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

Smoothed Analysis of the Height of Binary Search Trees

Bodo Manthey* Rüdiger Reischuk

Universität zu Lübeck
Institut für Theoretische Informatik
Ratzeburger Allee 160, 23538 Lübeck, Germany
manthey/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 $\Theta(\sqrt{n})$ for the expected height of binary search trees under partial permutations and partial alterations. That means 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.

ACM Computing Classification: E.1 [Data Structures]: Trees; F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems—sorting and searching; G.2.2 [Discrete Mathematics] Combinatorics—permutations and combinations.

*Supported by DFG research grant Re 672/3.

ISSN 1433-8092

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* [28, 31]. 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 neighbourhood of (worst case) instances.

Smoothed analysis takes into account that on the one hand a *typical* instance is not necessarily a *random* instance 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 (typically the uniform distribution). 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 x and a parameter p. The distribution $\Delta(x, p)$ should favour instances in the vicinity of x. That means, $\Delta(x, p)$ 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 d decreases exponentially in d. Thus, such probability distributions favour instances in the neighbourhood of the adversarial instance. For discrete 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.

Smoothed complexity can be interpreted as follows: If the smoothed complexity of an algorithm is low, then we must be unlucky to accidentally hit an instance on which our algorithm behaves poorly, even if the worst case complexity of our algorithm is bad. In this situation, worst case instances are isolated events.

While the smoothed complexity of continuous problems seems to be well understood, there are only few results on smoothed analysis of discrete problems. In this paper, we are concerned with smoothed analysis of an ordering problem: we examine the *smoothed height of binary search trees.*

Binary search trees are one of the most fundamental data structures and thus a building block 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 equals 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 randomly be perturbed and then the height of the binary search tree generated by the sequence thus obtained 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 byproduct, we also 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. [4]. Thus, the number of left-to-right maxima of a sequence is simply the length of the right-most path in the binary search tree grown from that sequence.

In smoothed analysis one analyses how fragile worst case instances are. We suggest to examine 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 yielding trees of logarithmic depth, but slightly perturbing the sequences yields trees of polynomial depth.

1.1 Previous 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 [26], in which an adversary adaptively chooses a sequence of bits and each is corrupted independently with some fixed probability. Their semi-random model inspired work on semi-random graphs [7, 16], which can be viewed as a forerunner of smoothed analysis of discrete problems.

Spielman and Teng introduced smoothed analysis as a hybrid of average case and worst case complexity [28, 31]. They showed that the simplex algorithm for linear programming with the shadow vertex pivot rule has polynomial smoothed complexity. That means, 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, e.g., several variants of linear programming [8, 30], properties of moving objects [10], online and other algorithms [5, 27], property testing [29], discrete optimisation [6, 25], graph theory [17], and computational geometry [11].

Banderier, Beier, and Mehlhorn [4] applied the concept of smoothed analysis to combinatorial problems. In particular, they analysed the number of left-to-right maxima of a sequence, which is the number of maxima seen when scanning a sequence from left to right. Here the worst case is the sequence $1, 2, \ldots, n$, which yields n left-to-right maxima. On average we expect $H_n = \sum_{i=1}^{n} i^{-1} \approx \ln n$ left-to-right maxima. The perturbation model used by Banderier et al. are *partial permutations*, where each element of the sequence is independently selected with a given probability $p \in [0, 1]$ 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 expected $O(\sqrt{(n/p)\log n})$ for $0 < p < 1$. On the other hand, they showed a lower bound of $\Omega(\sqrt{n/p})$ that holds for $1 < p \leq 1/2$.

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 2.3).

The study of binary search trees is one of the most fundamental problems in computer science since they are the building block for a large variety of data structures (see e.g. Aho et al. [1, 2] and Knuth [18]). Moreover, the height of $T(\sigma)$ is just the number of levels of recursion required by Quicksort if the first element of the sequence to be sorted is chosen as the pivot (see e.g. Cormen et al. [9]).

The worst case height of a binary search tree obtained in this way is obviously n: just take $\sigma = (1, 2, \ldots, n)$. (In this paper, the length of a path is the number of vertices and not the number of edges it contains.) 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. We briefly review some results. Let the random variable $H(n)$ denote the height of a binary search tree obtained from a random permutation. Robson [21] 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 [22]. Pittel [19] proved the existence of a $\gamma > 0$ with $\gamma = \lim_{n\to\infty} \frac{\mathbb{E}H(n)}{\ln(n)}$. Devroye [12] then proved that $\lim_{n\to\infty} \frac{\mathbb{E}H(n)}{\ln(n)} = \alpha$ with $\alpha \approx 4.31107$ being the larger root of

$\alpha \ln(2e/\alpha) = 1$. The variance of $H(n)$ was shown to be $O((\operatorname{llog} n)^2)$ by Devroye and Reed [13] and by Drmota [14]. Robson [23] proved that the expectation of the absolute value of the difference between the height of two random trees is constant. Thus, the height of the random trees is concentrated around the mean. A climax was the result discovered independently by Reed [20] and Drmota [15] that the variance of $H(n)$ actually is $O(1)$. Furthermore, Reed [20] 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 [24] 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 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 consider the height of binary search trees subject to slight perturbations (*smoothed height*), i.e. the expected height under limited randomness. The height of a binary search tree obtained from a sequence of elements only depends on the ordering of the elements. Thus, one should use a perturbation model, which in turn defines the "neighbourhood", that slightly perturbs 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. [4], rearrange some elements, i.e. randomly permute a small subset of the elements of the sequence. 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 some elements. *Partial deletions* remove some of the elements of the sequence without replacement. Thus, they shorten the input, but turn out to be useful for analysing the other two models.

For all three models, we prove matching lower and upper bounds for the expected height of binary search trees obtained from sequences that have been perturbed by one of the perturbation models. More precisely: for all $p \in (0,1)$ and all sequences of length n, the height of a binary search tree obtained via p-partial permutation is expected to be at most $6.7 \cdot (1-p) \cdot \sqrt{n/p}$ for sufficiently large n.

On the other hand, the height of a binary search tree obtained from the sorted sequence via p-partial permutation is at least $0.8 \cdot (1-p) \cdot \sqrt{n/p}$ in expectation. This matches the upper bound up to a constant factor.

For the number of left-to-right maxima under partial permutations or partial alterations, 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 $0.4 \cdot (1-p) \cdot \sqrt{n/p}$.

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

For partial deletions, we obtain $(1-p) \cdot n$ both as lower and upper bound. This result is straight-forward. The main reason for considering partial deletions is that we can bound the expected height subject to partial alterations and permutation by the expected height subject to partial deletions. The converse holds as well, we only have to blow up the sequences quadratically. We exploit this when considering the stability of the permutation models: we prove that partial deletions and thus partial permutations and partial alterations as well are quite unstable, i.e. can cause best case instances to become much worse. More precisely: there are sequences of length n that yield trees of depth $O(\log n)$, but the expected height of the tree obtained after smoothing is $\Omega(n^\epsilon)$ for some $\epsilon > 0$ that depends only on p.

1.3 Outline

In the next section, we introduce some basic notation. We define the perturbation models partial permutations, partial alterations, and partial deletions in Section 3. Then we show some basic properties of binary search trees (Section 4.1), partial permutations (Section 4.2), and partial alterations (Section 4.3). In Section 5 we show matching lower and upper bounds for the expected number of left-to-right maxima under perturbation. After that, we consider the smoothed height of binary search trees under partial permutations and partial alterations (Section 6). We prove matching lower and upper bounds for the expected height of binary search trees that hold for both perturbation models. Then we compare partial deletions with the two other models (Section 7). These results are exploited in Section 8, where we consider the stability of the perturbation models. Finally, we give some concluding remarks (Section 9).

2 Preliminaries

2.1 Notations

We denote by log and ln the logarithm to base 2 and e, respectively, while exp denotes the exponential function to base e. We abbreviate the twice iterated logarithm $\log \circ \log$ by llog. For any $x \in \mathbb{R}$, let $[x] = \{x - i \mid i \in \mathbb{N}, x - i > 0\}$. For instance, $[n] = \{1, 2, \ldots, n\}$ and $[n - \frac{1}{2}] = \{\frac{1}{2}, \frac{3}{2}, \ldots, n - \frac{1}{2}\}$ for $n \in \mathbb{N}$.

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}. When considering partial alterations, we have $\sigma^n_{\text{sort}} = (0.5, 1.5, \ldots, n - 0.5)$ instead (this will be clear from the context).

Let $\tau = (\tau_1, \ldots, \tau_t)$. We call τ a **subsequence of $\boldsymbol{\sigma}$** if there are numbers $i_1 < i_2 < \ldots < i_t$ 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. By abusing notation, we sometimes consider σ_μ as the set of elements at positions in μ, i.e. in this case $\sigma_\mu = \{\sigma_i \mid i \in \mu\}$. However, whether we consider σ_μ as a sequence or as a set will always be clear from the context. For $\mu \subseteq [n]$, we define $\overline{\mu} = [n] \setminus \mu$.

2.2 Probability Theory

We denote probabilities by $\mathbb{P}$ and expectations by $\mathbb{E}$. To bound large deviations, we will frequently use Chernoff bounds [3, Corollary A.7]. Let $p \in (0, 1)$ and let $X_1, X_2, \ldots, X_n$ be mutually independent random variables with $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p$ and $X = \sum_{i=1}^{n} X_i$. Clearly, $\mathbb{E}(X) = pn$. The probability that X deviates by more than a from its expectation is bounded from above by

$$\mathbb{P}(|X - p \cdot n| > a) < 2 \cdot \exp\left(-\frac{2a^2}{n}\right) . \tag{2.1}$$

We will frequently use the following lemma.

