

# Smoothed Analysis of Left-To-Right Maxima with Applications\*

Valentina Damerow<sup>†</sup>  
Deutsche Forschungsgemeinschaft  
Kennedyallee 40  
53175 Bonn, Germany

Bodo Manthey<sup>‡</sup>  
University of Twente  
Department of Applied Mathematics  
P. O. Box 217, 7500 AE Enschede, The Netherlands  
`b.manthey@utwente.nl`

Friedhelm Meyer auf der Heide  
Heinz Nixdorf Institute and Computer Science  
Department, University of Paderborn  
33102 Paderborn, Germany  
`fmadh@cs.upb.de`

Harald Räcke  
Department of Computer Science  
University of Warwick  
Coventry CV4 7AL, United Kingdom  
`H.Raecke@warwick.ac.uk`

Christian Scheideler  
Computer Science Department  
University of Paderborn  
Fürstenallee 11, 33102 Paderborn, Germany  
`scheideler@uni-paderborn.de`

Christian Sohler  
Technische Universität Dortmund  
Fakultät für Informatik  
Otto-Hahn-Straße 14, 44221 Dortmund, Germany  
`christian.sohler@tu-dortmund.de`

Till Tantau  
Universität zu Lübeck  
Institut für Theoretische Informatik  
Ratzeburger Allee 160, 23538 Lübeck, Germany  
`tantau@tcs.uni-luebeck.de`

August 25, 2009

A left-to-right maximum in a sequence of  $n$  numbers  $s_1, \dots, s_n$  is a number that is strictly larger than all preceding numbers. In this paper we present a smoothed analysis of the number of left-to-right maxima in the presence of additive random noise. We show that for every sequence of  $n$  numbers  $s_i \in [0, 1]$  that are perturbed by uniform noise from the interval  $[-\epsilon, \epsilon]$ , the expected number of left-to-right maxima is  $\Theta(\sqrt{n/\epsilon} + \log n)$  for  $\epsilon > 1/n$ . For Gaussian noise with standard deviation  $\sigma$  we obtain a bound of  $O(\frac{\log^{3/2} n}{\sigma} + \log n)$ .

---

\*The results of this work have been presented at the *11th Ann. European Symposium on Algorithms (ESA 2003)* [5] and the *33rd Int. Symposium on Mathematical Foundations of Computer Science (MFCS 2008)* [17].

<sup>†</sup>Work was done while the author was affiliated with the Heinz Nixdorf Institute and Computer Science Department, University of Paderborn, Germany.

<sup>‡</sup>Work done in part at the Department of Computer Science at Yale University (P. O. Box 208285, New Haven, CT 06520-8285, USA) supported by the Postdoc-Program of the German Academic Exchange Service (DAAD) and at the Department of Computer Science at Saarland University (P. O. Box 151150, 66041 Saarbrücken, Germany).

We apply our results to the analysis of the smoothed height of binary search trees and the smoothed number of comparisons in the quicksort algorithm and prove bounds of  $\Theta(\sqrt{n/\epsilon} + \log n)$  and  $\Theta(\frac{n}{\epsilon+1}\sqrt{n/\epsilon} + n \log n)$ , respectively, for uniform random noise from the interval  $[-\epsilon, \epsilon]$ . Our results can also be applied to bound the smoothed number of points on a convex hull of points in the two-dimensional plane and to *smoothed motion complexity*, a concept we describe in this paper. We bound how often one needs to update a data structure storing the smallest axis-aligned box enclosing a set of points moving in  $d$ -dimensional space.

## 1 Introduction

To explain the discrepancy between average-case and worst-case behavior of the simplex algorithm, Spielman and Teng introduced the notion of *smoothed analysis* [21]. Smoothed analysis interpolates between average-case and worst-case analysis: Instead of taking a worst-case instance or, as in average-case analysis, choosing an instance completely at random, we analyze the complexity of (possibly worst-case) objects subject to slight random perturbations. On the one hand, perturbations model that nature is not (at least, not always) adversarial. On the other hand, perturbations reflect the fact that data is often subject to measurement or rounding errors; even if the instance at hand was initially a worst-case instance, due to such errors we would probably get a less difficult instance in practice. Spielman and Teng [22] give a survey of results and open problems in smoothed analysis.

The number of left-to-right maxima of a sequence is the number of new maxima that we see when scanning the sequence from left-to-right. The number of left-to-right maxima is often considered in the context of permutations or in the analysis of basic algorithms [13, 15], such as finding a maximum in a sequence of numbers, where the number of left-to-right maxima is equal to the number of updates of the current maximum. In the worst case, every new element is greater than its predecessors, and we have  $n$  left-to-right maxima. On average, we expect to see  $H_n = \sum_{i=1}^n 1/i \approx \ln n$  left-to-right maxima, and the variance of this number is  $H_n - H_n^{(2)}$ , where  $H_n^{(2)} = \sum_{i=1}^n 1/i^2$  denotes the second-order harmonic numbers.

