigma_\mu) + \text{height}(\sigma,\overline{\mu})\,.$$

Proof. We have height$(\sigma) \leq$ height(σ, μ) + height$(\sigma, \overline{\mu}) \leq$ height(σ_μ) + height$(\sigma, \overline{\mu})$ according to Lemma 5. □

We can state equivalent lemmas for left-to-right maxima. Let σ be a sequence of length n and $\mu \subseteq [n]$. Then $\mathbf{ltrm}(\boldsymbol{\sigma}, \boldsymbol{\mu})$ denotes the **$\boldsymbol{\mu}$-restricted number of left-to-right maxima** of σ, i.e. the number of elements σ_i such that $i \in \mu$ and σ_i is a left-to-right maximum of σ. We omit the proof of the following lemma since it is almost identical to the proofs of the lemmas above.

Lemma 7. *Let σ be a sequence of length n and $\mu \subseteq [n]$. Then*

$$\begin{aligned} \mathrm{ltrm}(\sigma) &\leq \mathrm{ltrm}(\sigma, \mu) + \mathrm{ltrm}(\sigma, \overline{\mu}) \, , \\ \mathrm{ltrm}(\sigma, \mu) &\leq \mathrm{ltrm}(\sigma_\mu) \, , \textit{ and} \\ \mathrm{ltrm}(\sigma) &\leq \mathrm{ltrm}(\sigma_\mu) + \mathrm{ltrm}(\sigma, \overline{\mu}) \, . \end{aligned}$$

4.2 Properties of the Perturbation Models

Let us now prove some properties of partial permutations and partial alterations. The lemmas proved in this section are crucial for estimating the smoothed height and the smoothed number of left-to-right maxima under these models.

The following lemma states that the expected height under partial permutations and alterations depends merely on the elements that are left unmarked. The marked elements contribute at most $O(\log n)$ to the height. Thus, when estimating the expected height in the subsequent sections, we can restrict ourselves to considering the elements that are left unmarked.

Lemma 8. *Let σ be a sequence of length n and let $p \in (0, 1)$. Let $\mu \subseteq [n]$ be a random set of marked positions and σ' be the random sequence obtained from σ via p-partial permutation or p-partial alteration. Then*

$$\mathbb{E}(\mathrm{height}(\sigma')) \leq \mathbb{E}\left(\mathrm{height}(\sigma', \overline{\mu})\right) + O(\log n) \, .$$

Note that $\mathbb{E}(\mathrm{height}(\sigma')) = \mathbb{E}(\text{height-perm}_p(\sigma))$ in case of partial permutations and $\mathbb{E}(\mathrm{height}(\sigma')) = \mathbb{E}(\text{height-alter}_p(\sigma))$ in case of partial alterations.

Proof. In case of partial permutations, the elements of σ_μ are randomly permuted, while in case of partial alterations, they are drawn independently at random. In either case, $\mathbb{E}(\mathrm{height}(\sigma'_\mu)) \in O(\log n)$. The lemma follows from Lemma 6. □

Again we obtain an equivalent lemma for left-to-right maxima.

Lemma 9. *Under the assumptions of Lemma 8, we have*

$$\begin{aligned} \mathbb{E}(\text{ltrm-perm}_p(\sigma)) &\leq \mathbb{E}\left(\mathrm{ltrm}(\sigma', \overline{\mu})\right) + O(\log n) \textit{ and} \\ \mathbb{E}(\text{ltrm-alter}_p(\sigma)) &\leq \mathbb{E}\left(\mathrm{ltrm}(\sigma', \overline{\mu})\right) + O(\log n) \, . \end{aligned}$$

The following lemma gives an upper bound for the probability that no element in a fixed set of elements is among the first elements of the perturbed sequence.

Lemma 10. *Fix $\epsilon > 0$, and let $\alpha = 1 + n^{-\epsilon/8}$. Let $p = p(n)$ be the smoothing parameter with $p \geq n^{\epsilon-1}$ and $p \leq 1 - \epsilon$. Let σ be a sequence of length n, and let σ' be the random sequence obtained from σ by performing a p-partial permutation or a p-partial alteration.*

Let $k, \ell \in \mathbb{N}$ with $k, \ell \geq n^{-\epsilon/8} \cdot \sqrt{n/p}$. Let $A = [x - \frac{1}{2}, x + k - \frac{1}{2}) \subseteq [\frac{1}{2}, n + \frac{1}{2})$ for some $x \in [n]$.

Then

$$\mathbb{P}\big(\sigma'_{[\ell]} \cap A = \emptyset\big) \leq \exp\big(-\frac{k\ell p}{\alpha^3 n}\big) + 6 \cdot \exp\big(-\frac{n^{\epsilon/8}}{16}\big) .$$

In case of partial permutations, we can also choose $A = \{x, x+1, \ldots, x+k-1\}$.

Proof. According to Lemma 2, the probability that $|M_p^n \cap [\ell]| < \alpha^{-1} p\ell$, i.e. that too few of the first ℓ positions are marked, is bounded from above by $2 \cdot \exp(-n^{\epsilon/8}/16)$. In case of partial permutations, we also need to bound the probability of $|\sigma_{M_p^n} \cap A| < \alpha^{-1} pk$, i.e. that too few of the elements of A are marked, and of $|M_p^n| > \alpha pn$, i.e. that too many positions are marked overall. According to Lemma 2, both are bounded from above by $2 \cdot \exp(-n^{\epsilon/8}/16)$. Overall, the probability that any of these three events happens is at most $6 \cdot \exp(-n^{\epsilon/8}/16)$.

From now on, we assume that at least $\alpha^{-1} p\ell$ of the first ℓ positions of σ are marked, at least $\alpha^{-1} pk$ elements in A are marked, and at most αpn positions are marked overall.

The probability that a marked of the first ℓ elements of σ' does not assume a value in A is bounded from above by $\frac{\alpha pn - \alpha^{-1} pk}{\alpha pn} = 1 - \frac{k}{\alpha^2 n}$ in case of partial permutations. In case of partial alterations, the probability is $\frac{n-k}{n} \leq 1 - \frac{k}{\alpha^2 n}$. Thus, the probability that none of the first elements, at least $\alpha^{-1} p\ell$ of which are marked, assumes a value in A is bounded from above by

$$\left(1 - \frac{k}{\alpha^2 n}\right)^{\frac{p\ell}{\alpha}} = \left(\left(1 - \frac{k}{\alpha^2 n}\right)^{\frac{\alpha^2 n}{k}}\right)^{\frac{k\ell p}{\alpha^3 n}} \leq \exp\left(-\frac{k\ell p}{\alpha^3 n}\right) .$$

Overall, the probability that none of the elements of A is among the first elements of σ' is bounded from above by $\exp\big(-\frac{k\ell p}{\alpha^3 n}\big) + 6 \cdot \exp(-n^{\epsilon/8}/16)$ as claimed. □

In the proofs in the subsequent sections, we will exploit Lemma 10 only for $k \in O(\sqrt{n/p} \cdot \operatorname{polylog}(n)) \cap \Omega(\sqrt{n/p}/\operatorname{polylog}(n))$. For such values of k, we can get rid of the term $6 \cdot \exp(-n^{\epsilon/8}/16)$ for sufficiently large values of n.

Lemma 11. *Fix $\epsilon > 0$, $\beta > 1$, and $c \in \mathbb{N}$. Let $p = p(n)$ with $p \geq n^{\epsilon - 1}$ and $p \leq 1 - \epsilon$. Let $\ell = \ell(n) = a \cdot \sqrt{n/p}$ and $k = k(n) = b \cdot \sqrt{n/p}$ for $(\log n)^{-c} \leq a, b \leq (\log n)^c$. Let A be as in Lemma 10.*

Then $\mathbb{P}(\sigma'_{[\ell]} \cap A = \emptyset) \leq \exp(-ab/\beta)$ for all sufficiently large n.

Proof. We have $k\ell = \frac{abn}{p}$. Thus, $\mathbb{P}(\sigma'_{[\ell]} \cap A = \emptyset) \leq \exp(-ab/\alpha^3) + 6 \cdot \exp(-n^{\epsilon/8}/16)$ by Lemma 10.

For sufficiently large n, we have $\beta > \alpha$. Furthermore, $6 \cdot \exp(-n^{\epsilon/8}/16)$ decreases faster than $\exp(-ab/\alpha^3)$ since $a, b \leq (\log n)^c$. Thus, for sufficiently large n, we have $\exp(-ab/\alpha^3) + 6 \cdot \exp(-n^{\epsilon/8}/16) \leq \exp(-ab/\beta)$. □

5 Tight Bounds for Left-to-right Maxima

5.1 Partial Permutations

Theorem 12. *Fix* $\epsilon > 0$*. Let* $p = p(n)$ *with* $n^{\epsilon-1} \leq p \leq 1-\epsilon$*. Then for all sufficiently large* n *and for all sequences* σ *of length* n*,*

$$\mathbb{E}\big(\text{ltrm-perm}_p(\sigma)\big) \leq 3.6 \cdot (1-p) \cdot \sqrt{n/p}\,.$$

Proof. The basic idea for proving this theorem is to estimate the probability that one of the k largest elements of σ is among the first k elements, which would bound the number of left-to-right maxima by $2k$. We get the additional factor of $(1-p)$ since only unmarked elements have to be taken into account.

We will show an upper bound for $\mathbb{E}(\text{ltrm}(\sigma', \overline{\mu}))$. Then we obtain an upper bound for the number of left-to-right maxima by adding $O(\log n)$ according to Lemma 8.

Let σ be a permutation of $[n]$. Let $K_c = \lceil c\sqrt{n/p}\rceil$ for $c \in [\log n]$. Let $\alpha = 1 + n^{-\epsilon/8}$ and fix β with $1 < \beta < 1.001$. We have $\alpha < \beta$ for all sufficiently large n.