Lemma 2.1. *Let $k \in \mathbb{N}$, $\alpha > 1$ and $p \in [0, 1]$. Assume that we have mutually independent random variables $X_1, \ldots, X_k$ as above. Then*

$$\mathbb{P}\left((X > \alpha pk) \vee (X < \alpha^{-1} pk)\right) \leq 2 \cdot \exp\left(-2(1 - \alpha^{-1})^2 p^2 k\right) .$$

Proof. Since $\alpha - 1 \geq 1 - \alpha^{-1}$ for all $\alpha > 1$, let $a = (1 - \alpha^{-1})pk$. Then we apply Formula 2.1 and get

$$\begin{aligned} \mathbb{P}\left((X > \alpha pk) \vee (X < \alpha^{-1} pk)\right) &\leq \mathbb{P}\left(|X - pk| > (1 - \alpha^{-1})pk\right) \\ &< 2 \cdot \exp\left(-\frac{2(1 - \alpha^{-1})p^2k^2}{k}\right) \\ &= 2 \cdot \exp\left(-2(1 - \alpha^{-1})^2 p^2 k\right) . \end{aligned}$$

$\square$

2.3 Binary Search Trees and Left-to-right Maxima

Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ be a sequence. We obtain a **binary search tree $\boldsymbol{T(\sigma)}$** 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.

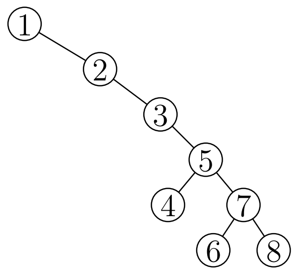


Figure 1: The binary search tree $T(\sigma)$ obtained from $\sigma = (1, 2, 3, 5, 7, 4, 6, 8)$. We have $\text{height}(\sigma) = 6$.

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

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

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($\boldsymbol{\sigma}$)** 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 just the length of the right-most path in the tree $T(\sigma)$.

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* itself.

Partial permutations implement the first possibility: 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 at random. Then the chosen elements are 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. Consider a sequence σ of length n and $p \in [0, 1]$. Every element of σ is marked independently of the others with probability p. To be more formally: the random variable 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 positions marked. If $i \in \mu$, then we say that position i and element σ_i are marked. Thus, σ_μ is the set (or sequence) of all marked elements.

We denote by **height-perm$_p(\sigma)$**, **height-alter$_p(\sigma)$**, and **height-del$_p(\sigma)$** the expected height of the binary search tree $T(\sigma')$ originated from the sequence σ' obtained by performing a p-partial permutation, alteration, and deletion, respectively, on σ (all three models will formally be defined in the following). Analogously, we denote by **ltrm-perm$_p(\sigma)$**, **ltrm-alter$_p(\sigma)$**, and **ltrm-del$_p(\sigma)$** the expected 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** has been introduced by Banderier et al. [4]. Given a random subset M_p^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 variable 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 $i \notin \mu$.

Thus, a p-partial permutation $\Pi(\sigma, M_p^n)$ of σ consists of two steps: randomly mark elements of σ as described above, i.e. randomly create a set $\mu = M_p^n \subseteq [n]$ of marked elements, and then randomly permute all the marked elements, i.e. the elements at positions in μ. Note that $i \in \mu$ does not necessarily mean that σ_i is at a position different from i in $\Pi(\sigma, \mu)$; the random permutation can of course map $\pi(i) = i$.

Example 3.1. *Figure 2 shows an example.*

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

Let us show that partial permutation are indeed a suitable perturbation model by proving that the distribution of $\Pi(\sigma, M_p^n)$ favours sequences close to σ. Therefore, we firstly have to introduce a metric on our sequences. Let σ and τ be two sequences of length n. Without loss of generality, we assume that both are permutations of $[n]$. Otherwise, we replace the jth smallest element of either sequence by j for $j \in [n]$. We define the distance $d(\sigma, \tau)$ between σ and τ as

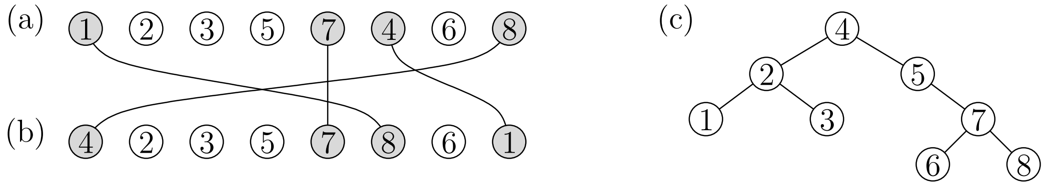


Figure 2: A partial permutation. (a) The sequence $\sigma = (1, 2, 3, 5, 7, 4, 6, 8)$ (Figure 1 shows $T(\sigma)$). The first, fifth, sixth, and eighth element is (randomly) marked, thus $\mu = M_p^n = \{1, 5, 6, 8\}$. (b) The marked elements are randomly permuted. The result is the sequence $\sigma' = \Pi(\sigma, \mu)$, in this case $\sigma' = (4, 2, 3, 5, 7, 8, 6, 1)$. (c) $T(\sigma')$ with $\text{height}(\sigma') = 4$.

$d(\sigma, \tau) = |\{i \mid \sigma_i \neq \tau_i\}|$, 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 $\Pi(\sigma, M_p^n) = \tau$ for some fixed τ depends only on $d(\sigma, \tau)$.

Lemma 3.2. *Let $p \in (0, 1)$, σ and τ be a permutations of $[n]$ with $d = d(\sigma, \tau)$. Then*

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

Proof. All d positions where σ and τ differ must be marked. This happens with probability p^d. The probability that $k - d$ ($k \geq d$) of the remaining positions are marked is $\binom{n-d}{k-d} \cdot p^{k-d} \cdot (1-p)^{n-k}$. Thus, the probability that k positions are marked, d of which are where σ and τ differ is $\binom{n-d}{k-d} \cdot p^k \cdot (1-p)^{n-k}$.

If k positions are marked overall, the probability that the "right" permutation is chosen is $1/k!$, which completes the proof. ☐

Let $\mathbb{P}_d = \sum_{k=d}^{n} p^k \cdot (1-p)^{n-k} \cdot \binom{n-d}{k-d} \cdot \frac{1}{k!}$ 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 fast to zero with increasing d. Thus, the distribution of $\Pi(\sigma, M_p^n)$ is highly concentrated around σ.

3.2 Partial Alterations

Let us now introduce ***p*-partial alterations**. For this perturbation model, we restrict the sequences of length n to be a permutation of $[n - \frac{1}{2}]$ (see Section 2.1).

Every element in M_p^n is replaced by a real number drawn uniformly and independently at random from $[0, n)$ to obtain a sequence σ'. With probability one, all elements in σ' are distinct.

Instead of considering permutations of $[n - \frac{1}{2}]$, we can also consider permutations of $[n]$ and draw the random values from $[\frac{1}{2}, n + \frac{1}{2})$. This does not change the results. Another possibility is to consider permutations of $[n]$ and draw the random values from $[0, n+1)$. This does not change the results by much. However, for technical reasons we consider partial alterations as introduced above.

Example 3.3. *Let* $\sigma = (0.5, 1.5, 2.5, 4.5, 6.5, 3.5, 5.5, 7.5)$ *(which is the sequence of Example 3.1 with* 0.5 *subtracted from each element) and* $\mu = \{1, 5, 6, 8\}$*. By replacing the marked elements with random numbers, we may obtain the sequence* $(3.96\ldots, 1.5, 2.5, 4.5, 7.22\ldots, 7.95\ldots, 5.5, 0.67\ldots)$*.*

Like partial permutations, partial alterations interpolate between worst case ($p = 0$) and average case ($p = 1$). Partial alterations are somewhat easier to analyse: the majority of the results on the average case height of binary search trees (see for instance Pittel [19] and Devroye [12]) is obtained not via random permutations but the binary search trees are grown from a sequence of n random variables that are uniformly and independently drawn from $[0, 1)$. There is no difference between partial permutations and partial alterations for $p = 1$. This seems to hold for all p; the bounds obtained for partial permutations and partial alterations are equal for all p.

3.3 Partial Deletions

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

Example 3.4. *The sequence* σ *and the marking* μ *as in Example 3.1 yield the sequence* $(2, 3, 5, 6)$*.*

Partial deletions are not really perturbing 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 obtained for partial permutations then carry over to partial permutations and partial alterations since the expected height with respect to these three models is closely related (Section 7).

4 Basic Properties

In this section, we state some basic properties of binary search trees (Section 4.1), partial permutations (Section 4.2), and partial alterations (Section 4.3) that we will exploit in subsequent sections.

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 σ be a sequence of length n. The **μ-restricted height of $T(\sigma)$**, denoted by **$\text{height}(\sigma, \mu)$**, is the maximum number of elements of σ_μ on a root-to-leaf path in $T(\sigma)$.