In this paper, we present a smoothed analysis of the number of left-to-right maxima in order to close the gap between the logarithmic average case and the linear worst case. Then we apply our findings to a smoothed analysis of the height of binary search trees, the number of comparisons needed by quicksort, the smoothed number of points on the convex hull, and properties of moving objects.

Binary search trees are one of the most fundamental data structures in computer science and the building blocks for a large variety of data structures. One of the most important parameters of binary search trees is their *height*. The worst-case height of a binary tree for  $n$  numbers is  $n$ . The average-case behavior has been the subject of a considerable amount of research, culminating in the result that the average-case height is  $\alpha \ln n - \beta \ln \ln n + O(1)$ , where  $\alpha \approx 4.311$  is the larger root of  $\alpha \ln(2e/\alpha) = 1$  and  $\beta = 3/(2 \ln(\alpha/2)) \approx 1.953$  [19]. Furthermore, the variance of the height is bounded by a constant [6, 19], and also all higher moments are bounded by constants [6]. Drmota [7] gives a recent survey.

Beyond being an important data structure, binary search trees play a central role in the analysis of divide-and-conquer algorithms like quicksort [15, Section 5.2.2]. While quicksort needs  $\Theta(n^2)$  comparisons in the worst case, the average number of comparisons is  $2n \log n - \Theta(n)$  with a variance of  $7n^2 + 4(n+1)^2 H_n^{(2)} - 2(n+1)H_n + 13n \approx (7 - \frac{2}{3}\pi^2) \cdot n^2 - 2n \ln n + O(n)$  as

mentioned by Fill and Janson [9]. Quicksort and binary search trees are closely related: The height of the tree  $T(s)$  obtained from a sequence  $s$  is equal to the number of levels of recursion required by quicksort to sort  $s$ . The number of comparisons, which corresponds to the total path length of  $T(s)$ , is at most  $n$  times the height of  $T(s)$ .

Binary search trees are also related to the number of left-to-right maxima: The number of left-to-right maxima of  $s$  is equal to the length of the rightmost path of the tree  $T(s)$ , which means that left-to-right maxima provide an easy-to-analyze lower bound for the height of binary search trees. Given the discrepancies between average-case and worst-case behavior of binary search trees, quicksort, and the number of left-to-right maxima, the question arises of what happens in between when the randomness is limited.

The convex hull of a set of points is the smallest convex polygon that contains all points in the set and computing this polygon is one of the basic problems in computational geometry. The average-case number of points on the convex hull of points has been studied extensively in the past [20]; we contribute upper and lower bounds on the *smoothed* number of points on the convex hull. Again, the analysis utilizes left-to-right maxima.

Besides analyzing binary search trees, quicksort, and the convex hull problem, we also introduce a new measure for the complexity of maintaining a geometric structure under motion, which we call *smoothed motion complexity*. The smoothed motion complexity of a basic geometric structure, namely the smallest axis-aligned bounding box of a set of moving points is closely related to left-to-right maxima.

**Our results.** We start our analyses with bounds for the number of left-to-right maxima of a sequence  $s_1, \dots, s_n$  of numbers  $s_i \in [0, 1]$  under various noise distributions in Section 3: First, we prove an upper bound of  $O(\sqrt{n/\epsilon} + \log n)$  for the smoothed number of left-to-right maxima if the noise for each of the  $n$  elements is drawn uniformly and independently from the interval  $[-\epsilon, \epsilon]$ . Thus, the average-case number of left-to-right maxima is obtained for  $\epsilon \in \Omega(n/\log^2 n)$ , and we have an improvement over the worst case already for  $\epsilon \in \omega(1/n)$ . For constant values of  $\epsilon$ , which corresponds to a perturbation by a constant percentage like 1%, we get a bound of  $O(\sqrt{n})$ .

After that, we prove a general lemma that bounds the smoothed number of left-to-right maxima for arbitrary noise distributions in Section 3.2. We apply this lemma to the case of Gaussian noise of standard deviation  $\sigma$ , for which we obtain a bound of  $O(\frac{\log^{3/2} n}{\sigma} + \log n)$ , which yields the average case already for  $\sigma \in \Omega(\sqrt{\log n})$  as shown in Section 3.3. Then we consider arbitrary unimodal noise distributions in Section 3.4: Let  $\varphi$  be the density of the noise distribution that attains its maximum at 0, then the smoothed number of left-to-right maxima is at most  $O(\sqrt{n \log n} \cdot \varphi(0) + \log n)$ . We conclude this section with a lower bound that depends on the noise distribution (Section 3.5). In particular, we obtain a lower bound of  $\Omega(\sqrt{n/\epsilon} + \log n)$  for the smoothed number of left-to-right maxima under uniform noise. This matches the upper bound of Section 3.1 up to constant factors.