We call a partial permutation **partially successful** if at least $\alpha^{-1}pK_c$ of the first K_c positions and of the K_c largest elements are marked for all $c \in [\log n]$ and at most αpn positions are marked overall. According to Lemma 2, the probability that a partial permutation is not partially successful is at most $P = (2 + 4 \cdot \log n) \cdot \exp(-n^{\epsilon/8}/16)$. If a partial permutation is not partially successful, we bound the number of left-to-right maxima by n.

We call σ' ***c*-successful** for $c \in [\log n]$ if the corresponding partial permutation is partially successful and one of the K_c largest elements $n, n-1, \ldots, n-K_c+1$ is among the first K_c elements in σ'.

Assume that σ' is c-successful and that $m \in \{n-K_c+1, \ldots, n\}$ is among the first K_c elements of σ'. The only unmarked elements that can contribute to $\text{ltrm}(\sigma', \overline{\mu})$ are those that are among the first K_c positions and those that are larger than m. All other unmarked elements are smaller than m and located behind m in σ', thus they are no left-to-right maxima. The expected number of unmarked elements larger than $n - K_c$ plus the expected number of unmarked positions among the first K_c positions is at most

$$2 \cdot \big(1 - \frac{p}{\alpha}\big) \cdot K_c \leq 2 \cdot \big(1 - p + p \cdot \frac{n^{-\epsilon/8}}{2}\big) \cdot (c\sqrt{n/p} + 1) = Q_c\,. \tag{1}$$

Hence, $\text{ltrm}(\sigma', \overline{\mu}) \leq Q_c$ if σ' is c-successful.

Let $c \in [\log n]$. The probability that a partially successful partial permutation is not c-successful is at most $\exp(-c^2/\beta)$ for sufficiently large n according to Lemma 11. In particular, the probability that σ' is not $(\log n)$-successful is at most $P' = \exp(-(\log n)^2/\beta)$. If σ' is not $(\log n)$-successful, we bound the number of left-to-right maxima by n.

If we restrict ourselves to partially successful partial permutations, we have

$\mathbb{P}(\text{ltrm}(\sigma', \overline{\mu}) > Q_c) \leq \exp(-c^2/\beta)$. Hence, we can bound $\text{ltrm}(\sigma', \overline{\mu})$ from above by

$$
\begin{aligned}
& \sum_{c=0}^{\log n - 1} Q_{c+1} \cdot \underbrace{\mathbb{P}(\text{no } c\text{-success but } (c+1)\text{-success})}_{\leq \mathbb{P}(\sigma' \text{ is not } c\text{-successful})} + n \cdot (P + P') \\
\leq \quad & 2 \cdot (1-p) \cdot \sqrt{n/p} \cdot \underbrace{\sum_{c \in \mathbb{N}} (c+1) \cdot e^{-\frac{c^2}{\beta}}}_{< 1.8 \text{ for } \beta < 1.001} \\
& \underbrace{+ \log n \cdot (1 - p + p\frac{n^{-\epsilon/8}}{2}) + \sum_{c \in [\log n]} pc\sqrt{n/p} \cdot \frac{n^{-\epsilon/8}}{2} + n \cdot (P + P')}_{\in O((\log n)^2 \cdot \sqrt{pn} \cdot n^{-\epsilon/8}); \text{ these terms are due to Inequality (1)}} \\
\leq \quad & C \cdot (1-p) \cdot \sqrt{n/p}
\end{aligned}
$$

for some $C < 3.6$ and all sufficiently large n. Thus, according to Lemma 8, we have $\mathbb{E}(\text{ltrm-perm}_p(\sigma)) \leq C \cdot (1-p) \cdot \sqrt{n/p} + O(\log n)$, which proves the theorem. $\square$

The following lemma is an improvement of the lower bound proof for the number of left-to-right maxima under partial permutations presented by Banderier et al. [3]. We obtain a lower bound with a much larger constant that holds in particular for all constant $p \in (0,1)$; the lower bound provided by Banderier et al. holds only for $p \leq 1/2$.

Lemma 13. *Fix $\epsilon > 0$, $\beta > 1$, and $c > 0$. Let $p = p(n)$ with $p \geq n^{\epsilon-1}$ and $p \leq 1-\epsilon$. Then for all sufficiently large n, there exist a sequence σ of length n with*

$$\mathbb{P}\big(\text{ltrm-perm}_p(\sigma) \geq c^2 \cdot (1-p) \cdot \sqrt{n/p}\big) \geq \exp(-c^2\beta) .$$

Proof. Fix $c' > c$ and β' with $1 < \beta' < \beta$ such that $c'^2\beta' < c^2\beta$. Thus, $\exp(-c'^2\beta') > \exp(-c^2\beta)$.

Let $K_{c'} = \lfloor c' \cdot \sqrt{n/p} \rfloor$, and let $\sigma = (n - K_{c'} + 1, n - K_{c'} + 2, \ldots, n, 1, 2, \ldots, n - K_{c'})$. We start by a sketch of the proof: The probability that none of the first $K_{c'}$ elements is moved further to the front is bounded from below roughly by $\exp(-c'^2\beta')$. In such a case, all unmarked elements among the first $K_{c'}$ elements are left-to-right maxima, and there are roughly $(1-p) \cdot K_{c'}$ such elements.

Let again $\alpha = 1 + n^{-\epsilon/8}$. Let P be probability that more than $\alpha c'\sqrt{np}$ of the first $K_{c'}$ elements are marked or that less than $\alpha^{-1}pn$ of all elements are marked. We have $P \leq 4 \cdot \exp(-n^{\epsilon/8}/16)$ according to Lemma 2.

Let μ be the set of marked positions and let $\mu_{c'} = \mu \cap [K_{c'}] = \{i_1, \ldots, i_x\}$ be the set of marked positions among the first $K_{c'}$ positions with $i_1 < i_2 < \ldots < i_x$. Let $y = |\mu|$ be the number of all marked positions. Let π be a random permutation of μ. We say that π is **successful** if $\pi(i) > i$ for all $i \in \mu_c$. Thus, under a successful permutation, all marked elements in $\{n - K_c + 1, \ldots, n\}$ are moved further to the back.

If π is successful, then all $K_{c'} - x$ unmarked elements in $\{n - K_{c'} + 1, \ldots, n\}$ are left-to-right maxima. Let us bound the probability from below that the random permutation π of μ is successful for a given μ: For i_x, $y - x$ positions are allowed and

x positions are not allowed; for i_{x-1}, $y-x$ positions are allowed (all in $\mu \setminus \mu_c$ plus one for position i_x minus one for position $\pi(i_x)$) and $x-1$ positions are not allowed; …; for i_1, $y-x$ positions are allowed and one position is not allowed. Thus, the probability that π is successful is at least

$$\left(\frac{y-x}{y}\right)^x = (\underbrace{(1-\frac{x}{y})^{\frac{y}{x}}}_{\geq e^{-1}\cdot(1-\frac{x}{y})})^{\frac{x^2}{y}} \geq \exp\left(\left(\ln\left(1-\frac{x}{y}\right)-1\right)\cdot\frac{x^2}{y}\right) ,$$

Provided that $x \leq \alpha c'\sqrt{np}$ and $y \geq \alpha^{-1}pn$, we obtain a probability that the random permutation is successful of at least

$$\begin{aligned} & \exp\left(\left(\ln\left(1-\frac{\alpha c'\sqrt{np}}{\alpha^{-1}pn}\right)-1\right)\cdot\frac{\alpha^2p^2(c'\sqrt{n/p})^2}{\alpha^{-1}pn}\right) \\ \geq\ & \exp\left(\left(\ln\left(1-\frac{\alpha^2c'}{\sqrt{pn}}\right)-1\right)\cdot\alpha^3c'^2\right) = Q\cdot\exp(-\alpha^3c'^2) \end{aligned}$$

for $Q = \big(1-\frac{\alpha^2c'}{\sqrt{pn}}\big)^{\alpha^3c'^2}$, which tends to one as n increases.

Thus, with a probability of at least $(1-P)\cdot Q\cdot\exp(-\alpha^3c'^2)$, all unmarked elements of $\{K_{c'}+1,\ldots,n\}$ are left-to-right maxima. Furthermore, we have $(1-P)\cdot Q\cdot\exp(-\alpha^3c'^2) \geq \exp(-c^2\beta)$ for sufficiently large n since $(1-P)\cdot Q$ tends to 1, $\alpha^3 < \beta$, and $c^2\beta > c'^2\beta'^2$.

The number of unmarked elements of $\{K_{c'}+1,\ldots,n\}$ is at least $(1-p/\alpha)K_{c'}$, which is bounded from below by $(1-p)\cdot c\cdot\sqrt{n/p}$ for large enough n since α tends to 1 and $c < c'$. □

A consequence of Lemma 13 is that $\exp(-c^2\alpha)\cdot c\cdot(1-p)\cdot\sqrt{n/p}$ for $c > 0$ is a lower bound for the number of left-to-right maxima. The term $\exp(-c^2\alpha)\cdot c$ assumes its maximum for $c = 1/\sqrt{2\alpha}$. By choosing α close to 1 and $c = 1/\sqrt{2\alpha}$, we obtain a lower bound of $0.4\cdot(1-p)\cdot\sqrt{n/p}$ for the expected number of left-to-right maxima under p-partial permutations. We can improve the lower bound by a more careful analysis.