Lemma 4.1. *For all sequences σ of length n and $\mu \subseteq [n]$, we have*

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

Proof. Consider any path of maximum length from the root to a leaf in $T(\sigma)$. This path consists of at most $\text{height}(\sigma, \mu)$ elements of σ_μ and at most $\text{height}(\sigma, \overline{\mu})$ 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 path from the root to a leaf in $T(\sigma_\mu)$. Assume that $a < b$. Then there exists a c prior to a and b in the sequence σ_μ with $a < c < b$. Thus, a and b do not lie on the same root-to-leaf path in the tree $T(\sigma)$ as well. Consider now any root-to-leaf path of $T(\sigma)$ with $\text{height}(\sigma, \mu)$ elements of σ_μ. Then all these elements from σ_μ lie on the same root-to-leaf path in $T(\sigma_\mu)$, which proves the second part of the lemma. □

Of course we have $\text{height}(\sigma, \mu) \leq \text{height}(\sigma)$ for all σ and μ. But $\text{height}(\sigma_\mu) \leq \text{height}(\sigma)$, which would imply $\text{height-del}_p(\sigma) \leq \text{height}(\sigma)$, does not hold in general: Consider $\sigma = (c, a, b, d, e)$ (we use letters and their alphabetical ordering instead of numbers for readability) and $\mu = \{2, 3, 4, 5\}$, then $\sigma_\mu = (a, b, d, e)$. Thus, $\text{height}(\sigma) = 3$ and $\text{height}(\sigma_\mu) = 4$. This will further be investigated in Section 8, when we consider the stability of the perturbation models.

For bounding the smoothed height from above, we will use the following lemma, which is an immediate consequence of Lemma 4.1.

Lemma 4.2. *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 $\text{height}(\sigma) \leq \text{height}(\sigma, \mu) + \text{height}(\sigma, \overline{\mu}) \leq \text{height}(\sigma_\mu) + \text{height}(\sigma, \overline{\mu})$ according to Lemma 4.1. □

For left-to-right maxima, we can state equivalent lemmas. Let σ be a sequence of length n and $\mu \subseteq [n]$. Then **$\text{ltrm}(\sigma, \mu)$** denotes the **μ-restricted number of left-to-right maxima** of σ, i.e. the number of elements σ_i for $i \in \mu$ such that $\sigma_i > \sigma_j$ for all $j \in [i - 1]$. We omit the proof of the following lemma since it is almost equal to the proofs of the lemmas above.

Lemma 4.3. *Let σ be a sequence of length n and $\mu \subseteq [n]$. Then*

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

4.2 Properties of Partial Permutations

Let us now prove some properties of partial permutations. The two lemmas proved in this section are crucial for estimating the smoothed height under partial permutations. In the next section, we prove counterparts of these lemmas for partial alterations that will play a similar role in estimating the height under partial alterations.

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

Lemma 4.4. *Let σ be a sequence of length n and $p \in (0,1)$. Let $\mu \subseteq [n]$ be the random set of marked positions and $\sigma' = \Pi(\sigma,\mu)$ be the random sequence obtained from σ via p-partial permutation. Then*

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

Proof. We have $\mathrm{height}(\sigma_\mu) \in O(\log n)$ since the elements at positions in μ are randomly permuted. Then the lemma follows from Lemma 4.2. □

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

Lemma 4.5. *Under the assumptions of Lemma 4.4, we have*

$$\text{ltrm-perm}_p(\sigma) \leq \mathbb{E}\left(\mathrm{ltrm}(\sigma',\overline{\mu})\right) + O(\log n) \ .$$

The following lemma bounds the probability from above that no element of a fixed set of elements is permuted to a position of a fixed set of positions.

Lemma 4.6. *Let $p \in (0,1)$, $\alpha > 1$, $n \in \mathbb{N}$ be sufficiently large, and σ be a sequence of length n with elements from $[n]$. Let $\sigma' = \Pi(\sigma, M_p^n)$.*

Let $\ell = a\sqrt{n/p}$ and $k = b\sqrt{n/p}$ with $a,b \in \Omega((\mathrm{polylog}\, n)^{-1}) \cap O(\mathrm{polylog}\, n)$. Let $A = \sigma'_{[\ell]}$ be the set of the first ℓ elements of σ'. Let $B \subseteq [n]$ be any subset with $|B| = k$.

Then $\mathbb{P}(A \cap B = \emptyset) \leq \exp(-ab/\alpha)$.

Proof. We choose β with $1 < \beta^3 < \alpha$ arbitrarily. According to Lemma 2.1, the probability P that

- $|M_p^n \cap [\ell]| < \beta^{-1}p\ell$, i.e. that too few of the first ℓ positions are marked,
- $|\sigma_{M_p^n} \cap B| < \beta^{-1}pk$, i.e. that too few of the elements of B are marked, or
- $|M_p^n| > \beta pn$, i.e. that too many positions are marked overall

is $O(\exp(-n^\epsilon))$ for fixed $p \in (0,1)$, $\beta > 1$, and appropriately chosen $\epsilon > 0$. This holds since $a, b \in \Omega((\operatorname{polylog} n)^{-1})$.

From now on, assume that at least $\beta^{-1}p\ell$ of the first ℓ positions of σ are marked, at least $\beta^{-1}pk$ elements in B are marked, and at most βpn positions are marked overall. The probability that then no element from B is in A is at most

$$\begin{aligned} \left(\frac{\beta pn - \beta^{-1}p\ell}{\beta pn}\right)^{\beta^{-1}pk} &= \left(1 - \frac{\ell}{\beta^2 n}\right)^{\beta^{-1}pk} \\ = \left(\left(1 - \frac{\ell}{\beta^2 n}\right)^{\frac{\beta^2 n}{\ell}}\right)^{\frac{\ell}{\beta^2 n}\cdot\beta^{-1}pk} &\leq \exp\left(-\frac{\ell}{\beta^2 n}\cdot\beta^{-1}pk\right) = \exp\left(-\frac{ab}{\beta^3}\right) . \end{aligned}$$

Overall, $\mathbb{P}(A \cap B = \emptyset) \leq \exp(-ab/\beta^3) + P \leq \exp(-ab/\alpha)$ for sufficiently large n since $a, b \in O(\operatorname{polylog} n)$. □

4.3 Properties of Partial Alterations

Partial alterations fulfil roughly the same properties as partial permutations. We state the lemmas and restrict ourselves to pointing out the differences in the proofs.

Lemma 4.7. *Let σ be a sequence of length n with elements from $[n - \frac{1}{2}]$ and $p \in (0,1)$. Let σ' be the random sequence obtained from σ via p-partial alteration and μ be the random set of marked positions. Then*

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

The following lemma is the counterpart of Lemma 4.6 above.

Lemma 4.8. *Let $p \in (0,1)$, $\alpha > 1$, and $n \in \mathbb{N}$ be sufficiently large, and σ be a sequence with elements from $[n - \frac{1}{2}]$. Let σ' be the random sequence obtained from σ by performing a p-partial alteration.*

Let $\ell = a\sqrt{n/p}$ and $k = b\sqrt{n/p}$ with $a, b \in \Omega((\operatorname{polylog} n)^{-1}) \cap O(\operatorname{polylog} n)$. Let $A = \sigma'_{[\ell]}$ and $B = [x, x+k) \subseteq [0, n)$.

Then $\mathbb{P}(A \cap B = \emptyset) \leq \exp(-ab/\alpha)$.

Proof. The proof is similar to the proof of Lemma 4.6. Choose β arbitrarily with $1 < \beta < \alpha$. Assume that at least $\beta^{-1}p\ell$ of the first ℓ positions of σ are marked. Then the probability that no element in A assumes a value of B is at most

$$\left(\frac{n-k}{n}\right)^{\beta^{-1}p\ell} = \left(\left(1-\frac{k}{n}\right)^{\frac{n}{k}}\right)^{ab/\beta} \leq \exp(-ab/\beta) .$$

The remainder of the proof is as in the proof of Lemma 4.6. □

5 Tight Bounds for the Smoothed Number of Left-To-Right Maxima

5.1 Partial Permutations

Theorem 5.1. *Let $p \in (0,1)$. Then for all sufficiently large n and for all sequences σ of length n, we have*

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

Proof. According to Lemma 4.5, it suffices to show

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

for some $C < 3.6$, where $\mu \subseteq [n]$ is the random set of marked positions and σ' is the sequence obtained via randomly permuting the elements of σ_μ. Then $\text{ltrm-perm}_p(\sigma) \leq C(1-p)\sqrt{n/p}+O(\log n) \leq 3.6(1-p)\sqrt{n/p}$. We assume without loss of generality that σ is a permutation of $[n]$.

Let $K_c = c\sqrt{n/p}$ for $c \in [\log n]$. In this and the following proofs, we assume that K_c is a natural number for the sake of readability. If K_c is not a natural number, then we can replace K_c by $\lceil K_c \rceil$. The proofs remain valid.

Choose α with $1 < \alpha < 1.001$. Let P denote the probability that less than $\alpha^{-1}pK_c$ of the first K_c positions are marked or that less than $\alpha^{-1}pK_c$ of the K_c largest elements are marked for some $c \in [\log n]$ or that overall more than αpn elements are marked. P tends exponentially fast to zero as n increases by Lemma 2.1.

From now on, we assume that for all $c \in [\log n]$, at least $\alpha^{-1}pK_c$ of the K_c first positions and of the K_c largest elements are marked. In this case, we say that the partial permutation is **partially successful**. If a partial permutation is not partially successful, we bound the expected number of left-to-right maxima by n.