In Section 4, we exploit our results for left-to-right maxima to prove the (asymptotically) same bounds for the smoothed height of binary search trees under uniform noise: If the noise is drawn uniformly from  $[-\epsilon, \epsilon]$ , then the smoothed height of the binary search tree obtained from the perturbed sequence is  $O(\sqrt{n/\epsilon} + \log n)$ . As for left-to-right maxima under uniform noise, the average case is obtained for  $\epsilon \in \Omega(n/\log^2 n)$ .

We analyze the number of comparisons quicksort needs to sort a perturbed sequence under uniform noise in Section 5: The smoothed number of comparisons is  $\Theta(\frac{n}{\epsilon+1} \sqrt{n/\epsilon} + n \log n)$ .

For the smoothed number of points on the convex hull in the plane, studied in Section 6, we obtain upper and lower bounds that do not quite match. However, we show at least that while the upper bound for the smoothed number of points on the convex hull is poly-logarithmic for Gaussian noise, it is  $\Omega(\sqrt[5]{n}/\epsilon)$  for uniform noise.

Finally, we introduce the concept of *smoothed motion complexity* in Section 7 as a realistic measure for maintaining a geometric data structure under motion. We illustrate our results on the example of the axis-aligned bounding box for a set of points moving in  $d$ -dimensional space.

**Related work.** The first smoothed analysis of quicksort, due to Banderier, Beier, and Mehlhorn [2], uses a perturbation model different from the one used in the present paper, namely a *discrete perturbation model*. Such models take discrete objects like permutations as input and again yield discrete objects like another permutation. Banderier et al. used *p-partial permutations*, which work as follows: An adversary chooses a permutation of the numbers  $\{1, \dots, n\}$  as sequence, every element of the sequence is marked independently with a probability of  $p$ , and then the marked elements are randomly permuted. Banderier et al. showed that the number of comparisons subject to  $p$ -partial permutations is  $O(\frac{n}{p} \cdot \log n)$ . Furthermore, they proved bounds on the smoothed number of left-to-right maxima in this model.

Manthey and Reischuk [16] analyzed the height of binary search trees under  $p$ -partial permutations. They proved an upper bound of  $O((1-p) \cdot \sqrt{n/p})$  and an asymptotically matching lower bound for the smoothed tree height. For the number of left-to-right maxima, they proved asymptotically matching bounds.

Special care must be taken when defining perturbation models for discrete inputs: The perturbation should favor instances in the neighborhood of the adversarial instance, which requires a suitable definition of neighborhood in the first place, and the perturbation should preserve the global structure of the adversarial instance. Partial permutations have the first feature [16, Lemma 3.2], but destroy much of the global order of the adversarial sequence.

The concept of *smoothed motion complexity* is closely related to the concept of *kinetic data structures* (KDS), which was introduced by Basch et al. [3]. In kinetic data structures the (near) future motion of all objects is known and can be specified by so-called pseudo-algebraic functions of time specified by linear functions or low-degree polynomials. This specification is called a *flight plan*. The goal is to maintain the description of a combinatorial structure as the objects move according to this flight plan. Interesting kinetic data structures have been developed, for instance for connectivity of discs [10] and rectangles [12], convex hulls [3], proximity problems [4], and collision detection for simple polygons [14]. Basch et al. [3] developed a KDS to maintain a bounding box of a moving point set in  $\mathbb{R}^d$ . The number of updates that these data structures need is  $O(n \log n)$ , which is close to the  $\Theta(n)$  updates that may be needed in any case in the worst-case. Agarwal and Har-Peled [1] showed that it is possible to maintain a  $(1 + \epsilon)$ -approximation of such a bounding box. The advantage of this approach is that the motion complexity of this approximation is only  $O(1/\sqrt{\epsilon})$ . The average case motion complexity has also been considered in the past. It has been shown that if  $n$  particles are drawn independently from the unit square, then the expected number of combinatorial changes in the convex hull is  $\Theta(\log^2(n))$ , in the Voronoi diagram  $\Theta(n^{3/2})$  and in the closest pair  $\Theta(n)$  [23].