Theorem 14. *Fix $\epsilon > 0$. Let $p = p(n) \in (0,1)$ with $n^{\epsilon-1} \leq p \leq 1-\epsilon$. For all sufficiently large n, there exists a sequence σ of length n with*

$$\mathbb{E}\big(\text{ltrm-perm}_p(\sigma)\big) \geq 0.6\cdot(1-p)\cdot\sqrt{n/p}\,.$$

Proof. We use the same notation as in the proof of Lemma 13. The key observation for improving the lower bound is the following: If none of the marked of the largest K_c elements is among the first γK_c elements of σ' for $\gamma \in [0,1]$, then we have $\gamma\cdot(1-p)\cdot K_c$ left-to-right maxima in expectation. The probability for this is at least $\exp(-c^2\gamma/\beta)$ for any fixed $\beta > 1$ and sufficiently large n.

We consider γ at discrete values in $[0,1]$. Then the expected number of left-to-right maxima after performing a p-partial permutation is at least

$$\sum_{\gamma\in C} 0.01\cdot(1-p)\cdot K_c\cdot\exp\big(-\frac{c^2\cdot\gamma}{\beta}\big)\,.$$

Setting $c = 1.12$ and β sufficiently close to 1 completes the proof. □

Theorem 14 also yields the same lower bound for the tree height since the number of left-to-right maxima of a sequence is a lower bound for the height of the binary search tree obtained from that sequence. We can, however, prove a stronger lower bound for the height of binary search trees (Theorem 22).

Another consequence of Lemma 13 is that there is no constant c such that the number of left-to-right maxima is at most $c \cdot (1-p) \cdot \sqrt{n/p}$ with high probability, i.e. with a probability of at least $1 - n^{-\Omega(1)}$. Thus, the bounds proved for the expected tree height or the number of left-to-right maxima cannot be generalised to bounds that hold with high probability. A bound for the tree height that holds with high probability can be obtained from Lemma 11, as we will show in Theorem 20. Clearly, this bound holds for the number of left-to-right maxima as well.

5.2 Partial Alterations

We obtain the same upper bound for the expected number of left-to-right maxima under partial alterations.

Theorem 15. *Fix $\epsilon > 0$. Let $p = p(n) \in (0,1)$ with $n^{\epsilon-1} \leq p \leq 1-\epsilon$. Then for all sufficiently large n and for all sequences σ of length n,*

$$\mathbb{E}\big(\text{ltrm-alter}_p(\sigma)\big) \leq 3.6 \cdot (1-p) \cdot \sqrt{n/p} \,.$$

Proof. The proof is similar to the proof of Theorem 12. The sequence σ' obtained from σ via p-partial alteration is called c-successful if at least one of the first K_c elements of σ' assumes a value in the interval $[n - K_c + \frac{1}{2}, n + \frac{1}{2})$. If this happens, we can bound $\text{ltrm}(\sigma', \overline{\mu})$ by $\big(1 - \frac{p}{\alpha}\big) \cdot 2K_c$. The probability that we do not have c-success is at most $\exp(-c^2/\beta)$ by Lemma 11. The remainder of the proof proceeds in the same way as the proof of Theorem 12. □

Let us now prove the counterpart for partial alterations of Lemma 13.

Lemma 16. *Fix $\epsilon > 0$, $\beta > 1$, and $c > 0$. Let $p = p(n)$ with $p \geq n^{\epsilon-1}$ and $p \leq 1-\epsilon$. Then for all sufficiently large n, there exist a sequence σ of length n with*

$$\mathbb{P}\big(\text{ltrm-alter}_p(\sigma) \geq c^2 \cdot (1-p) \cdot \sqrt{n/p}\big) \geq \exp(-c^2\beta) \,.$$

Proof. We choose c' and β' as in the proof of Lemma 13. Let again $K_{c'} = \lfloor c' \cdot \sqrt{n/p} \rfloor$ and $\sigma = (n - K_c + 1, n - K_c + 2, \ldots, n, 1, 2, \ldots, n - K_c)$. Let $\alpha = 1 + n^{-\epsilon/8}$, and let $P \leq 2 \cdot \exp(-n^{\epsilon/8}/16)$ be probability that more than $\alpha c' \sqrt{np}$ of the first $K_{c'}$ elements are marked (see Lemma 2).

Let $\mu_{c'}$ be the set of marked positions among the first K_c positions, $x = |\mu_{c'}|$ its cardinality, and $\mu_{c'} = \{i_1, \ldots, i_x\}$. We say that σ' is successful if $\sigma'_{i_j} \leq n - c' \cdot \sqrt{n/p} + \frac{1}{2}$ for all $j \in [x]$. If σ' is successful, then all unmarked elements among the first $K_{c'}$ elements if σ' are left-to-right maxima.

The probability that σ' is successful is at least

$$\left(\frac{n - c' \cdot \sqrt{n/p}}{n}\right)^x = \underbrace{((1 - \frac{c'}{\sqrt{np}})^{\frac{\sqrt{np}}{c'}}}_{\geq e^{-1}\cdot(1-\frac{c'}{\sqrt{np}})})^{\frac{xc'}{\sqrt{np}}}$$

$$\geq \exp\left(\left(\ln\left(1 - \frac{c'}{\sqrt{np}}\right) - 1\right) \cdot \frac{xc'}{\sqrt{np}}\right) = Q \cdot \exp\left(-\frac{xc'}{\sqrt{np}}\right)$$

for $Q = \left(1 - \frac{c'}{\sqrt{np}}\right)^{\frac{xc'}{\sqrt{np}}}$. Provided that $x \leq \alpha c' \sqrt{np}$, Q tends to 1, and we obtain a lower bound for the probability that σ' is successful of $Q \cdot \exp(c'^2\alpha)$. Thus, with a probability of at least $(1 - P) \cdot Q \cdot \exp(-\alpha^3 c'^2)$, all unmarked elements of $\{K_{c'} + 1, \ldots, n\}$ are left-to-right maxima.

The proofs that $Q \cdot (1 - P) \cdot \exp(c'^2\alpha) \geq \exp(-c^2\beta)$ and that the number of unmarked elements of $\{K_{c'} + 1, \ldots, n\}$ is at least $(1 - p) \cdot c \cdot \sqrt{n/p}$ for large enough n follow the same lines as in the proof of Lemma 13. □

From the above lemma, we obtain $0.4 \cdot (1 - p) \cdot \sqrt{n/p}$ as a lower bound for the number of left-to-right maxima. As for partial permutations, this bound is obtained by choosing α close to 1 and $c = 1/\sqrt{2\alpha}$. Again, we can improve the constant in the lower bound. The proof is almost identical to the proof of Theorem 14.

Theorem 17. *For all $p \in (0, 1)$ and all sufficiently large n, there exists a sequence σ of length n with*

$$\text{ltrm-alter}_p(\sigma) \geq 0.6 \cdot (1 - p) \cdot \sqrt{n/p}\,.$$

As for partial permutations, a consequence of Lemma 16 is that we cannot achieve a bound of $O((1-p) \cdot \sqrt{n/p})$ that holds with high probability for the number of left-to-right maxima or the height of binary search trees, but we can show that the height after p-partial alteration is $O(\sqrt{(n/p) \cdot \log n})$ with high probability (Theorem 24).

6 Tight Bounds for Binary Search Trees

6.1 Partial Permutations

Let us now prove one of the main results of this work, namely an upper bound for the expected height of binary search trees obtained from sequences under partial permutations.

Theorem 18. *Let $p \in (0, 1)$. Then for all sufficiently large n and all sequences σ of length n, we have*

$$\mathbb{E}\big(\text{height-perm}_p(\sigma)\big) \leq 6.7 \cdot (1 - p) \cdot \sqrt{n/p}\,.$$

Proof. The idea is to divide the sequence into blocks $B_1, B_2, \ldots$, where B_d is of size $cd^2\sqrt{n/p}$ for some $c > 0$. Each block B_d is further divided into d^4 parts $A_d^1, \ldots, A_d^{d^4}$, each consisting of $cd^{-2}\sqrt{n/p}$ elements. If on every root-to-leaf path in the tree obtained from the perturbed sequence, there are elements of at most two such A_d^i for every d, then the height is at most

$$\sum_{d=1}^{\infty} 2 \cdot \underbrace{cd^{-2}\sqrt{n/p}}_{\text{size of an } A_d^i} = (c\pi^2/3) \cdot \sqrt{n/p}\,.$$

The probability that this does not happen decreases exponentially in c, which will be shown later on. We obtain the upper bound claimed in the theorem mainly by carefully applying this bound and by exploiting the fact that only a fraction of $(1-p)$ of the elements are unmarked. Marked elements contribute at most $O(\log n)$ to the expected height of the tree according to Lemma 8. Thus, it suffices to show

$$\mathbb{E}(\text{height}(\sigma', \overline{\mu})) \leq C \cdot (1-p) \cdot \sqrt{n/p}$$

for some fixed $C < 6.7$, where $\mu \subseteq [n]$ is the random set of marked positions and σ' is the sequence obtained by randomly permuting the elements of σ_μ. Then, for all sufficiently large n,

$$\text{height-perm}_p(\sigma) \leq C \cdot (1-p) \cdot \sqrt{n/p} + O(\log n) \leq 6.7 \cdot (1-p) \cdot \sqrt{n/p}\,.$$

Choose β arbitrarily with $1 < \beta < 1.01$. Let

$$D(d) = \sum_{i=1}^{d-1} i^2 = \frac{(d-1)\cdot(d-1/2)\cdot d}{3}\,.$$

Then $D(d) \geq d^3/8$ for $d \geq 2$.

Let $c \in [\log n]$ and $K_c = c \cdot \sqrt{n/p}$. We divide a prefix of the sequence σ into blocks $B_1, B_2, \ldots, B_{(\log n)^2}$. The block B_d consists of $d^2 K_c$ elements (we deal with the case the K_c is not integral in a moment): B_1 contains the elements of σ at the first K_c positions, B_2 contains the elements of σ at the next $4K_c$ positions, and so on. Let $B = \bigcup_{d=1}^{(\log n)^2} B_d$ be the set of elements that are contained in any B_d. Let $d' = (\log n)^2 + 1$ and $D' = D(d') \geq (\log n)^6/8$. We have $|B| = D' \cdot K_c \geq \frac{1}{8} \cdot (\log n)^6 \cdot K_c$.