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

Assume that σ' is c-successful and that $x \in \{n - K_c + 1, \ldots, n\}$ is among the first K_c elements of σ'. Then only the unmarked among the first K_c positions and the unmarked among elements larger than x can contribute to $\mathrm{ltrm}(\sigma', \overline{\mu})$. All other unmarked elements are smaller than x and located after x 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 $2 \cdot (1-p) \cdot K_c = Q_c$. Thus, we have $\mathbb{E}(\mathrm{ltrm}(\sigma', \overline{\mu})) \leq Q_c$ if σ' is c-successful.

The probability that a partially successful partial permutation is not c-successful for $c \in O(\log n)$ is bounded from above by $\exp(-c^2/\alpha)$ according to Lemma 4.6. Particularly, the probability that σ' is not $\log n$-successful is at most $P' = \exp(-(\log n)^2/\alpha)$. If σ' is not $\log n$-successful, we bound the number of left-to-right maxima by n.

Thus, restricted to partially successful partial permutations, we have

$$\mathbb{P}(\text{ltrm-perm}_p(\sigma) > Q_c) \leq \exp(-c^2/\alpha)\,.$$

Hence, we can bound $\mathrm{ltrm}(\sigma', \overline{\mu})$ from above by

$$\begin{aligned}
& \sum_{c=0}^{\log n} Q_{c+1} \cdot \mathbb{P}(\sigma' \text{ is not } c\text{-successful but } (c+1)\text{-successful}) + n \cdot (P + P') \\
\leq\ & 2 \cdot (1-p) \cdot \sqrt{n/p} \cdot \underbrace{\sum_{c \in \mathbb{N}} (c+1) \cdot e^{-\frac{c^2}{\alpha}}}_{< 1.8 \text{ for } \alpha < 1.001} + n \cdot (P + P') \\
\leq\ & C \cdot (1-p) \cdot \sqrt{n/p}
\end{aligned}$$

for some $C < 3.6$, which proves the theorem. □

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. [4]. This way we get a lower bound with a much larger constant that holds for all $p \in (0, 1)$; the lower bound provided by Banderier et al. holds only for $p \leq 1/2$.

Lemma 5.2. *Let $p \in (0,1)$, $\alpha > 1$, and $c > 0$. Then for all sufficiently large n, there exist sequences σ of length n with*

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

Proof. Let $K_c = c\sqrt{n/p}$. Let $\sigma = (n - K_c + 1, n - K_c + 2, \ldots, n, 1, 2, \ldots, n - K_c)$. Choose β arbitrarily with $1 < \beta^3 < \alpha$. Let P denote the probability that more than $\beta p K_c$ of the first K_c elements or less than $\beta^{-1} pn$ of the remaining elements are selected. P tends exponentially fast to zero as n increases (Lemma 2.1).

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

If f is successful, then all $K_c - x$ unmarked elements of $\{n - K_c + 1, \ldots, n\}$ are left-to-right maxima. Provided that at most $\beta p K_c$ of the first K_c elements are marked, i.e. $x \le \beta p K_c$, the expectation of $K_c - x$ is at least $(1 - p)K_c$.

Let us bound the probability from below that the random permutation of μ is successful for a given μ: For i_x, there are y positions allowed and x positions not allowed, for i_{x-1}, there are y positions allowed (all in $\mu \setminus \mu_c$ plus i_x minus $f(i_x)$) and $x - 1$ positions not allowed, …, for i_1, there are y positions allowed and one position not allowed. Thus, the probability that the random permutation is successful is at least

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

Provided that $x \le \beta p K_c$ and $x + y \ge y \ge \beta^{-1} p n$, we obtain that the probability that the random function is successful is at least

$$\begin{aligned} & \exp\left(\left(\ln\left(1 - \frac{\beta p K_c}{\beta^{-1} p n}\right) - 1\right) \cdot \frac{\beta^2 p^2 K_c^2}{\beta^{-1} p n}\right) \\ = \; & \exp\left(\left(\ln\left(1 - \frac{\beta^2 c}{\sqrt{pn}}\right) - 1\right) \cdot \beta^3 c^2\right) = Q \cdot \exp(-\beta^3 c^2) \end{aligned}$$

for $Q = \left(1 - \frac{\beta^2 c}{\sqrt{pn}}\right)^{\beta^3 c^2}$, which tends to one as n increases. Thus, with probability at least $(1-P)\cdot Q \cdot \exp(-\beta^3 c^2)$, all unmarked elements of $\{K_c+1, \ldots, n\}$ are left-to-right maxima. Furthermore, we have $(1 - P) \cdot Q \cdot \exp(-\beta^3 c^2) \ge \exp(-c^2 \alpha)$ for sufficiently large n. Since the expectation of the number of unmarked elements among the first K_c elements is at least $(1 - p)K_c$, the lemma is proved. □

By choosing α sufficiently close to 1 and $c = \sqrt{1/2}$, we immediately get the following theorem from Lemma 5.2.

Theorem 5.3. *For all $p \in (0, 1)$ and all sufficiently large n, there exists a sequence σ of length n with*

$$\text{ltrm-perm}_p(\sigma) \ge 0.4 \cdot (1 - p) \cdot \sqrt{n/p}\,.$$

Theorem 5.3 also yields the same lower bound for height-perm$_p(\sigma)$ since the number of left-to-right maxima of a sequence bounds the height of the binary search tree obtained from that sequence from below. However, for the smoothed height of binary search trees, we can prove a stronger lower bound (Theorem 6.3).

Another consequence of Lemma 5.2 is that there does not exist a 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 probability at least $1 - n^{-\Omega(1)}$. Thus, the bounds proved for the expectation of the tree height or the number of left-to-right maxima cannot be generalised to bounds that hold with high probability. A bound that holds with high probability can directly be obtained from Lemma 4.6: Let σ' be the sequence obtained from σ via p-partial permutation. Then $\text{height}(\sigma') \in O(\sqrt{(n/p) \cdot \log n})$ with probability at least $1 - n^{-\Omega(1)}$. The same holds for $\text{ltrm}(\sigma')$.

5.2 Partial Alterations

As for the height of binary search trees, we obtain the same upper bound for the expected number of left-to-right maxima under partial alterations.

Theorem 5.4. *Let $p \in (0,1)$. Then for all sufficiently large n and for all sequences σ of length n (where σ is a permutation of $[n - \frac{1}{2}]$), we have*

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

Proof. The main difference between the proof of this theorem and the proof of Theorem 5.1 is that we have to use Lemma 4.8 instead of Lemma 4.6.

The sequence σ' obtained from σ via p-partial alteration is called c-successful if there is at least one element of the interval $[n - K_c, n)$ among the first K_c elements of σ'. The remainder of the proof goes the same way as the proof of Theorem 5.1. □

Let us now prove the counterpart for partial alterations of Lemma 5.2.

Lemma 5.5. *Let $p \in (0,1)$, $\alpha > 1$, and $c > 0$. Then for all sufficiently large n, there exist sequences σ of length n with*

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

Proof. Let $K_c = c\sqrt{n/p}$. Let $\sigma = (K_c + 1, K_c + 2, \ldots, n, 1, 2, \ldots, K_c)$. Choose β arbitrarily with $1 < \beta < \alpha$. Let P denote the probability that more than $\beta p K_c$ are marked. P tends exponentially fast to zero as n increases (Lemma 2.1).

Let μ be the set of marked positions and $\mu_c = \mu \cap [K_c]$ be the set of marked positions among the first K_c. Let $\mu_c = \{i_1, \ldots, i_x\}$ with $i_1 < \ldots < i_x$, i.e. $|\mu_c| = x$. We have $\sigma_{i_j} = n - K_c + i_j - \frac{1}{2}$ for all $j \in [x]$. Let σ' be the sequence obtained from σ by replacing all marked elements by random numbers from $[0, n)$. We say that σ' is successful if $\sigma'_{i_j} \leq n - K_c$. If σ' is successful, then all $K_c - x$ unmarked elements among the first K_c elements of σ are left-to-right maxima.

The probability that σ' is successful is at least

$$\left(\frac{n-K_c}{n}\right)^x = (\underbrace{(1-\frac{K_c}{n})^{\frac{n}{K_c}}}_{\geq e^{-1}\cdot(1-\frac{K_c}{n})})^{\frac{xK_c}{n}} \geq \exp\left(\left(\ln\left(1-\frac{K_c}{n}\right)-1\right)\cdot xK_c/n\right) .$$

Provided that $x \leq \beta p K_c$, we obtain that the probability that the random function is successful is at least

$$\begin{aligned} & \exp\left(\left(\ln\left(1-\frac{\beta p K_c}{n}\right)-1\right)\cdot \beta p K_c^2/n\right) \\ = \; & \exp\left(\left(\ln\left(1-\frac{\beta c}{\sqrt{pn}}\right)-1\right)\cdot \beta c^2\right) = Q\cdot\exp(-\beta c^2) \end{aligned}$$

for $Q = \left(1-\frac{\beta c}{\sqrt{pn}}\right)^{\beta c^2}$, which tends to one as n increases. Thus, with probability at least $(1-P)\cdot Q\cdot\exp(-\beta c^2)$, all unmarked among the first K_c elements are left-to-right maxima. The expectation of the number of unmarked elements among the first K_c elements is at least $(1-p)K_c$. Furthermore, for sufficiently large n, we have $(1-P)\cdot Q\cdot\exp(-\beta c^2) \geq \exp(-\alpha c^2)$, which proves the lemma. □

From the above lemma, we obtain the same lower bound for the number of left-to-right maxima as for partial permutations.