## 2 Preliminaries

Intervals of the real axis are denoted by  $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ . To denote an interval that does not include an endpoint, we replace the square bracket next to the endpoint by a parenthesis. We denote *sequences* of real numbers by  $s = (s_1, \dots, s_n)$ , where  $s_i \in \mathbb{R}$ . For  $U = \{i_1, \dots, i_\ell\} \subseteq \{1, \dots, n\}$  with  $i_1 < i_2 < \dots < i_\ell$  let  $s_U = (s_{i_1}, s_{i_2}, \dots, s_{i_\ell})$  denote the *subsequence* of  $s$  of the elements at positions in  $U$ .

The  $n$ th harmonic number will be denoted by  $H_n = \sum_{i=1}^n \frac{1}{i}$ .

We denote probabilities by  $\mathbb{P}$  and expected values by  $\mathbb{E}$ . To bound large deviations from the expected value, we will use the Chernoff bound [18, Sect. 4.1]: Let  $X_1, \dots, X_n \in \{0, 1\}$  be independent random variables with  $\mathbb{P}(X_i = 1) = p = 1 - \mathbb{P}(X_i = 0)$ . Let  $X = \sum_{i=1}^n X_i$ . Then  $\mathbb{E}(X) = pn$  and, for every  $\delta > 0$ , we have

$$\mathbb{P}(X > (1 + \delta) \cdot pn) < \left( \frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^{pn}.$$

Throughout the paper, for simplicity we will assume that numbers like  $\sqrt{n}$  are integers and we do not write down the tedious floor and ceiling functions that are actually necessary. Since we are interested in asymptotic bounds, this does not affect the validity of the proofs.

### 2.1 Left-To-Right Maxima, Binary Search Trees, and Quicksort

Let  $s$  be a sequence of length  $n$  consisting of pairwise distinct elements. For the following definitions, let  $R = \{i \in \{1, \dots, n\} \mid s_i > s_1\}$  be the set of positions of elements greater than  $s_1$ , and let  $L = \{i \in \{1, \dots, n\} \mid s_i < s_1\}$  be the set of positions of elements smaller than  $s_1$ .

The number of *left-to-right maxima* of  $s$  is the number of maxima seen when scanning  $s$  from left to right: let  $\text{ltrm}(s) = 1 + \text{ltrm}(s_R)$ , and let  $\text{ltrm}(s) = 0$  when  $s$  is the empty sequence.

From  $s$ , we obtain a *binary search tree*  $T(s)$  by iteratively inserting the elements  $s_1, \dots, s_n$  into the initially empty tree as follows: The root of  $T(s)$  is  $s_1$ . The left subtree of the root  $s_1$  is  $T(s_L)$ , and the right subtree of  $s_1$  is  $T(s_R)$ . The *height* of  $T(s)$  is the maximum number of nodes on any root-to-leaf path of  $T(s)$ :  $\text{height}(s) = 1 + \max\{\text{height}(s_L), \text{height}(s_R)\}$ , and let  $\text{height}(s) = 0$  when  $s$  is the empty sequence. The number of left-to-right maxima of  $s$  is equal to the length of the rightmost path of  $T(s)$ , so  $\text{ltrm}(s) \leq \text{height}(s)$ .

*Quicksort* is the following sorting algorithm: Given  $s$ , we construct  $s_L$  and  $s_R$ . To do this, all elements of  $(s_2, \dots, s_n)$  have to be compared to  $s_1$ , which is called the *pivot element*. Then we sort  $s_L$  and  $s_R$  recursively to obtain  $s'_L$  and  $s'_R$ , respectively. Finally, we output  $s' = (s'_L, s_1, s'_R)$ . The number  $\text{qs}(s)$  of comparisons needed to sort  $s$  is thus  $\text{qs}(s) = (n - 1) + \text{qs}(s_L) + \text{qs}(s_R)$  if  $s$  has a length of  $n \geq 1$ , and  $\text{qs}(s) = 0$  when  $s$  is the empty sequence.

### 2.2 Perturbation Models

Let  $\varphi: \mathbb{R} \rightarrow \mathbb{R}_+$  be a density function. Given a sequence  $s$  of  $n$  numbers chosen by an adversary from the interval  $[0, 1]$ , we draw a *noise*  $\nu_i$  for each  $i \in \{1, \dots, n\}$  independently according to the density function  $\varphi$ . Then we obtain the perturbed sequence  $\bar{s} = (\bar{s}_1, \dots, \bar{s}_n)$  by adding  $\nu_i$  to  $s_i$ , that is,  $\bar{s}_i = s_i + \nu_i$ . Note that  $\bar{s}_i$  need no longer be an element of  $[0, 1]$ . All elements of  $\bar{s}$  are distinct with probability 1.