Every block B_d is further divided into d^4 subsets $A_d^1, \ldots, A_d^{d^4}$ of elements as follows: A_d^1 contains the K_c/d^2 smallest elements of B_d, A_d^2 contains the K_c/d^2 next smallest elements of B_d, …, and $A_d^{d^4}$ contains the K_c/d^2 largest elements of B_d. Figure 3(a) illustrates the division of σ into blocks $B_1, B_2, \ldots, B_{(\log n)^2}$ and subsets A_d^i for $d \in [(\log n)^2]$ and $i \in [d^2]$.

Finally, we divide $[n]$ into $\log n \cdot \sqrt{np}$ subsets $C_1, \ldots, C_{\log n \cdot \sqrt{np}}$ with $C_j = \{\frac{\sqrt{n/p}}{\log n} \cdot (j-1)+1, \ldots, \frac{\sqrt{n/p}}{\log n} \cdot j\}$. Thus, C_1 contains the $(\log n)^{-1} \cdot \sqrt{n/p}$ smallest numbers of $[n]$, C_2 contains the $(\log n)^{-1} \cdot \sqrt{n/p}$ next smallest numbers of $[n]$, …, and $C_{\log n \cdot \sqrt{np}}$ contains the $(\log n)^{-1} \cdot \sqrt{n/p}$ largest elements of $[n]$.

If K_c, the sizes K_c/d^2 of the subsets A_d^i, or the size $\sqrt{n/p}/\log n$ of the subsets C_j are not integral, we have to replace them by $\lceil K_c \rceil$, $\lceil K_c/d^2 \rceil$, and $\lceil \sqrt{n/p}/\log n \rceil$, respectively. We allow an overlap of one of two subsets A_d^i and A_d^{i+1} as well as of two subsets C_j and C_{j+1}. Augmenting the sizes of the subsets only increases the success probabilities in the following (though only marginally). Since there are only poly-logarithmically many subsets A_d^i, this would increase the tree height only by an additional $O(\operatorname{polylog} n)$. Furthermore, every root-to-leaf path contains only elements of at most two sets C_j. Thus, replacing $\sqrt{n/p}/\log n$ by $\lceil \sqrt{n/p}/\log n \rceil$ would increase the tree height by at most two. The overall increase due to rounding up is $O(\operatorname{polylog} n)$, which is negligible.

Let again $\alpha = 1 + n^{-\epsilon/8}$. We call a set of k positions or elements **partially successful** in μ and σ' if at least $\alpha^{-1}pk$ and at most αpk positions or elements of this set are marked. We say that μ and σ' are partially successful if the following properties are fulfilled:

- for all $c \in [\log n]$, $d \in [(\log n)^2]$, and $i \in [d^4]$, A_d^i is partially successful in μ and σ', and
- for all $j \in [\log n \sqrt{np}]$, C_j is partially successful in μ and σ'.

The probability that one particular of these sets is not partially successful is at most $\exp(-n^{\epsilon/8}/16)$ according to Lemma 2. Since the number of such sets is polynomially bounded in n, the probability P that at least one of these sets is not partially successful is bounded from above by $\exp(-n^{\epsilon/8}/32)$ for all sufficiently large n.

If μ and σ' are not partially successful, we bound the height of $T(\sigma')$ by n.

From now on, we consider the case that μ and σ' are partially successful. When speaking about partial success, we occasionally do not mention σ' or μ.

We call a subset A_d^i ***c*-successful** if at least one element of A_d^i is permuted to one of the $D(d) \cdot c \cdot \sqrt{n/p}$ positions that precede B_d. Thus, for all $d \in [(\log n)^2]$, $d \geq 2$, and $i \in [d^4]$, we have

$$\mathbb{P}(A_d^i \text{ is not successful}) \leq \exp(-d^{-2}cD(d)c\beta^{-1}) \leq \exp\left(-\frac{c^2 d}{8\beta}\right) \tag{2}$$

according to Lemma 11: There are $d^{-2}c\sqrt{n/p}$ elements in A_d^i, $D(d)c\sqrt{n/p}$ positions that precede B_d, and $D(d) \geq d^3/8$.

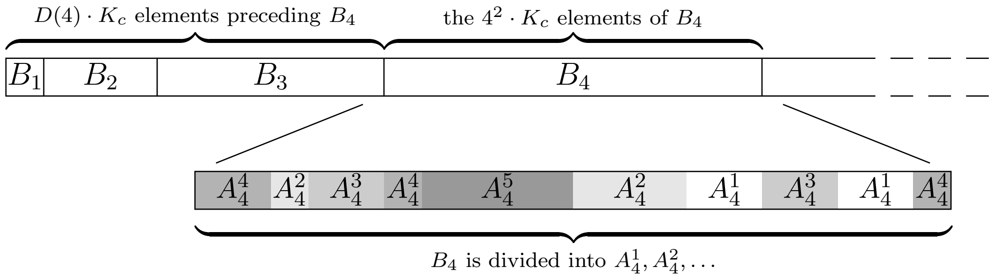
We call a block B_d (for $d \geq 2$) ***c*-successful** if all subsets $A_d^1, \ldots, A_d^{d^4}$ of B_d are c-successful. The probability that B_d is not c-successful is at most $d^4 \cdot \exp(-c^2 d/(8\beta))$ according to Inequality (2) since there are d^4 subsets of B_d. Figure 3 illustrates c-success.

A subset C_j is called ***c*-successful** if at least one element of C_j is among the first $D'c\sqrt{n/p}$ positions of σ'. The probability that a fixed C_j is not c-successful is at most $\exp(-\frac{cD'}{\beta \log n}) \leq \exp(-\frac{c(\log n)^5}{8\beta})$. The probability that any C_j is not c-successful is bounded from above by

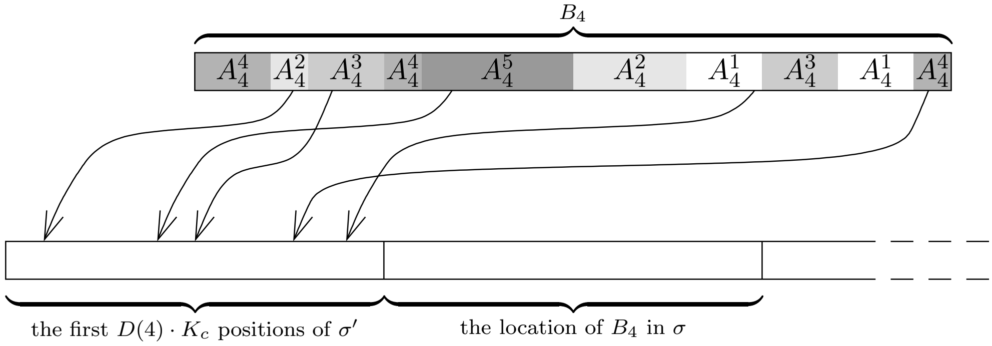
$$\log n \cdot \sqrt{np} \cdot \exp\left(-\frac{c(\log n)^5}{8\beta}\right) \leq d'^4 \cdot \exp\left(-\frac{c^2 d'}{8\beta}\right) \tag{3}$$

for all sufficiently large n. Although the upper bound of $d'^4 \cdot \exp(-c^2 d'/(8\beta))$ is not tight, it suffices for the calculations below.

Finally, we say that **σ' is *c*-successful** if all blocks $B_1, B_2, \ldots, B_{(\log n)^2}$ are c-successful and all subsets $C_1, \ldots, C_{\log n\sqrt{np}}$ are c-successful.



(a) Dividing the first $D' \cdot K_c$ elements of σ into blocks $B_1, \dots, B_{(\log n)^2}$. The subset A_4^1 contains the $K_c/4$ smallest elements of B_4, …, and A_4^{16} contains the $K_c/4$ largest elements of B_4. (For readability, B_4 is divided into only five subsets in the illustration.)



(b) A subset A_4^i is c-successful if at least one element of A_4^i is among the first $D(4) \cdot K_c$ elements of σ'. The block B_4 is c-successful if all A_4^i are c-successful.

Figure 3: The division of σ into blocks and subsets (shown here for B_4).

Let $c \geq 5$. The probability that σ' is not c-successful is at most

$$\begin{aligned} & \sum_{2 \leq d \leq (\log n)^2} d^4 \cdot \exp\bigl(-c^2 d/(8\beta)\bigr) + \mathbb{P}(\text{some } C_j \text{ is not } c\text{-successful}) \\ \leq & \sum_{2 \leq d \leq (\log n)^2 + 1} d^4 \cdot \exp\bigl(-c^2 d/(8\beta)\bigr) \;\leq\; \sum_{d \geq 2} \Bigl(\exp\bigl(-c^2/(16\beta)\bigr)\Bigr)^d \\ = & \; \frac{\exp\bigl(-c^2/(16\beta)\bigr)^2}{1 - \exp\bigl(-c^2/(16\beta)\bigr)} = E(c, \beta)\,. \end{aligned} \tag{4}$$

The first inequality holds due to Inequality (3), the second inequality holds since $c \geq 5$. If σ' is not $(\log n)$-successful, which happens with a probability of at most $E(\log n, \beta)$, we bound the height of $T(\sigma')$ by n.

Let $Q_c = \bigl(c \cdot \frac{\pi^2}{3} + \frac{2}{\log n}\bigr) \cdot (1 - \alpha^{-1} p) \cdot \sqrt{n/p}$.

Lemma 19. *If σ' is c-successful, then* $\mathrm{height}(\sigma', \overline{\mu}) \leq Q_c$.