Theorem 5.6. *For all $p \in (0,1)$ and all sufficiently large n, there exists a sequence σ of length n with*

$$\text{ltrm-alter}_p(\sigma) \geq 0.4\cdot(1-p)\cdot\sqrt{n/p} .$$

As for partial permutations, a consequence of Lemma 5.5 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 again, we obtain from Lemma 4.8 that for all sequences, the height and the number of left-to-right maxima under partial alterations is in $O(\sqrt{(n/p)\cdot\log n})$ with probability at least $1-n^{-\Omega(1)}$.

6 Tight Bounds for the Smoothed Height of Binary Search Trees

In this section, we consider the smoothed height of binary search trees under the perturbation models partial permutation and partial alteration.

6.1 Partial Permutations

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

Theorem 6.1. *Let $p \in (0,1)$. Then for all sufficiently large n and all sequences σ of length n, we have*

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

Proof. According to Lemma 4.4, 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 via randomly permuting the elements of σ_μ. Then $\text{height-perm}_p(\sigma) \leq C \cdot (1-p) \cdot \sqrt{n/p} + O(\log n) \leq 6.7 \cdot (1-p) \cdot \sqrt{n/p}$ for sufficiently large n.

Choose α arbitrarily with $1 < \alpha < 1.01$. Without loss of generality, we assume that σ is a permutation of $[n]$.

Let $c \in [\log n]$ and $K_c = c\sqrt{n/p}$. We define $D(d) = \sum_{i=1}^{d-1} i^2 = \frac{1}{3} \cdot (d-1) \cdot (d - \frac{1}{2}) \cdot d$. Then $D(d) \geq d^3/8$ for $d \geq 2$.

We divide 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: 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. Thus,

$$B_d = \sigma_{[D(d+1) \cdot K_c]} \setminus \sigma_{[D(d) \cdot K_c]}\,.$$

Let $B = \bigcup_{d=1}^{(\log n)^2} B_d$ be the set of elements that are contained in any B_d. We have $|B| = D((\log n)^2 + 1) \cdot K_c \geq \frac{1}{8} \cdot (\log n)^6 \cdot K_c$.

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 $d^{-2}K_c$ smallest elements of B_d, A_d^2 contains the $d^{-2}K_c$ second smallest elements of B_d, and so on. The subset A_d^1 contains the $d^{-2}K_c$ smallest elements of B_d, A_d^2 the $d^{-2}K_c$ second smallest elements of B_d, $\ldots$, and $A_d^{d^4}$ contains the $d^{-2}K_c$ 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 the numbers in $[n]$ into $\log n \cdot \sqrt{pn}$ subsets $C_1, \ldots, C_{\log n \cdot \sqrt{pn}}$ with

$$C_j = \left\{ \frac{\sqrt{n/p}}{\log n} \cdot (j-1) + 1, \ldots, \frac{\sqrt{n/p}}{\log n} \cdot j \right\}\,.$$

Thus, C_1 contains the $(\log n)^{-1} \cdot \sqrt{n/p}$ smallest numbers of $[n]$, C_2 contains the $(\log n)^{-1} \cdot \sqrt{n/p}$ second smallest numbers of $[n]$, and so on.

Let $\eta = 1 + n^{-1/6}$. We call a set of positions or elements of cardinality k **partially successful** in μ and σ' if at least $\eta^{-1}pk$ and at most ηpk 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
- for all $j \in [\log n\sqrt{pn}]$, C_j is partially successful in μ.

There are only polynomially many sets of elements that must be partially successful, and every such set is of cardinality $\Omega\left(\sqrt{n/p}/\operatorname{polylog} n\right)$. Thus, there exists some $\epsilon > 0$ such that the probability that μ and σ are partially successful is $O(\exp(-n^{\epsilon}))$ according to Lemma 2.1. Let P denote this probability. If μ is not partially successful, we bound the height of σ' by n.

From now on, we assume that μ is partially successful.

We call a subset A_d^i for $d \geq 2$ and $i \in [d^4]$ ***c*-successful** if at least one element of A_d^i is permuted to one of the $D(d)c\sqrt{n/p}$ positions that precede B_d. Thus, the probability that a fixed A_d^i is not successful is at most $\exp(-c^2 D(d) d^{-2} \alpha^{-1}) \leq \exp(-c^2 d/(8\alpha))$ according to Lemma 4.6: there are $d^{-2}c\sqrt{n/p}$ elements in A_d^i and $D(d)c\sqrt{n/p}$ positions that precede B_d.

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 $d^4 \cdot \exp(-c^2 d/(8\alpha))$ since there are d^4 subsets $A_d^1, \ldots, A_d^{d^4}$ of B_d. Figures 3(a) and 3(b) illustrate c-success.

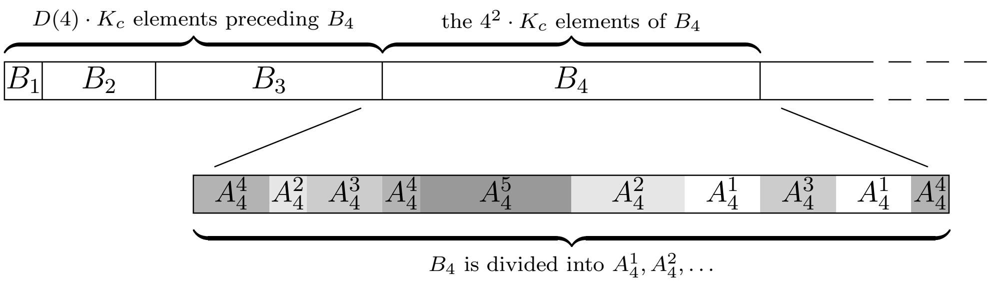
Let $d' = (\log n)^2 + 1$ and $D' = D(d') \geq (\log n)^6/8$. A subset C_j is called ***c*-successful** if at least one element of C_j is among the first $D'c\sqrt{n/p}$ positions of σ'. The probability that a fixed C_j is not c-successful is at most $\exp(-\frac{cD'}{\alpha \log n}) \leq \exp(-\frac{c(\log n)^5}{8\alpha})$. The probability that any C_j is not c-successful is bounded from above by

$$\log n \cdot \sqrt{np} \cdot \exp\left(-\frac{c(\log n)^5}{8\alpha}\right) \leq d'^4 \cdot \exp\left(-\frac{c^2 d'}{8\alpha}\right) \tag{6.1}$$

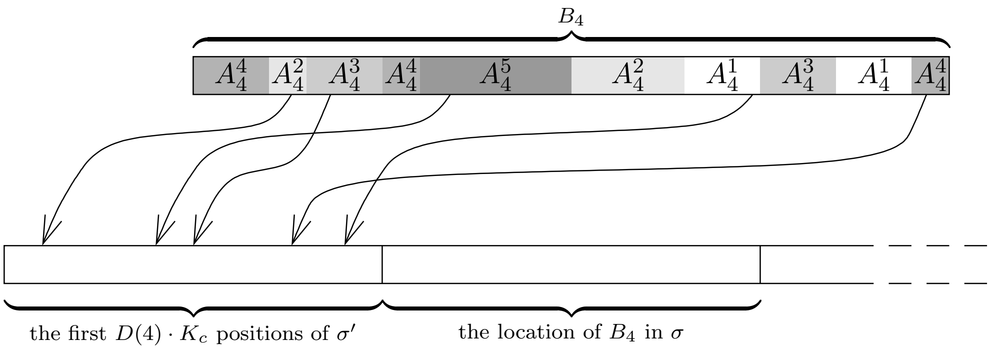
for sufficiently large n.

Finally, we say that $\boldsymbol{\sigma'}$ **is *c*-successful** if

- all blocks $B_1, B_2, \ldots, B_{(\log n)^2}$ are c-successful and
- all subsets $C_1, \ldots, C_{\log n\sqrt{pn}}$ are c-successful.



(a) Dividing the first $D' \cdot K_c$ elements of σ into blocks $B_1, \ldots, B_{(\log n)^2}$. For instance, the block B_4 is further divided into subsets $A_4^1, \ldots, A_4^{16}$, where 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

$$\begin{aligned} & \sum_{d=2}^{(\log n)^2} d^4 \cdot \exp(-c^2 d/(8\alpha)) + \mathbb{P}(\text{some } C_j \text{ is not } c\text{-successful}) \\ \leq \; & \sum_{d=2}^{\infty} d^4 \cdot \exp(-c^2 d/(8\alpha)) \;\leq\; \sum_{d=2}^{\infty} \big(\exp(-c^2/(16\alpha))\big)^d \\ = \; & \frac{\exp(-c^2/(16\alpha))^2}{1 - \exp(-c^2/(16\alpha))} = E(c, \alpha) \,. \end{aligned} \tag{6.2}$$

The first inequality holds due to Formula 6.1, the second inequality holds since $c \geq 5$. If σ' is not $\log n$-successful, which happens with probability at most $E(\log n, \alpha) \leq \exp(-(\log n)^2/(16\alpha))$, we bound the height of $T(\sigma')$ by n.

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

Claim 6.2. *If σ' is c-successful, then* $\operatorname{height}(\sigma', \overline{\mu}) \leq Q_c$.