For a given permutation model  $\varphi$ , we define the random variables  $\text{height}_\varphi(s)$ ,  $\text{qs}_\varphi(s)$ , and  $\text{ltrm}_\varphi(s)$ , which denote the smoothed search tree height, smoothed number of quicksort comparisons, and smoothed number of left-to-right maxima, respectively, when the sequence  $s$

is perturbed according to  $\varphi$ . Since the adversary is choosing  $s$ , our goal are bounds for  $\max_{s \in [0,1]^n} \mathbb{E}(\text{height}_\varphi(s))$ ,  $\max_{s \in [0,1]^n} \mathbb{E}(\text{qs}_\varphi(s))$ , and  $\max_{s \in [0,1]^n} \mathbb{E}(\text{ltrm}_\varphi(s))$ . In the following, we sometimes write  $\text{height}(\bar{s})$  instead of  $\text{height}_\varphi(s)$  if  $\varphi$  is clear from the context. Since  $\bar{s}$  is random,  $\text{height}(\bar{s})$  is also a random variable. Similarly, we will use  $\text{ltrm}(\bar{s})$  and  $\text{qs}(\bar{s})$ .

The choice of the interval sizes is arbitrary since the model is invariant under scaling if we scale  $\varphi$  accordingly. This is summarized in the following lemma, which we will exploit a couple of times in the following.

**Lemma 2.1.** *Let  $b > a$ , and let  $\varphi'$  be defined by  $\varphi'(x) = \varphi((b-a) \cdot x)$ . Then*

$$\max_{s \in [a,b]^n} \mathbb{E}(\text{ltrm}_\varphi(s)) = \max_{s \in [0,1]^n} \mathbb{E}(\text{ltrm}_{\varphi'}(s)).$$

*For other smoothed quantities analogous equalities hold.*

We can always shift  $\varphi$  to obtain a probability distribution with mean 0 without changing anything. Thus, we restrict ourselves in the following to consider distributions with mean 0.

We mainly consider the following two probability distributions: First, the uniform distribution on the interval  $[-\epsilon, \epsilon]$ , which is denoted by  $U[\epsilon]$  and has density  $x \mapsto \frac{1}{2\epsilon}$  if  $x \in [-\epsilon, \epsilon]$  and  $x \mapsto 0$  otherwise. Second, the Gaussian distribution with standard deviation  $\sigma$  and mean 0, which has density

$$x \mapsto \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

and is denoted by  $N[\sigma]$ . We also consider *unimodal* distributions, which are distributions whose density functions assume their maximum at 0 and are monotonically decreasing to both sides. (They are not necessarily symmetric.) In particular,  $U[\epsilon]$  and  $N[\sigma]$  are unimodal distributions.

## 3 Smoothed Number of Left-To-Right Maxima

### 3.1 Uniform Noise

We start our analyses with the smoothed number of left-to-right maxima under uniform noise. Our aim for the present section is to prove the following upper bound on the smoothed number of left-to-right maxima. This bound is asymptotically equal to our lower bound for uniform noise (Corollary 3.8, Section 3.5).

**Theorem 3.1.**

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{ltrm}_{U[\epsilon]}(s)) \in O(\sqrt{n/\epsilon} + \log n).$$

The following notation will be helpful: For  $j \leq 0$ , let  $s_j = \nu_j = 0$ . This allows us to define  $\delta_i = s_i - s_{i-\sqrt{n\epsilon}}$  for all  $i \in \{1, \dots, n\}$ . We define  $I_i = \{j \in \{1, \dots, n\} \mid i - \sqrt{n\epsilon} \leq j < i\}$  to be the set of the  $|I_i| = \min\{i-1, \sqrt{n\epsilon}\}$  positions that precede  $i$ .

To prove Theorem 3.1, we first use a “bubble-sorting argument” to show that the adversary should choose a sorted sequence. Note that this is not as obvious as it may seem: In Section 3.2 (Theorem 3.4) we show that this bubble-sorting argument does not apply to all distributions. Also note that we can assume that  $\epsilon \geq 1/n$  because  $\epsilon < 1/n$  implies a bound of  $O(n)$ , which is always true since we can have at most  $O(n)$  left-to-right maxima in the worst case.

**Lemma 3.2.** For all  $\epsilon > 0$ , and for every sequence  $s$  and its sorted version  $\hat{s}$ , we have

$$\mathbb{E}(\text{ltrm}_{U[\epsilon]}(s)) \leq \mathbb{E}(\text{ltrm}_{U[\epsilon]}(\hat{s})).$$

*Proof.* We prove the lemma by “bubble-sorting”  $s$ . If  $s$  is already sorted, then there is nothing to show. Otherwise, there exist adjacent  $s_i$  and  $s_{i+1}$  with  $s_i > s_{i+1}$ . Our aim is to show that  $\mathbb{E}(\text{ltrm}_{U[\epsilon]}(s)) \leq \mathbb{E}(\text{ltrm}_{U[\epsilon]}(s'))$  where  $s'$  is obtained from  $s$  by swapping  $s_i$  and  $s_{i+1}$ . Then the claim follows by iteratively applying this argument.