Proof. Consider the way in which $T(\sigma')$ is built iteratively from σ'. Let $d \geq 2$. After inserting the first $D(d) \cdot K_c$ elements, the partial tree $\tilde{T}$ grown so far contains at least one element of A_d^i for every $i \in [d^4]$. Except for elements of $\tilde{T}$, there cannot be elements from both $A_d^{j^-}$ and $A_d^{j^+}$ for $j^- < i < j^+$ that lie on the same root-to-leaf path of $T(\sigma')$: Let $x \in A_d^i$ be part of $\tilde{T}$, then all elements of $A_d^{j^-}$ that are not part of $\tilde{T}$ are to the left of x in $T(\sigma')$, while all elements of $A_d^{j^+}$ that are not part of $\tilde{T}$ are to the right of x in $T(\sigma')$. Thus, except for elements of $\tilde{T}$, only elements of two consecutive subsets A_d^i and A_d^{i+1} can lie on the same root-to-leaf path of $T(\sigma')$. For every i, there are at most $2 \cdot d^{-2} \cdot K_c$ such elements.

For every d and i, there are at most $(1 - \alpha^{-1} p) \cdot d^{-2} \cdot K_c$ unmarked elements in A_d^i since σ' is partially successful. Thus for every d, at most $2 \cdot (1 - \alpha^{-1} p) \cdot d^{-2} \cdot K_c$ unmarked elements of B_d are on the same root-to-leaf path in $T(\sigma')$.

Let $\overline{B} = [n] \setminus B$ be the set of elements of σ that are not contained in any A_d^i. There cannot be unmarked elements from both $C_{k^-} \cap \overline{B}$ and $C_{k^+} \cap \overline{B}$ for $k^- < j < k^+$ on the same root-to-leaf path in σ' since there is at least one element of C_j among the first $D' \cdot K_c$ elements of σ'. Thus, there are at most $2 \cdot (1 - \alpha^{-1} p) \cdot \frac{\sqrt{n/p}}{\log n}$ unmarked elements of $\overline{B}$ on the same root-to-leaf path in $T(\sigma')$. The maximum number of unmarked elements on any root-to-leaf path in $T(\sigma')$ is thus at most

$$
\begin{aligned}
& \sum_{d=1}^{(\log n)^2} 2 \cdot (1 - \alpha^{-1} p) \cdot c d^{-2} \cdot \sqrt{n/p} + 2 \cdot (1 - \alpha^{-1} p) \cdot (\log n)^{-1} \cdot \sqrt{n/p} \\
\leq \; & \Big(2c \cdot \sum_{d \geq 1} d^{-2} + 2/\log n\Big) \cdot (1 - \alpha^{-1} p) \cdot \sqrt{n/p} = Q_c \, .
\end{aligned}
$$

□

According to Lemma 19 and Formula (4), we have $\mathbb{P}(\text{height}(\sigma', \overline{\mu}) > Q_c) \leq E(c, \alpha)$ for $5 \leq c \leq \log n$ in case of partial success. Hence, we can bound the expectation of $\text{height}(\sigma', \overline{\mu})$ from above by

$$
\begin{aligned}
& Q_5 + \sum_{5 \leq c \leq \log n} Q_{c+1} \cdot \underbrace{\mathbb{P}(\sigma' \text{ is not } c\text{-successful but } (c+1)\text{-successful})}_{\leq \mathbb{P}(\sigma' \text{ is not } c\text{-successful})} \\
& + \underbrace{n \cdot (P + E(\log n, \beta))}_{=X} \\
\leq \; & \underbrace{(1 - \alpha^{-1} p)}_{\leq (1-p) + n^{-\epsilon/8} p} \cdot \sqrt{n/p} \cdot \underbrace{\left(5 + \sum_{c=5}^{\infty} \left(\frac{\pi^2}{3}(c+1) + \frac{2}{\log n}\right) \cdot E(c, \beta)\right)}_{=Y \in O(1)} + X \\
\leq \; & \underbrace{(1 - p) \cdot \sqrt{n/p}}_{=Z} \cdot Y + \underbrace{n^{\frac{1}{2} - \frac{\epsilon}{8}} \cdot \sqrt{p} \cdot Y + X}_{\in o(Z)} \\
= \; & Z \cdot \underbrace{\Big(5 + \frac{\pi^2}{3} \cdot \overbrace{\sum_{c \geq 5} (c+1) \cdot E(c, \beta)}^{< 0.5 \text{ for } \beta < 1.01}\Big)}_{< C \text{ for some } C < 6.7 \text{ and } \beta < 1.01} + o(Z) \leq C \cdot (1 - p) \cdot \sqrt{n/p}
\end{aligned}
$$

for all sufficiently large n and $\beta < 1.01$. The second inequality holds since $\alpha^{-1} \geq 1 - n^{-\epsilon/8}$. The equality holds because $Z \cdot \sum_{c=5}^{\infty} \frac{2E(c,\beta)}{\log n} \in O(Z/\log n) \subseteq o(Z)$. Finally, $\sum_{c \geq 5} (c+1) \cdot E(c, \beta) < 0.5$ for $\beta < 1.01$ can be be shown by adding up the first few terms and bounding the terms for larger c by a geometric series. This completes the proof. □

An upper bound for the height of binary search trees under partial permutation and partial alteration that holds with high probability can be obtained by applying Lemma 11.

Theorem 20. *Fix $\epsilon > 0$ and $\beta > 1$. Let $p = p(n) \in (0, 1)$ with $n^{\epsilon-1} \leq p \leq 1 - \epsilon$, $c > 0$, and let $n \in \mathbb{N}$ be sufficiently large. Let σ be a sequence of length n, and let $c > 0$. Then*

$$\mathbb{P}\left(\text{height-perm}_p(\sigma) > c \cdot \sqrt{(n/p) \cdot \ln n}\right) \leq n^{-(c/3)^2/\beta+0.5} .$$

Proof. Fix $\tilde{c} < c/3$ and $\beta' < \beta$ such that and $\tilde{c}^2/\beta' > (c/3)^2/\beta$. Let $K_{\tilde{c}} = \lceil \tilde{c} \cdot \sqrt{(n/p) \cdot \ln n}\rceil$, which is bounded from above by $c/3 \cdot \sqrt{(n/p) \cdot \ln n}$ for sufficiently large n. Let B_1 be the set of the $K_{\tilde{c}}$ smallest elements of σ, let B_2 be the set of the $K_{\tilde{c}}$ next smallest elements of σ, ..., and let $B_{n/K_{\tilde{c}}}$ be the set of the $K_{\tilde{c}}$ largest elements of σ. (In case that $K_{\tilde{c}}$ is not integral, the sets $B_1, \ldots, B_{n/K_{\tilde{c}}}$ are allowed to overlap by one element.) In the following, let σ' be the random sequence obtained from σ by a p-partial permutation.

Lemma 21. *Assume that for every i, at least one element of B_i is among the first $K_{\tilde{c}}$ elements of σ'. Then* $\text{height}(\sigma') \leq c \cdot \sqrt{(n/p) \cdot \ln n}$.

Proof. Consider the way in which $T(\sigma')$ is built iteratively from σ'. After inserting the first $K_{\tilde{c}}$ elements, the partial tree $\tilde{T}$ grown so far has a height of at most $K_{\tilde{c}}$. The tree $\tilde{T}$ contains at least one element of every B_i. Except for elements of $\tilde{T}$, there cannot be elements from both B_{j^-} and B_{j^+} for $j^- < i < j^+$ that lie on the same root-to-leaf path of $T(\sigma')$: Let $x \in B_i$ be part of $\tilde{T}$, then all elements of B_{j^-} that are not part of $\tilde{T}$ are to the left of x in $T(\sigma')$, while all elements of B_{j^+} that are not part of $\tilde{T}$ are to the right of x in $T(\sigma')$.

It follows that except for elements of $\tilde{T}$, only elements of two consecutive blocks B_i and B_{i+1} can lie on the same root-to-leaf path of $T(\sigma')$. For every i, there are at most $2 \cdot K_{\tilde{c}}$ such elements, yielding a height of at most $2 \cdot K_{\tilde{c}}$. Together with the first $K_{\tilde{c}}$ elements, which build $\tilde{T}$, we obtain $\text{height}(\sigma') \leq 3 \cdot K_{\tilde{c}} \leq c \cdot \sqrt{(n/p) \cdot \ln n}$. □

What remains to be estimated is the probability that there is an i such that no element of B_i is among the first $K_{\tilde{c}}$ elements. For every i, the probability that no element of B_i is among the first $K_{\tilde{c}}$ elements in σ' is at most $\exp(-(\tilde{c}^2/\beta') \cdot \ln n) = n^{-\tilde{c}^2/\beta'}$ by Lemma 11. Thus, the probability that there is any B_i such that no element of B_i is among the first $K_{\tilde{c}}$ elements of σ' is at most

$$(n/K_{\tilde{c}}) \cdot n^{-\tilde{c}^2/\beta'} \leq \tilde{c}^{-1} \cdot \sqrt{p/\ln n} \cdot n^{-\tilde{c}^2/\beta'+0.5} \leq n^{-(c/3)^2/\beta+0.5}$$

for all sufficiently large n. □

From Theorem 20, we immediately obtain that the probability that the height is greater than $3.7 \cdot \sqrt{(n/p) \cdot \ln n}$ is at most $1/n$.

As a counterpart to Theorem 18, we prove the following lower bound. Interestingly, the lower bound is obtained for the sorted sequence, which is not the worst case for the expected number of left-to-right maxima under partial permutation; the expected number of left-to-right maxima of the sequence obtained by partially permuting the sorted sequence is only logarithmic [3].