Proof of Claim 6.2. Consider any path from the root to a leaf in $T(\sigma')$. This path cannot contain unmarked elements from both A_d^{i-1} and A_d^{i+1} for $d \geq 2$ and $2 \leq i \leq d^4 - 1$ since there is at least one element of A_d^i that stands before all unmarked elements of A_d^{i-1} and A_d^{i+1}.

It is possible that unmarked elements from A_d^i and A_d^{i+1} are on the same root-to-leaf path in $T(\sigma')$. For every d and i, there are at most $(1-\eta^{-1}p)cd^{-2}\sqrt{n/p}$ unmarked elements in A_d^i since σ' is partially successful. Thus, for every d, at most $2(1-\eta^{-1}p)cd^{-2}\sqrt{n/p}$ 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 $C_{j-1} \cap \overline{B}$ and $C_{j+1} \cap \overline{B}$ on the same root-to-leaf path in σ' since there is at least one element of C_j among the first $D'c\sqrt{n/p}$ elements of σ'. Thus, there are at most $2(1-\eta^{-1}p)\frac{\sqrt{n/p}}{\log n}$ elements of $\overline{B} \cap \bigcup_{i=1}^{\log n \cdot \sqrt{np}} C_j$ on the same root-to-leaf path in $T(\sigma')$.

Overall, the maximum number of elements on any root-to-leaf path in $T(\sigma')$ can be bounded from above by

$$\begin{aligned}
& \sum_{d=1}^{(\log n)^2} 2 \cdot (1-\eta^{-1}p) \cdot c \cdot d^{-2} \cdot \sqrt{n/p} + 2 \cdot (1-\eta^{-1}p) \cdot (\log n)^{-1} \cdot \sqrt{n/p} \\
\leq \; & \left(2c \cdot \sum_{d=1}^{\infty} \frac{1}{d^2} + \frac{2}{\log n}\right) \cdot (1-\eta^{-1}p) \cdot \sqrt{n/p} \\
= \; & \left(c \cdot \frac{\pi^2}{3} + \frac{2}{\log n}\right) \cdot (1-\eta^{-1}p) \cdot \sqrt{n/p} = Q_c \, ,
\end{aligned}$$

which proves the claim. $\square$

According to Claim 6.2 and Formula 6.2, we have

$$\mathbb{P}\left(\text{height}(\sigma', \overline{\mu}) > Q_c\right) \leq E(c, \alpha)$$

for $5 \leq c \leq \log n$. Furthermore,

$$\eta^{-1} = \frac{1}{1+n^{-1/6}} = 1 - \frac{n^{-1/6}}{1+n^{-1/6}} \geq 1 - n^{-1/6} \, . \tag{6.3}$$

Hence, we can bound the expectation of $\text{height}(\sigma', \overline{\mu})$ from above by

$$
\begin{aligned}
& Q_5 + \sum_{c=5}^{\log n} Q_{c+1} \cdot \mathbb{P}(\sigma' \text{ is not } c\text{-successful but } (c+1)\text{-successful}) \\
& + \underbrace{n \cdot (P + E(\log n, \alpha))}_{=X} \\
\leq\ & \underbrace{(1 - \eta^{-1} p)}_{\leq 1-(1-n^{-1/6})p} \cdot \sqrt{n/p} \cdot \underbrace{\left(5 + \sum_{c=5}^{\infty} \left(\frac{\pi^2}{3}(c+1) + \frac{2}{\log n} \right) \cdot E(c, \alpha) \right)}_{=Y \in O(1)} + X \\
\leq\ & \underbrace{(1-p) \cdot \sqrt{n/p}}_{=Z} \cdot Y + \underbrace{n^{2/6} \cdot \sqrt{p} \cdot Y + X}_{\in o(Z)} \\
=\ & Z \cdot \left(\underbrace{5 + \frac{\pi^2}{3} \cdot \overbrace{\sum_{c \geq 5} (c+1) \cdot E(c, \alpha)}^{< 0.5 \text{ for } \alpha < 1.01}}_{= C < 6.7 \text{ for } \alpha < 1.01} \right) + o(Z) \\
\leq\ & C \cdot (1-p) \sqrt{n/p}
\end{aligned}
$$

for sufficiently large n and $\alpha < 1.01$. The second inequality holds due to Formula 6.3. The first equal sign holds since $Z \cdot \sum_{c=5}^{\infty} \frac{2E(c,\alpha)}{\log n} \in o(Z)$. This completes the proof. ☐

As a counterpart to the above theorem, we prove the following lower bound. Interestingly, the lower bound is obtained for the sorted sequence, which is not a worst case for the expected number of left-to-right maxima; the expected number of left-to-right maxima of the sequence obtained by partially permuting the sorted sequence is only logarithmic [4].

Theorem 6.3. *For all $p \in (0, 1)$ and all sufficiently large $n \in \mathbb{N}$, we have*

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

Proof. Let $c > 0$ be any constant and $K_c = c\sqrt{n/p}$. Let σ' be the sequence obtained from σ^n_{sort} via p-partial permutation. We say that σ' is ***c*-successful** if all marked elements among the first K_c elements of σ^n_{sort} are permuted further to the back. According to Lemma 5.2, we have

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

for arbitrarily chosen $\alpha > 1$ and sufficiently large n. If σ' is c-successful and x elements among the first K_c elements are unmarked, then $\text{height}(\sigma') \geq x$. Let $Q = (1-p) \cdot \sqrt{n/p}$ for short. Analogously to Lemma 5.2, we obtain

$$\mathbb{P}\left(\text{height}(\sigma') \geq cQ\right) \geq \exp(-c^2\alpha)$$

for $\alpha > 1$ and sufficiently large n. The idea is now to consider c-success for all $c \in \{0.1, 0.2, \ldots, 9.9, 10\} = C$. Thus,

$$\begin{aligned}
\mathbb{E}(\text{height}(\sigma')) &\geq Q \cdot \sum_{c \in C} c \cdot \mathbb{P}(cQ \leq \text{height}(\sigma') < (c+1)Q) \\
&\geq Q \cdot \sum_{c \in C} 0.01 \cdot \mathbb{P}(\text{height}(\sigma') \geq cQ) \\
&\geq Q \cdot \sum_{c \in C} 0.01 \cdot \exp(-c^2\alpha) \geq 0.8 \cdot Q
\end{aligned}$$

for sufficiently large n and $\alpha < 1.01$, which proves the theorem. □

6.2 Partial Alterations

The following theorem is obtained via a proof similar to the proof of Theorem 6.1.

Theorem 6.4. *Let $p \in (0, 1)$. Then for all sufficiently large n and all sequences σ of length n (where σ is a permutation of $[n - \frac{1}{2}]$), we have*

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

Proof. The main difference between the proof of this theorem and the proof of Theorem 6.1 is that we have to use Lemma 4.8 instead of Lemma 4.6. The blocks B_d and C_j and the subsets A_d^i are defined in the same way. Now we have for each subset A_i^d numbers $a_d^i = \lfloor \min A_i^d \rfloor$ and $b_d^i = \lceil \max A_i^d \rceil$. We say that A_i^d is ***c*-successful** if among the first $D(d)c\sqrt{n/p}$ elements there is at least one element of 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 like for A_i^d. The remainder of the proof goes along the same lines as the proof of Theorem 6.1. □

We also get the same lower bound for the height of the binary search trees under partial alterations. Again, the lower bound is obtained for the sorted sequence.

Theorem 6.5. *For all $p \in (0, 1)$ and all sufficiently large $n \in \mathbb{N}$, we have*

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

Proof. The proof is almost equal to the proof of Theorem 6.3. The only difference is that we have to use Lemma 5.5 instead of Lemma 5.2. □

7 Comparing Partial Deletions with Permutations and Alterations

In this section, we justify the consideration of partial deletions. 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, although containing less elements. Thus, the expected height under partial deletions yields upper bounds (up to an additional $O(\log n)$ term) for the expected height under partial permutations and alterations. On the other hand, 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.

For the sake of completeness, we start by providing matching upper and lower bounds for the height of binary search trees under partial deletions.

Proposition 7.1. *For all $p \in [0,1]$, $n \in \mathbb{N}$, and sequences σ of length n, we have*

$$\text{height-del}_p(\sigma) \leq (1-p) \cdot n\,.$$

Moreover,

$$\text{height-del}_p(\sigma^n_{\text{sort}}) = (1-p) \cdot n\,.$$

Proof. Let σ' be the sequence obtained from σ via p-partial deletion. Then σ' consists of $(1-p) \cdot n$ elements in expectation. The number of elements is an upper bound for the number of left-to-right maxima.

The second proposition holds obviously. □

The following lemma is an immediate consequence of Lemmas 4.4, 4.5, and 4.7, we therefore omit its proof.

Lemma 7.2. *For all sequences σ of length n and $p \in [0,1]$,*

$$\begin{array}{lll} \text{height-perm}_p(\sigma) & \leq & \text{height-del}_p(\sigma) + O(\log n) \quad \textit{and} \\ \text{ltrm-perm}_p(\sigma) & \leq & \text{ltrm-del}_p(\sigma) + O(\log n)\,. \end{array}$$

If σ is a permutation of $[n - \frac{1}{2}]$, then

$$\begin{array}{lll} \text{height-alter}_p(\sigma) & \leq & \text{height-del}_p(\sigma) + O(\log n) \quad \textit{and} \\ \text{ltrm-alter}_p(\sigma) & \leq & \text{ltrm-del}_p(\sigma) + O(\log n)\,. \end{array}$$

Thus, we can bound the expected height under partial permutations or alterations from above by the expected height under partial deletions. The reverse is not true (Theorems 6.1 and 6.4 and Proposition 7.1). But we can bound the expected height under partial deletions by the expected height under partial permutations or alterations by padding the sequences considered.