After perturbation with  $\nu$ , we obtain  $\bar{s}$  and  $\bar{s}'$ , where  $\bar{s}'_i = s'_i + \nu_{i+1} = s_{i+1} + \nu_{i+1}$  and  $\bar{s}'_{i+1} = s'_{i+1} + \nu_i = s_i + \nu_i$ . Now we analyze the number of left-to-right maxima of  $\bar{s}$  and  $\bar{s}'$ . To do this, let  $\delta = s_i - s_{i+1} > 0$ . We distinguish between two cases. First, we condition on  $\nu_i \in [-\epsilon, \epsilon - \delta]$  and  $\nu_{i+1} \in [\delta - \epsilon, \epsilon]$ . In this case, both  $(\bar{s}_i, \bar{s}_{i+1})$  and  $(\bar{s}'_i, \bar{s}'_{i+1})$  are pairs of random numbers, all of which lie uniformly in the interval  $[s_i - \epsilon, s_{i+1} + \epsilon]$ . Then the expected numbers of left-to-right maxima of  $\bar{s}$  and  $\bar{s}'$  are equal. Second, we condition on the event that  $\nu_i \in (\epsilon - \delta, \epsilon]$  or  $\nu_{i+1} \in [-\epsilon, \delta - \epsilon)$ . In either case, both  $\bar{s}_i > \bar{s}_{i+1}$  and  $\bar{s}'_i < \bar{s}'_{i+1}$  hold. Then  $\bar{s}_{i+1}$  cannot be a left-to-right maximum in  $\bar{s}$ , and if  $\bar{s}_i$  is a left-to-right maximum in  $\bar{s}$ , then so is  $\bar{s}'_{i+1}$  in  $\bar{s}'$ . Since the case distinction is exhaustive, the lemma is proved.  $\square$

We can now embark on the proof of Theorem 3.1.

*Proof of Theorem 3.1.* By Lemma 3.2 we can restrict ourselves to proving the lemma for sorted sequences  $s$ . We estimate the probability that a given  $\bar{s}_i$  for  $i \in \{1, \dots, n\}$  is a left-to-right maximum. Then the bound follows by the linearity of expectation. To bound the probability that  $\bar{s}_i$  is a left-to-right maximum (ltrm), consider the following computation:

$$\mathbb{P}(\bar{s}_i \text{ is an ltrm}) \leq \mathbb{P}(\forall j \in I_i: \nu_j < \bar{s}_i - s_{i-\sqrt{n}\epsilon}) \quad (1)$$

$$\leq \mathbb{P}(\epsilon < \bar{s}_i - s_{i-\sqrt{n}\epsilon}) + \int_{-\epsilon}^{\epsilon - \delta_i} \mathbb{P}(\forall j \in I_i: \nu_j < s_i + x - s_{i-\sqrt{n}\epsilon}) \cdot \frac{1}{2\epsilon} dx \quad (2)$$

$$\leq \frac{\delta_i}{2\epsilon} + \int_{-\epsilon}^{\epsilon} \mathbb{P}(\forall j \in I_i: \nu_j < x) \cdot \frac{1}{2\epsilon} dx \quad (3)$$

$$\leq \frac{\delta_i}{2\epsilon} + \mathbb{P}(\forall j \in I_i: \nu_j < \nu_i) = \frac{\delta_i}{2\epsilon} + \frac{1}{|I_i| + 1}. \quad (4)$$

To see that (1) holds, assume that  $\bar{s}_i$  is a left-to-right maximum. Then  $\bar{s}_i - s_{i-\sqrt{n}\epsilon}$  must be larger than the noises of all the elements in the index range  $I_i$ , for if the noise  $\nu_j$  of some element  $s_j$  were larger than  $\bar{s}_i - s_{i-\sqrt{n}\epsilon}$ , then  $\bar{s}_j = s_j + \nu_j$  would be larger than  $s_j + \bar{s}_i - s_{i-\sqrt{n}\epsilon}$ . Since the sequence is sorted, we would get  $s_j + \bar{s}_i - s_{i-\sqrt{n}\epsilon} \geq \bar{s}_i$ , and  $\bar{s}_i$  would not be a left-to-right maximum.