Theorem 22. *Fix $\epsilon > 0$. Let $p = p(n)$ with $p \geq n^{\epsilon-1}$ and $p \leq 1 - \epsilon$. Then, for all sufficiently large $n \in \mathbb{N}$,*

$$\mathbb{E}\big(\text{height-perm}_p(\sigma^n_{\text{sort}})\big) \geq 0.8 \cdot (1-p) \cdot \sqrt{n/p}\,.$$

Proof. The proof is similar to the proof of Lemma 13, except that we consider the sorted sequence. Fix $\beta < 1.01$, $c' > c$ and β' with $1 < \beta' < \beta$ such that $c'^2\beta' < c^2\beta$. Let again $K_{c'} = \lfloor c' \cdot \sqrt{n/p} \rfloor$ and $\alpha = 1 + n^{\epsilon/8}$.

Let σ' be the sequence obtained from σ^n_{sort} via p-partial permutation. We say that σ' is **c'-successful** if at least $(p/\alpha)K_{c'}$ of the first $K_{c'}$ elements are marked and all of these elements are permuted further to the back. According to the proof of Lemma 13, we have

$$\mathbb{P}(\sigma' \text{ is } c'\text{-successful}) \geq \exp(-c'^2\beta') \geq \exp(-c^2\beta)$$

for sufficiently large n. If σ' is c-successful, then height(σ') is at least the number of unmarked elements among the first $K_{c'}$ elements. There are $(1 - p/\alpha)K_{c'}$ such unmarked elements, which can be bounded from below by $(1-p)K_c$ for all sufficiently large n. Thus, we obtain

$$\mathbb{P}\big(\text{height}(\sigma') \geq (1-p)K_c\big) \geq \exp(-c^2\beta)$$

for all sufficiently large n. We compute a lower bound for the expected height of $T(\sigma')$ by considering c-success at discrete points in $C = \{0.1, 0.2, \ldots, 9.9, 10\}$. To use more values for c does not make much sense since the changes in the result are negligible. Let $Q = (1-p) \cdot \sqrt{n/p}$. We obtain

$$\begin{aligned}
\mathbb{E}(\text{height}(\sigma')) &\geq Q \cdot \sum_{c \in C} c \cdot \mathbb{P}\big(cQ \leq \text{height}(\sigma') < (c+0.1) \cdot Q\big) \\
&\geq Q \cdot \sum_{c \in C} 0.1 \cdot \mathbb{P}(\text{height}(\sigma') \geq cQ) \\
&\geq Q \cdot \underbrace{\sum_{c \in C} 0.1 \cdot \exp(-c^2\beta)}_{\geq 0.8 \text{ for } \beta < 1.01} \;\geq\; 0.8 \cdot Q
\end{aligned}$$

for sufficiently large n and $\beta < 1.01$. □

6.2 Partial Alterations

As for the number of left-to-right maxima, we obtain the same upper bound for the height of binary search trees under partial alterations. The following theorem is obtained via a proof similar to the proof of Theorem 18.

Theorem 23. *Fix $\epsilon > 0$. Let $p = p(n)$ with $p \geq n^{\epsilon-1}$ and $p \leq 1 - \epsilon$. Then for all sufficiently large n and all sequences σ of length n,*

$$\text{height-alter}_p(\sigma) \leq 6.7 \cdot (1-p) \cdot \sqrt{n/p}\,.$$

Proof. The blocks B_d and C_j and the subsets A_d^i are defined in the same way. Now, for each subset A_d^i, we have numbers $a_d^i = \min A_d^i - \frac{1}{2}$ and $b_d^i = \max A_d^i + \frac{1}{2}$. We say that A_d^i is ***c*-successful** if at least one of the first $D(d) \cdot c \cdot \sqrt{n/p}$ elements is from the interval $[a_d^i, b_d^i)$. The term c-successful for blocks B_d is defined in the same way as in the previous proof. For subsets C_j, the term c-successful is defined just as for A_i^d. The remainder of the proof proceeds along the same lines as the proof of Theorem 18. □

We also get the same bound for the height of binary search trees under partial alterations that holds with high probability.

Theorem 24. *Fix $\epsilon > 0$ and $\beta > 1$. Let $p = p(n) \in (0,1)$ with $n^{\epsilon-1} \leq p \leq 1-\epsilon$, $c > 0$, and let $n \in \mathbb{N}$ be sufficiently large. Let σ be a sequence of length n, and let $c > 0$. Then*

$$\mathbb{P}\left(\text{height-alter}_p(\sigma) > c \cdot \sqrt{(n/p) \cdot \ln n}\right) \leq n^{-(c/3)^2/\beta+0.5} .$$

Proof. There are basically two differences to the proof of Theorem 20. First, we have to estimate the probability that for every i, at least one of the first $K_{\tilde{c}}$ elements assumes a value in the interval $[(i-1) \cdot K_{\tilde{c}}, i \cdot K_{\tilde{c}}]$. Second, we have to take the marked elements into account: It might happen that many of the marked elements assume values in the same interval $[(i-1) \cdot K_c, i \cdot K_c]$ for a certain i. Then we cannot argue as in Lemma 19. However, the probability that height of a tree grown from a random permutation is larger than $\delta \cdot \log n$ is at most $n^{\delta \cdot \ln(2e/\delta)-1}$ [13, Lemma 3.1]. We plug in $\delta = \ln n$. Thus, using very coarse estimations, the probability that the height of a tree grown from a random permutation of n elements exceeds $O((\log n)^2)$ is bounded from above by $n^{-\ln n}$ for all sufficiently large n. In particular, the probability that the marked elements contribute more than $O((\log n)^2)$ to the tree height is at most $n^{-\ln n}$.

Fix again $\tilde{c} < c/3$ and $\beta' < \beta$ such that $\tilde{c}^2/\beta' > (c/3)^2/\beta$. Let again $K_{\tilde{c}} = \lceil \tilde{c} \cdot \sqrt{(n/p) \cdot \ln n} \rceil$. Let σ' be the sequence obtained from σ by performing a p-partial alteration. For every i, the probability that no element of the first $K_{\tilde{c}}$ elements of σ' assumes a value in the interval $[(i-1) \cdot K_{\tilde{c}}, i \cdot K_{\tilde{c}}]$ is at most $n^{-\tilde{c}^2/\beta'}$ by Lemma 11. Thus, the probability that there is any such i is at most $n^{-\tilde{c}^2/\beta'+0.5}$ for all sufficiently all large n. Furthermore, the probability that the unmarked elements contribute more than $O((\log n)^2)$ to the height is at most $n^{-\ln n}$ for sufficiently large n. Thus, for all sufficiently large n, the probability that the tree height exceeds $3K_{\tilde{c}} + O((\log n)^2)$ is at most $n^{-\tilde{c}^2/\beta'+0.5} + n^{-\ln n} \leq n^{-(c/3)^2/\beta+0.5}$. □

We obtain the same lower bound for the height of binary search trees under partial alterations. Again, the lower bound is obtained for the sorted sequence. The proof is almost identical to the proof of Theorem 22. The only difference is that we have to use the proof of Lemma 16 instead of Lemma 13.

Theorem 25. *For all $p \in (0,1)$ and all sufficiently large $n \in \mathbb{N}$,*

$$\text{height-alter}_p(\sigma_{\text{sort}}^n) \geq 0.8 \cdot (1-p) \cdot \sqrt{n/p} .$$

7 Comparing Partial Deletions with Partial Permutations and Alterations

For the sake of completeness, let us mention tight bounds for the tree height and the number of left-to-right maxima under partial deletions: For all sequences σ of length n, we have $\text{height-del}_p(\sigma) \leq (1-p) \cdot n$ and $\text{ltrm-del}_p(\sigma) \leq (1-p) \cdot n$. On the other hand, $\text{height-del}_p(\sigma^n_{\text{sort}}) = \text{ltrm-del}_p(\sigma^n_{\text{sort}}) = (1-p) \cdot n$.

Partial deletions turn out to be the worst of the three models: Trees are usually expected to be higher under partial deletions than under partial permutations or alterations, even though they contain fewer elements. The expected height under partial deletions yields upper bounds (up to an additional $O(\log n)$) for the expected height under partial permutations and alterations. Furthermore, we prove that lower bounds for the expected height under partial deletions yield slightly weaker lower bounds for permutations and alterations. The main advantage of partial deletions over partial permutations and partial alterations is that partial deletions are much easier to analyse.

By Lemmas 8 and 9, the expected height and number of left-to-right maxima under partial permutations or alterations can be bounded from above by their counterpart under partial deletions. More precisely: For all sequences σ of length n and for all $p \in [0, 1]$,

$$\begin{aligned}
\mathbb{E}\big(\text{height-perm}_p(\sigma)\big) &\leq \mathbb{E}\big(\text{height-del}_p(\sigma)\big) + O(\log n) \ , \\
\mathbb{E}\big(\text{ltrm-perm}_p(\sigma)\big) &\leq \mathbb{E}\big(\text{ltrm-del}_p(\sigma)\big) + O(\log n) \ , \\
\mathbb{E}\big(\text{height-alter}_p(\sigma)\big) &\leq \mathbb{E}\big(\text{height-del}_p(\sigma)\big) + O(\log n) \ , \text{ and} \\
\mathbb{E}\big(\text{ltrm-alter}_p(\sigma)\big) &\leq \mathbb{E}\big(\text{ltrm-del}_p(\sigma)\big) + O(\log n) \ .
\end{aligned}$$

The converse is not true; this follows from the upper bounds for the height of binary search trees under partial permutations and partial alterations (Theorems 18 and 23) and the lower bound under partial deletions. But we can find a bound for the expected height under partial deletions by the expected height under partial permutations or alterations by padding the sequences considered.

Lemma 26. *Let $p \in (0, 1)$ be fixed, and let σ be a sequence of length n with* $\text{height}(\sigma) = d$ *and* $\mathbb{E}(\text{height-del}_p(\sigma)) = d'$.