Lemma 7.3. *Let $p \in (0,1)$ be fixed and let σ be a sequence of length n with* $\text{height}(\sigma) = d$ *and* $\text{height-del}_p(\sigma) = d'$*. Then there exists a sequence $\tilde{\sigma}$ of length $O(n^2)$ with* $\text{height}(\tilde{\sigma}) = d + O(\log n)$ *and* $\text{height-perm}_p(\tilde{\sigma}) \in \Omega(d')$.

Proof. Without loss of generality, 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 via partial deletions.

Choose $K = n^2p$ and construct $\tilde{\sigma}$ from σ as follows: the first n items of $\tilde{\sigma}$ are just σ, which we call 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 constant probability (see the proof of Lemma 5.2), say c, all elements marked in the head are permuted into the tail.

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

The following is the analogue of the above lemma for partial alterations. Since its proof is similar to the proof of the previous lemma (the mere difference is that we have to use the proof of Lemma 5.5 instead of Lemma 5.2), we omit it.

Lemma 7.4. *Let $p \in (0,1)$ be fixed and σ be a sequence of length n with elements of $[n - \frac{1}{2}]$. Let* $\text{height}(\sigma) = d$ *and* $\text{height-del}_p(\sigma) = d'$*. Then there exists a sequence $\tilde{\sigma}$ of length $O(n^2)$ with* $\text{height}(\tilde{\sigma}) = d + O(\log n)$ *and* $\text{height-alter}_p(\tilde{\sigma}) \geq \Omega(d')$.

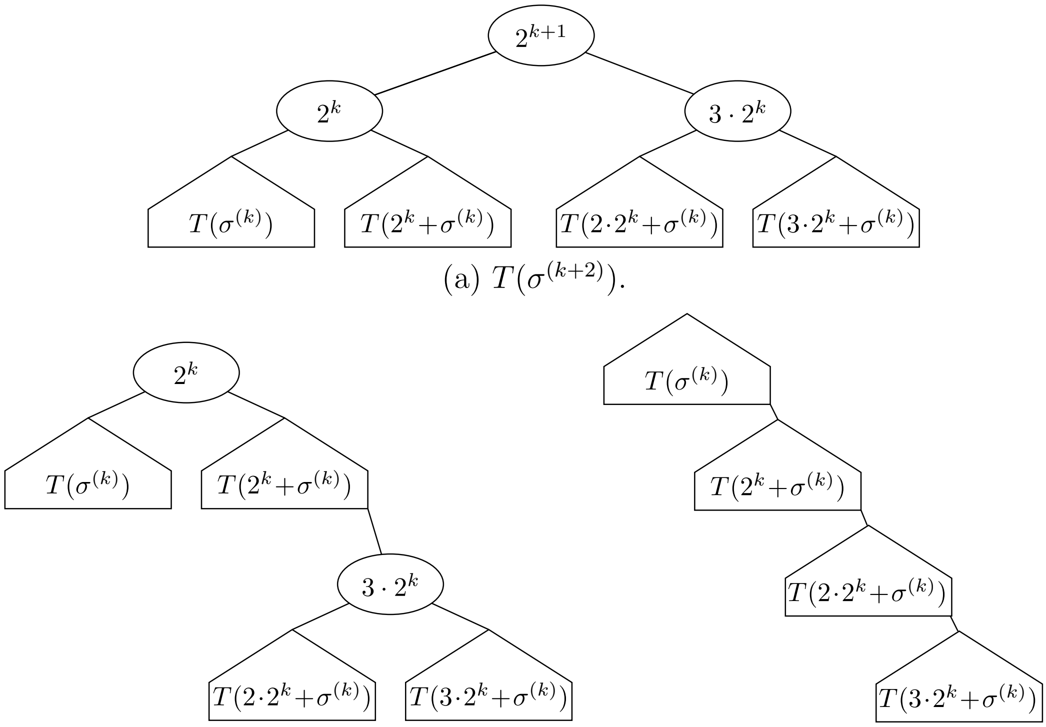
8 The (In-)Stability of Perturbations

Having shown that worst case instances become much better by smoothing, 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 sequence:

- $\sigma^{(1)} = (1)$.
- $\sigma^{(k+1)} = (2^k, \sigma^{(k)}, 2^k + \sigma^{(k)})$, 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 $\text{height}(\sigma^{(k)}) = \text{ltrm}(\sigma^{(k)}) = k \in \Theta(\log n)$.



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

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

(c) Removing further the roots 2^k and $3 \cdot 2^k$ of $T(\sigma^{(k+1)})$ and $T(2^{k+1} + \sigma^{(k+1)})$, respectively, quadruplicates the height.

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

Let us estimate the expected number of left-to-right maxima after partial deletion, bounding the expected height of the binary search tree from below. 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 and illustrated in Figure 4.

Theorem 8.1. *Let $p \in (0,1)$. Then for all $k \in \mathbb{N}$,*

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

Proof. Let $\ell(k) = \text{ltrm-del}_p(\sigma^{(k)})$ 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 $\ell(k-1)$. If the root is not deleted, we expect $1 + \ell(k-1)$ left-to-right maxima: One is the root and $\ell(k-1)$ are expected in the right subtree. We do not get any new maximum from the left subtree since all elements in the left subtree are smaller than the root. We have $\ell(1) = 1 - p$ since with probability p, the single element will be deleted. Overall, we have

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

□

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 the same sequence, we get the following result.

Corollary 8.2. *For all $p \in (0,1)$ and all $k \in \mathbb{N}$,*

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

We conclude that there are some best case instances that are quite fragile under partial deletions: From logarithmic height they "jump" to height $\Omega(n^\epsilon)$, for constant $\epsilon > 0$, via smoothing. Thus, the height increases exponentially.

We can transfer this result to partial permutations and partial alterations due to Lemmas 7.3 and 7.4. Therefore, we consider sequences $\tilde{\sigma}^{(k)}$ which are obtained from $\sigma^{(k)}$ as described in the proof of Lemma 7.3.

Corollary 8.3. *For any fixed $p \in (0,1)$, we have*

$$\begin{array}{lcl} \text{height}(\tilde{\sigma}^{(k)}) & \in & O(\log n) \ , \\ \text{height-perm}_p(\tilde{\sigma}^{(k)}) & \in & \Omega\big(n^\epsilon\big) \ , \textit{ and} \\ \text{height-alter}_p(\tilde{\sigma}^{(k)}) & \in & \Omega\big(n^\epsilon\big) \end{array}$$

for some fixed $\epsilon > 0$.

For the sake of completeness, let us mention that the number of left-to-right-maxima is, at least asymptotically for any fixed p, as fragile as possible: 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 straight-forward: 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. With constant probability, this n will be marked and moved behind the first $\Theta(\sqrt{n/p})$ elements in case of partial permutations. 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 $\sqrt{n}$ 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. Thus, the height of binary search trees under limited randomness differs significantly from both the average and the worst case. One direction for future work is of course improving the constants of the bounds.

Interestingly, the sorted sequence σ^n_{sort} turns out to be a worst case for the smoothed height of binary search trees in the sense that the lower bounds are obtained for σ^n_{sort} (Theorems 6.3 and 6.5). This contrasts the fact that the expected number of left-to-right maxima of σ^n_{sort} under p-partial permutations is roughly $O(\log n)$ [4]. We believe that for the height of binary search trees, σ^n_{sort} is indeed a worst case.

Conjecture 9.1. *For all $p \in [0,1]$, all $n \in \mathbb{N}$, and every sequence σ of length n,*

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

We have 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 have performed $5\,000$ partial permutations of σ^n_{sort}. Furthermore, we did the same for $n \in \{100\,000, 500\,000\}$ and $p \in \{0.05, 0.10, \ldots, 0.95\}$. (See Appendix A for more details.) The results lead to the following conjecture. Proving this conjecture would immediately improve our lower bound. Provided that Conjecture 9.1 holds, this would also yield an improved upper bound for the height of binary search trees under partial permutations.

Conjecture 9.2. *For $p \in (0,1)$ and sufficiently large n,*

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

for some constant $\gamma \approx 1.8$.

Throughout this work, the bounds obtained for partial permutations and partial alterations are equal. Moreover, the proofs used for obtaining these bounds are almost identical. We suspect that this always holds for binary search trees.

Conjecture 9.3. *For all $p \in [0,1]$ and σ,*

$$\text{height-perm}_p(\sigma) \approx \text{height-alter}_p(\sigma) \,.$$

Beyond partial permutations and alterations, one could consider other perturbation models for sequences. From a more abstract point of view, a future research direction is 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 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. [4]), routing algorithms, and other data structures. Hopefully, this will shed some light on the discrepancy between worst case and average case behaviour of these algorithms.

Acknowledgements

We thank Jan Arpe for valuable discussions and comments.

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] Noga Alon, Joel H. Spencer, and Paul Erdős. *The Probabilistic Method.* John Wiley & Sons, 1992.

[4] 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.