For (2), first observe that  $\nu_j < \bar{s}_i - s_{i-\sqrt{n}\epsilon}$  is surely the case for all  $j \in I_i$  if  $\epsilon < \bar{s}_i - s_{i-\sqrt{n}\epsilon}$ . So, consider the case  $\epsilon \geq \bar{s}_i - s_{i-\sqrt{n}\epsilon} = \delta_i + \nu_i$ . Then  $\nu_i \in [-\epsilon, \epsilon - \delta_i]$  and we can rewrite  $\mathbb{P}(\forall j \in I_i: \nu_j < \delta_i + \nu_i)$  as  $\int_{-\epsilon}^{\epsilon - \delta_i} \mathbb{P}(\forall j \in I_i: \nu_j < \delta_i + x) \cdot \frac{1}{2\epsilon} dx$ , where  $\frac{1}{2\epsilon}$  is the density of the random variable  $\nu_i$ . For (3) observe that  $\epsilon < \bar{s}_i - s_{i-\sqrt{n}\epsilon}$  is equivalent to  $\epsilon - \delta_i < \nu_i$  and the probability of this is  $\delta_i/(2\epsilon)$ . Furthermore, we performed an index shift in the integral. In (4), we replaced the integral by a probability once more and get the final result.

Let us bound  $\sum_{i=1}^n \delta_i$ . We have  $\sum_{i=1}^n \delta_i = \sum_{i=1}^n (s_i - s_{i-\sqrt{n}\epsilon}) = \sum_{i=n-\sqrt{n}\epsilon+1}^n s_i \leq \sqrt{n}\epsilon$ . The second equality holds since most  $s_i$  cancel themselves out and  $s_i = 0$  for  $i \leq 0$ . The inequality

holds since there are  $\sqrt{n\epsilon}$  summands. We complete the proof by bounding  $1/(|I_i| + 1) = 1/\min\{i, \sqrt{n\epsilon} + 1\}$  by  $1/i + 1/\sqrt{n\epsilon}$  and summing over all  $i$ :

$$\mathbb{E}(\text{ltrm}_\epsilon(s)) \leq \sum_{i=1}^n \frac{\delta_i}{2\epsilon} + \sum_{i=1}^n \frac{1}{|I_i| + 1} \leq \frac{\sqrt{n\epsilon}}{2\epsilon} + \sum_{i=1}^n \frac{1}{i} + \frac{n}{\sqrt{n\epsilon}} \in O\left(\sqrt{n/\epsilon} + \log n\right).$$

□

### 3.2 General Noise

The goal of this section is to provide a tool to show upper bounds for the smoothed number of left-to-right maxima under arbitrary distributions. Throughout this section, we assume that  $\nu_1, \dots, \nu_n$  are distributed according to a density function  $\varphi$  with probability distribution  $\Phi$ . Then for a given sequence  $s$  the numbers  $\bar{s}_i$  are random variables with probability distribution  $\mathbb{P}(\bar{s}_i \leq x) = \Phi(x - s_i)$  and density  $\varphi(x - s_i)$ .

We can write the probability that the  $i$ th element of  $\bar{s}$  is a left-to-right maximum as

$$\mathbb{P}(\bar{s}_i \text{ is an ltrm of } \bar{s}) = \int_{-\infty}^{\infty} \varphi(x - s_i) \cdot \prod_{j=1}^{i-1} \Phi(x - s_j) dx. \quad (5)$$

The main idea for computing the integral (5) is now to divide the interval  $[0, 1]$  into  $m = 1/\delta$  smaller intervals of length  $\delta$ . Here  $\delta$  is a small parameter that we will specify later on. Then the sequence  $s$  of unperturbed input elements is partitioned into  $m$  subsequences  $s^{(1)}, \dots, s^{(m)}$ , where  $s^{(\ell)}$  contains all elements that lie in the  $\ell$ th such interval of length  $\delta$ , i.e.,  $s^{(\ell)} = s_{S_\ell}$  and

$$S_\ell = \{i \in \{1, \dots, n\} \mid s_i \in ((\ell - 1) \cdot \delta, \ell\delta]\}$$

with  $S_1$  also containing all  $i$  with  $s_i = 0$ .

Now we use the inequality

$$\mathbb{E}(\text{ltrm}_\varphi(s)) \leq \sum_{\ell=1}^m \mathbb{E}(\text{ltrm}_\varphi(s^{(\ell)})), \quad (6)$$

which holds not only for our choice of  $s^{(1)}, \dots, s^{(m)}$ , but for any  $s^{(1)}, \dots, s^{(m)}$  such that each element of  $s$  appears in at least one of the subsequences.