Then there exists a sequence $\tilde{\sigma}$ of length $O(n^2)$ with $\text{height}(\tilde{\sigma}) = d + O(\log n)$, $\mathbb{E}(\text{height-perm}_p(\tilde{\sigma})) \in \Omega(d')$, *and* $\mathbb{E}(\text{height-alter}_p(\tilde{\sigma})) \in \Omega(d')$.

Proof. We assume that σ is a permutation of $[n]$. The idea is to attach a tail of sufficiently many elements greater than n to the sequence such that all marked elements that are greater than or equal to n will be permuted to this tail. Thus, the overall structure of the remaining elements from $[n]$ will be as if a partial deletion has been carried out.

Choose $K = n^2 p$ and construct $\tilde{\sigma}$ from σ as follows: the first n items of $\tilde{\sigma}$ are just σ; we call this the **head** of $\tilde{\sigma}$. The last $K - n$ items of $\tilde{\sigma}$, which we call the **tail** of $\tilde{\sigma}$, are numbers greater than n such that these numbers build a tree of height $O(\log(K - n)) = O(\log n)$. With a constant probability, say, c, all elements marked in the head are permuted into the tail (see the proof of Lemma 13).

Consider the tree obtained from the first n elements after partial permutation under the assumption that all marked head elements are now in the tail. This tree is almost identical to the tree obtained from σ via partial deletion when the same elements are marked. The difference is that the tree contains some elements greater than n, which only increase the length of the right-most path. Thus, height-perm$_p(\tilde{\sigma})$ is at least cd', which proves the lemma.

The result for partial alterations follows in the same way. We only have to use the proof of Lemma 16 instead of Lemma 13. □

8 The (In-)Stability of Perturbations

Having shown that worst-case instances become much better when smoothed, we now provide a family of best-case instances for which smoothing results in an exponential increase in height.

We consider the following family of sequences:

- $\sigma^{(1)} = (1)$.
- $\sigma^{(k+1)} = (2^k, \sigma^{(k)}, 2^k + \sigma^{(k)})$, where $c + \sigma = (c + \sigma_1, \ldots, c + \sigma_n)$ for a sequence σ of length n.

For instance, $\sigma^{(2)} = (2, 1, 3)$ and $\sigma^{(3)} = (4, 2, 1, 3, 6, 5, 7)$. Let $n = 2^k - 1$. Then $\sigma^{(k)}$ contains the numbers $1, 2, \ldots, n$, and we have $\text{height}(\sigma^{(k)}) = \text{ltrm}(\sigma^{(k)}) = k = \log(n + 1)$.

Let us estimate the expected number of left-to-right maxima after partial deletion, thus obtaining a lower bound for the expected height of the binary search tree. Deleting the first element of $\sigma^{(k)}$ roughly doubles the number of left-to-right maxima in the resulting sequence. This is the basic idea behind the following theorem; the idea is illustrated in Figure 4.

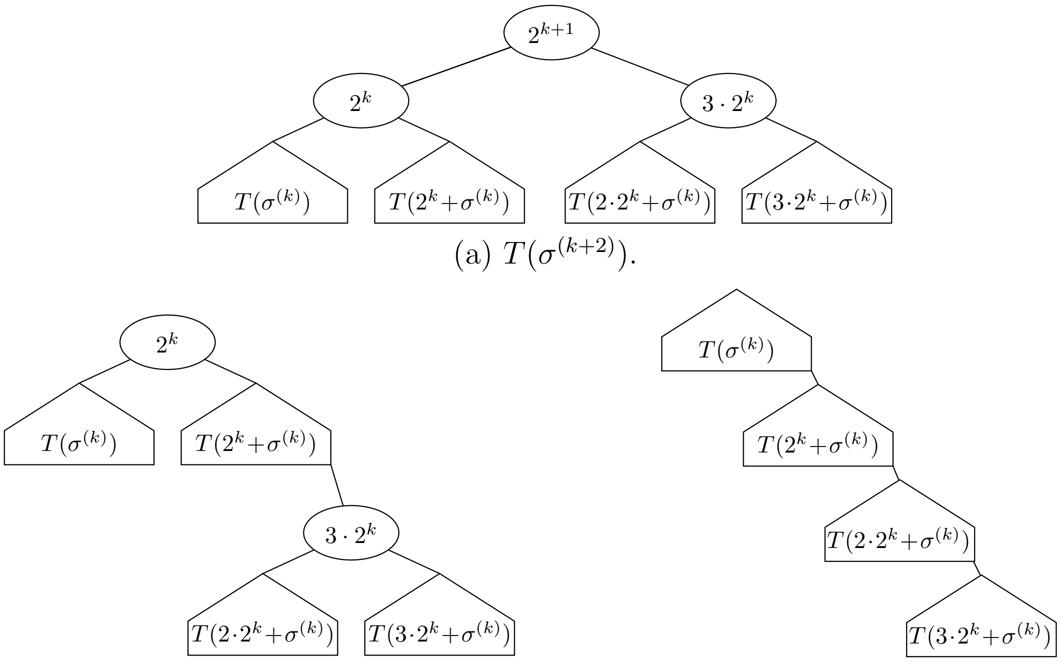
Theorem 27. *Let $p \in (0, 1)$. Then for all $k \in \mathbb{N}$,*

$$\mathbb{E}\big(\text{ltrm-del}_p(\sigma^{(k)})\big) = \frac{1-p}{p} \cdot \big((1+p)^k - 1\big) .$$

Proof. Let $\ell_k = \mathbb{E}(\text{ltrm-del}_p(\sigma^{(k)}))$ for short. The root 2^{k-1} is deleted with probability p. Then the expected number of left-to-right maxima is just the expectation for the left subtree plus the expectation for the right subtree since all elements in the left subtree are smaller and occur earlier than all elements in the right subtree. Both expectations are ℓ_{k-1}. If the root is not deleted, we expect $1 + \ell_{k-1}$ left-to-right maxima: One is the root and ℓ_{k-1} are expected in the right subtree. The left subtree does not contribute any other maxima since all elements in the left subtree are smaller than the root. We have $\ell_1 = 1 - p$ since the single element will be deleted with probability p. Overall,

$$\begin{aligned} \ell_k &= p \cdot 2\ell_{k-1} + (1-p) \cdot (1 + \ell_{k-1}) \\ &= (1+p) \cdot \ell_{k-1} + (1-p) = (1-p) \cdot \sum_{i=0}^{k-1} (1+p)^i = \frac{1-p}{p} \cdot \big((1+p)^k - 1\big) . \end{aligned}$$

□



(a) $T(\sigma^{(k+2)})$.

(b) Removing the root 2^{k+1} roughly doubles the height.

(c) Removing also the roots 2^k of $T(\sigma^{(k+1)})$ and $3 \cdot 2^k$ of $T(2^{k+1}+\sigma^{(k+1)})$ increases the height by a factor of four.

Figure 4: Removing root elements increases the height and the number of left-to-right maxima.

Corollary 28. *For all $p \in (0,1)$ and all $k \in \mathbb{N}$,*

$$\mathbb{E}\big(\text{height-del}_p(\sigma^{(k)})\big) \geq \frac{1-p}{p} \cdot \big((1+p)^k - 1\big) .$$

We conclude that there are some best-case instances that are quite fragile under partial deletions: From logarithmic height they "jump" via smoothing to a height of $\Omega(n^{\log(1+p)})$. (We have $\frac{1-p}{p} \cdot ((1+p)^k - 1) \in \Theta(n^{\log(1+p)})$ for fixed $p \in (0,1)$.) Thus, the height increases exponentially.

We can transfer this result to partial permutations and partial alterations due to Lemma 26. Therefore, we consider sequences $\tilde{\sigma}^{(k)}$, which are obtained from $\sigma^{(k)}$ as described in the proof of Lemma 26.

Corollary 29. *Let $p \in (0,1)$ be fixed. Then* $\text{height}(\tilde{\sigma}^{(k)}) \in O(\log n)$ *and there exists a constant $\epsilon > 0$ with*

$$\mathbb{E}\big(\text{height-perm}_p(\tilde{\sigma}^{(k)})\big) \in \Omega(n^\epsilon) \quad \textit{and} \quad \mathbb{E}\big(\text{height-alter}_p(\tilde{\sigma}^{(k)})\big) \in \Omega(n^\epsilon) .$$

For the sake of completeness, let us mention that the number of left-to-right-maxima is maximally fragile, at least asymptotically for any fixed p: There are sequences with one left-to-right maximum for which the expected number of left-to-right maxima after partial permutation is $\Omega(\sqrt{n})$. The same holds for partial alterations. For partial deletions, the number can jump from 1 to $\Omega(n)$. The proofs are straightforward: Take an adversarial sequence of length $n-1$ for proving lower bounds for the expected number of left-to-right maxima under any of these perturbation models and add an n at the front of the sequence. For partial permutations, this n will be marked and moved behind the first $\Theta(\sqrt{n/p})$ elements with constant probability. For the other two models, the proof is similar.

9 Conclusions

We have analysed the height of binary search trees obtained from perturbed sequences and obtained asymptotically tight lower and upper bounds of roughly $\Theta(\sqrt{n})$ for the height under partial permutations and alterations. This stands in contrast to both the worst-case and the average-case height of n and $\Theta(\log n)$, respectively. One direction for future work is of course improving the constants of the bounds. Another direction is generalising the results to the case that p decreases faster than $n^{\epsilon-1}$.

Interestingly, the sorted sequence σ^n_{sort} turns out to be the worst-case for the smoothed height of binary search trees in the sense that the lower bounds are obtained for σ^n_{sort} (Theorems 22 and 25). This is in contrast to the fact that the expected number of left-to-right maxima of σ^n_{sort} under p-partial permutations is roughly $O(\log n)$ [3]. We believe that for the height of binary search trees, σ^n_{sort} is indeed the worst case.