[5] Luca Becchetti, Stefano Leonardi, Alberto Marchetti-Spaccamela, Guido Schäfer, and Tjark Vredeveld. Average case and smoothed competitive analysis of the multi-level feedback algorithm. In *Proc. of the 44th Ann. IEEE Symp. on Foundations of Computer Science (FOCS)*, pages 462–471. IEEE Computer Society, 2003.

[6] René Beier and Berthold Vöcking. Typical properties of winners and losers in discrete optimization. In *Proc. of the 36th Ann. ACM Symp. on Theory of Computing (STOC)*, pages 343–352. ACM Press, 2004.

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

[13] 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.

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

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

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

[17] 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.

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

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

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

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

[22] 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.

[23] 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.

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

[25] 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.

[26] 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.

[27] Guido Schäfer and Naveen Sivadasan. Topology matters: Smoothed competitiveness of metrical task systems. In Volker Diekert and Michel Habib, editors, *Proc. of the 22th Int. Symp. on Theoretical Aspects of Computer Science (STACS)*, volume 2996 of *Lecture Notes in Computer Science*, pages 489–500. Springer, 2004.

[28] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. In *Proc. of the 33rd Ann. ACM Symp. on Theory of Computing (STOC)*, pages 296–305. ACM Press, 2001.

[29] 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.

[30] 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.

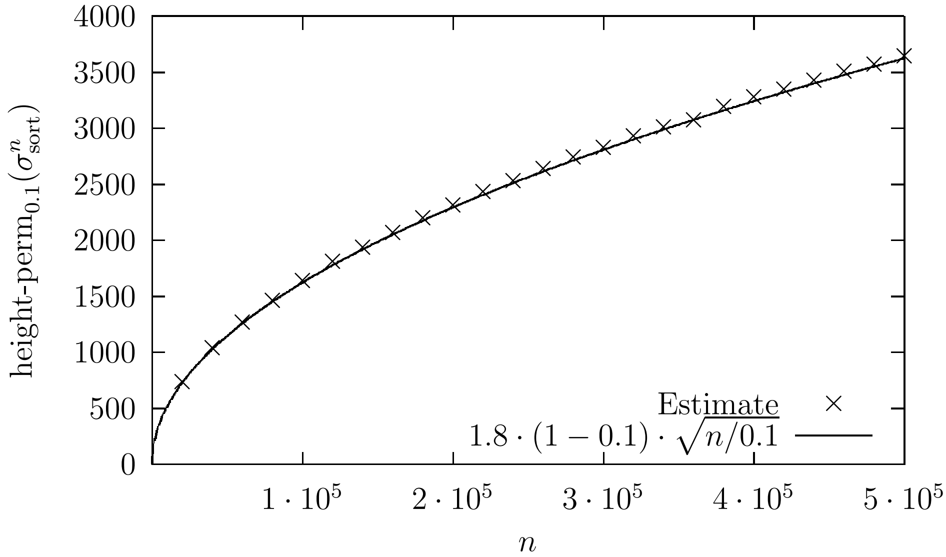
[31] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, 2004.

A Experimental Results

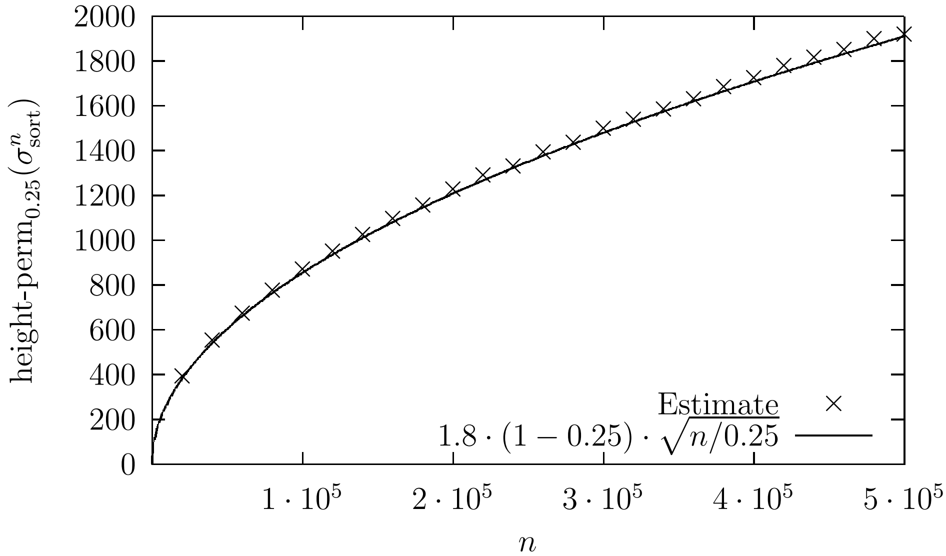
For $n \in \{20\,000, 40\,000, \ldots, 500\,000\}$ and $p \in \{0.1, 0.25\}$, we have randomly performed 5 000 p-partial permutations on σ^n_{sort}. Then we have estimated the expected height height-perm$_p(\sigma^n_{\text{sort}})$ as the average height of the trees generated by the sequences thus obtained. Figure 5 shows the results compared to $1.8 \cdot (1-p) \cdot \sqrt{n/p}$.

We have performed the same experiment for $n \in \{100\,000, 500\,000\}$ and $p \in \{0.05, 0.10, \ldots, 0.95\}$. Figure 6 shows the results, again compared to $1.8 \cdot (1-p) \cdot \sqrt{n/p}$.

These experiments lead us to Conjecture 9.2.

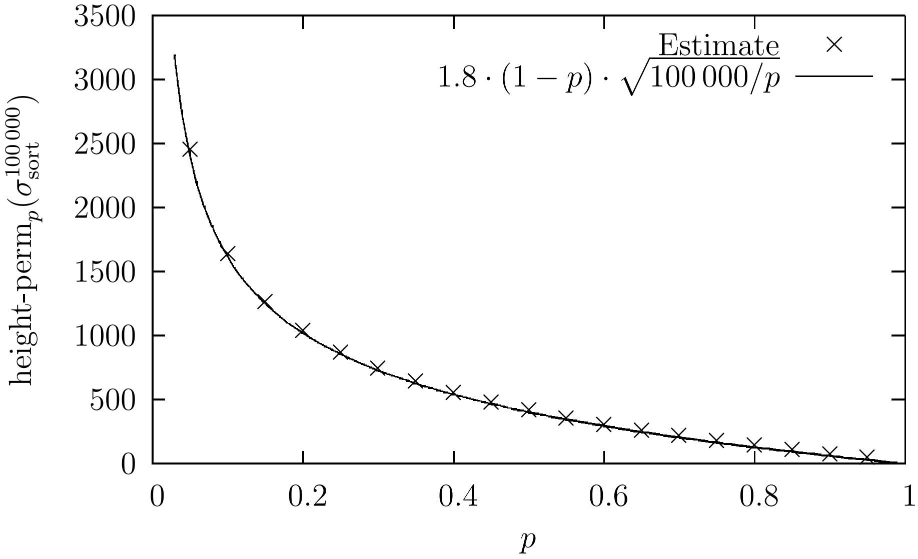


(a) $p = 0.1$.

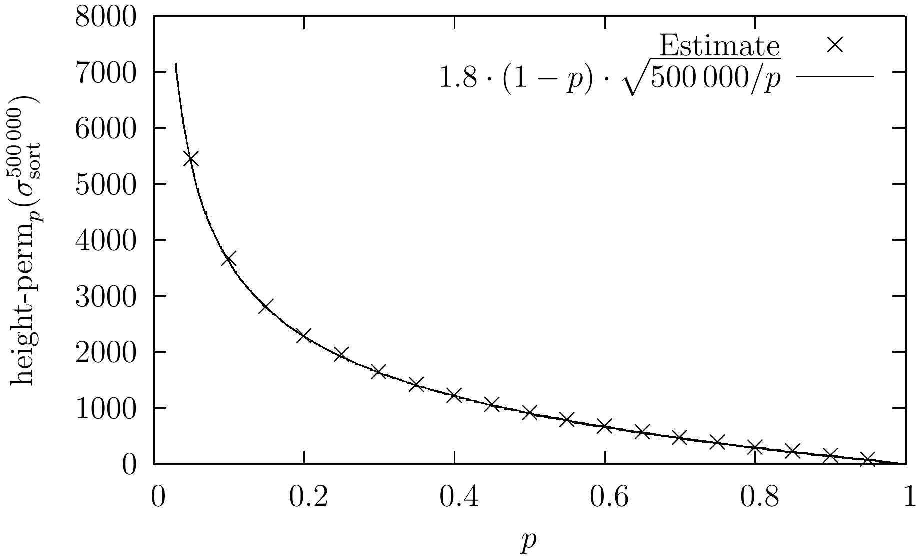


(b) $p = 0.25$.

Figure 5: Experimental data for $n \in \{20\,000, 40\,000, \ldots, 500\,000\}$ and $p \in \{0.1, 0.25\}$ compared to $1.8 \cdot (1-p) \cdot \sqrt{n/p}$.



(a) $n = 100\,000$.



(b) $n = 500\,000$.

Figure 6: Experimental data, in dependence of p, for $p \in \{0.05, 0.10, \ldots, 0.95\}$ and $n \in \{100\,000, 500\,000\}$ compared to $1.8 \cdot (1-p) \cdot \sqrt{n/p}$.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'