Conjecture 30. *For all $p \in [0,1]$, all $n \in \mathbb{N}$, and every sequence σ of length n,*

$$\begin{aligned} \mathbb{E}\big(\text{height-perm}_p(\sigma)\big) &\leq \mathbb{E}\big(\text{height-perm}_p(\sigma^n_{\text{sort}})\big) \quad \textit{and} \\ \mathbb{E}\big(\text{height-alter}_p(\sigma)\big) &\leq \mathbb{E}\big(\text{height-alter}_p(\sigma^n_{\text{sort}})\big)\ . \end{aligned}$$

We performed experiments to estimate the constants in the bounds for the height of binary search trees. For all $n \in \{20\,000, 40\,000, \ldots, 500\,000\}$ and $p \in \{0.1, 0.25\}$, we performed 5 000 partial permutations of σ^n_{sort}. We did the same thing for $n \in \{100\,000, 500\,000\}$ and $p \in \{0.05, 0.10, \ldots, 0.95\}$. The results led to the following conjecture. Proving this conjecture would immediately improve our lower bound. Provided that Conjecture 30 holds as well, we would also obtain an improved upper bound for the height of binary search trees under partial permutations.

Conjecture 31. *For $p \in (0,1)$ and some constant $\gamma \approx 1.8$,*

$$\mathbb{E}\big(\text{height-perm}_p(\sigma^n_{\text{sort}})\big) = (\gamma + o(1)) \cdot (1-p) \cdot \sqrt{n/p}\ .$$

Throughout this paper, the bounds obtained for partial permutations and partial alterations are equal. Moreover, the proofs used to obtain these bounds are almost identical. We suspect that this is always true for binary search trees.

Conjecture 32. *For all $p \in [0,1]$ and σ,*

$$\mathbb{E}\big(\text{height-perm}_p(\sigma)\big) \approx \mathbb{E}\big(\text{height-alter}_p(\sigma)\big)\ .$$

In addition to partial permutations and alterations, one could consider other perturbation models for sequences. From a more abstract point of view, a future research direction would be to characterise the properties of perturbation models that lead to upper or lower bounds that are asymptotically different from the average or worst case.

Apart from lower and upper bounds, we have also examined the stability of perturbations, i.e. how much higher a tree can become if the underlying sequence is perturbed. It turns out that all three perturbation models are unstable.

Finally, we are interested in generalising these results to other problems based on permutations, like sorting algorithms (Quicksort under partial permutations has already been investigated by Banderier et al. [3]), routing algorithms, and other algorithms and data structures. Hopefully, this will shed some light on the discrepancy between the worst-case and average-case complexity of these problems.

Acknowledgements

We thank Jan Arpe and Martin Böhme for valuable discussions and comments and the anonymous referees for their helpful suggestions.

References

[1] Alfred V. Aho, John E. Hopcroft, and Jeffrey D. Ullman. *The Design and Analysis of Computer Algorithms*. Addison-Wesley, 1974.

[2] Alfred V. Aho, John E. Hopcroft, and Jeffrey D. Ullman. *Data Structures and Algorithms*. Addison-Wesley, 1983.

[3] Cyril Banderier, René Beier, and Kurt Mehlhorn. Smoothed analysis of three combinatorial problems. In Branislav Rovan and Peter Vojtás, editors, *Proc. of the 28th Int. Symp. on Mathematical Foundations of Computer Science (MFCS)*, volume 2747 of *Lecture Notes in Computer Science*, pages 198–207. Springer, 2003.

[4] Vikas Bansal, Friedhelm Meyer auf der Heide, and Christian Sohler. Labeling smart dust. In Susanne Albers and Tomasz Radzik, editors, *Proc. of the 12th Ann. European Symp. on Algorithms (ESA)*, volume 3221 of *Lecture Notes in Computer Science*, pages 77–88. Springer, 2004.

[5] Luca Becchetti, Stefano Leonardi, Alberto Marchetti-Spaccamela, Guido Schäfer, and Tjark Vredeveld. Average case and smoothed competitive analysis of the multilevel feedback algorithm. *Mathematics of Operations Research*, 31(1):85–108, 2006.

[6] René Beier and Berthold Vöcking. Typical properties of winners and losers in discrete optimization. *SIAM Journal on Computing*, 35(4):855–881, 2006.

[7] Avrim Blum and Joel Spencer. Coloring random and semi-random k-colorable graphs. *Journal of Algorithms*, 19(2):204–234, 1995.

[8] Avrim L. Blum and John D. Dunagan. Smoothed analysis of the perceptron algorithm for linear programming. In *Proc. of the 13th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 905–914. SIAM, 2002.

[9] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms.* MIT Press, 2nd edition, 2001.

[10] Valentina Damerow, Friedhelm Meyer auf der Heide, Harald Räcke, Christian Scheideler, and Christian Sohler. Smoothed motion complexity. In Giuseppe Di Battista and Uri Zwick, editors, *Proc. of the 11th Ann. European Symp. on Algorithms (ESA)*, volume 2832 of *Lecture Notes in Computer Science*, pages 161–171. Springer, 2003.

[11] Valentina Damerow and Christian Sohler. Extreme points under random noise. In Susanne Albers and Tomasz Radzik, editors, *Proc. of the 12th Ann. European Symp. on Algorithms (ESA)*, volume 3221 of *Lecture Notes in Computer Science*, pages 264–274. Springer, 2004.

[12] Amit Deshpande and Daniel A. Spielman. Improved smoothed analysis of the shadow vertex simplex method. In *Proc. of the 46th Ann. IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 349–356. IEEE Computer Society, 2005.

[13] Luc Devroye. A note on the height of binary search trees. *Journal of the ACM*, 33(3):489–498, 1986.

[14] Luc Devroye and Bruce Reed. On the variance of the height of random binary search trees. *SIAM Journal on Computing*, 24(6):1157–1162, 1995.

[15] Michael Drmota. An analytic approach to the height of binary search trees. *Algorithmica*, 29(1–2):89–119, 2001.

[16] Michael Drmota. An analytic approach to the height of binary search trees II. *Journal of the ACM*, 50(3):333–374, 2003.

[17] Uriel Feige and Joe Kilian. Heuristics for semirandom graph problems. *Journal of Computer and System Sciences*, 63(4):639–671, 2001.

[18] Abraham D. Flaxman and Alan M. Frieze. The diameter of randomly perturbed digraphs and some applications. In Klaus Jansen, Sanjeev Khanna, José D. P. Rolim, and Dana Ron, editors, *Proc. of the 8th Int. Workshop on Randomization and Computation (RANDOM)*, volume 3122 of *Lecture Notes in Computer Science*, pages 345–356. Springer, 2004.

[19] Donald E. Knuth. *Sorting and Searching*, volume 3 of *The Art of Computer Programming*. Addison-Wesley, 2nd edition, 1998.

[20] Bodo Manthey and Rüdiger Reischuk. Smoothed analysis of binary search trees. In Xiaotie Deng and Dingzhu Du, editors, *Proc. of the 16th Ann. Int. Symp. on Algorithms and Computation (ISAAC)*, volume 3827 of *Lecture Notes in Computer Science*, pages 483–492. Springer, 2005.

[21] Boris Pittel. On growing random binary trees. *Journal of Mathematical Analysis and Applications*, 103(2):461–480, 1984.

[22] Bruce Reed. The height of a random binary search tree. *Journal of the ACM*, 50(3):306–332, 2003.

[23] John Michael Robson. The height of binary search trees. *The Australian Computer Journal*, 11(4):151–153, 1979.

[24] John Michael Robson. The asymptotic behaviour of the height of binary search trees. Technical Report TR-CS-81-15, The Australian National University, Department of Computer Science, Canberra, 1981.

[25] John Michael Robson. On the concentration of the height of binary search trees. In Pierpaolo Degano, Roberto Gorrieri, and Alberto Marchetti-Spaccamela, editors, *Proc. of the 24th Int. Coll. on Automata, Languages and Programming (ICALP)*, volume 1256 of *Lecture Notes in Computer Science*, pages 441–448. Springer, 1997.

[26] John Michael Robson. Constant bounds on the moments of the height of binary search trees. *Theoretical Computer Science*, 276(1–2):435–444, 2002.

[27] Heiko Röglin and Berthold Vöcking. Smoothed analysis of integer programming. In Michael Jünger and Volker Kaibel, editors, *Proc. of the 11th Int. Conf. on Integer Programming and Combinatorial Optimization (IPCO)*, volume 3509 of *Lecture Notes in Computer Science*, pages 276–290. Springer, 2005.

[28] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *Journal of Computer and System Sciences*, 33(1):75–87, 1986.

[29] Guido Schäfer and Naveen Sivadasan. Topology matters: Smoothed competitiveness of metrical task systems. *Theoretical Computer Science*, 241(1–3):216–246, 2005.

[30] Raimund Seidel and Cecilia R. Aragon. Randomized search trees. *Algorithmica*, 16(4–5):464–497, 1996.

[31] Daniel A. Spielman. The smoothed analysis of algorithms. In Maciej Liśkiewicz and Rüdiger Reischuk, editors, *Proc. of the 15th Int. Symp. on Fundamentals of Computation Theory (FCT)*, volume 3623 of *Lecture Notes in Computer Science*, pages 17–18. Springer, 2005.

[32] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis: Motivation and discrete models. In Frank Dehne, Jörg-Rüdiger Sack, and Michiel Smid, editors, *Proc. of the 8th Workshop on Algorithms and Data Structures (WADS)*, volume 2748 of *Lecture Notes in Computer Science*, pages 256–270. Springer, 2003.

[33] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of termination of linear programming algorithms. *Mathematical Programming, Series B*, 97(1–2):375–404, 2003.

[34] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.

[35] Vijay V. Vazirani. *Approximation Algorithms*. Springer, 2001.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/