For small enough  $\delta$ , the elements of a subsequence  $s^{(\ell)}$  behave almost as in the usual average case. The reason for this is that the unperturbed input elements lie very close together, so the order of the elements is dominated by the perturbation and not by the original values of the elements.

Let  $n_\ell = |S_\ell|$  be the number of elements of  $s^{(\ell)}$ . Without loss of generality, we analyze  $s^{(1)} = (s'_1, \dots, s'_{n_1})$  in the following. The probability that  $s'_k$  is a left-to-right maximum of  $\bar{s}^{(1)}$  (where  $\bar{s}^{(\ell)}$  is obtained from  $s^{(\ell)}$  by perturbation) is

$$\begin{aligned} \mathbb{P}(s'_k \text{ is an ltrm of } \bar{s}^{(1)}) &= \int_{-\infty}^{\infty} \varphi(x - s'_k) \cdot \prod_{j=1}^{k-1} \Phi(x - s'_j) dx \\ &\leq \int_{-\infty}^{\infty} \varphi(x) \cdot \Phi(x + \delta)^{k-1} dx. \end{aligned} \quad (7)$$



The inequality follows from the observation that the minimum and maximum element of  $s^{(1)}$  differ by at most  $\delta$ , from substituting  $x$  by  $x + s'_k$ , and from the monotonicity of  $\Phi$ .

We can easily bound this integral from above if  $\varphi(x) \leq r \cdot \varphi(x + \delta)$ . In this way, we lose at most a factor of  $r$ , which we will specify later on. Let

$$\mathcal{Z}_{\delta,r}^\varphi = \{x \in \mathbb{R} \mid \varphi(x) > r \cdot \varphi(x + \delta)\} \subseteq \mathbb{R}$$

be the set of numbers for which the ratio  $\varphi(x)/\varphi(x + \delta)$  exceeds  $r$ . Let  $\mathcal{Z}$  be the probability of the set  $\mathcal{Z}_{\delta,r}^\varphi$ , i. e., the probability that a random number drawn according to  $\varphi$  assumes a value in  $\mathcal{Z}_{\delta,r}^\varphi$ :

$$\mathcal{Z} = \int_{\mathcal{Z}_{\delta,r}^\varphi} \varphi(x) \, dx.$$

We can now formulate and prove the main lemma of this section. We will apply this lemma in the following two sections to prove upper bounds on the smoothed number of left-to-right maxima for Gaussian noise (Section 3.3) and for arbitrary unimodal noise distributions (Section 3.4).

**Lemma 3.3.** *Let  $\varphi$  be a continuous, integrable probability distribution, and let  $\delta, r > 0$ . Let  $\mathcal{Z}_{\delta,r}^\varphi = \{x \in \mathbb{R} \mid \varphi(x) > r \cdot \varphi(x + \delta)\}$ . Then*

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{ltrm}_\varphi(x)) \leq r \cdot \lceil 1/\delta \rceil \cdot H_n + n \cdot \mathcal{Z}.$$

*Proof.* Let  $m = \lceil 1/\delta \rceil$ . Without loss of generality, we consider the input subsequence  $s^{(1)}$ . Now (7) yields an upper bound for the probability that the  $k$ th element of  $\bar{s}^{(1)}$  is a left-to-right maximum. This yields

$$\begin{aligned} \mathbb{P}(s'_k \text{ is an ltrm of } \bar{s}^{(1)}) &\leq \int_{\mathbb{R}} \varphi(x) \cdot \Phi(x + \delta)^{k-1} \, dx \\ &\leq \int_{\mathbb{R} \setminus \mathcal{Z}_{\delta,r}^\varphi} \varphi(x + \delta) \cdot \frac{\varphi(x)}{\varphi(x + \delta)} \cdot \Phi(x + \delta)^{k-1} \, dx + \int_{\mathcal{Z}_{\delta,r}^\varphi} \varphi(x) \, dx \\ &\leq r \cdot \int_{-\infty}^{\infty} \varphi(x + \delta) \cdot \Phi(x + \delta)^{k-1} \, dx + \mathcal{Z} \\ &= \frac{r}{k} + \mathcal{Z}. \end{aligned}$$

Thus,

$$\mathbb{E}(\text{ltrm}_\varphi(s^{(1)})) \leq \sum_{k=1}^{n_1} \left( \frac{r}{k} + \mathcal{Z} \right) = r \cdot H_{n_1} + n_1 \cdot \mathcal{Z}.$$

The expected number of left-to-right maxima of  $\bar{s}^{(2)}, \dots, \bar{s}^{(m)}$  can be bounded from above in the same way. Thus, (6) yields

$$\mathbb{E}(\text{ltrm}_\varphi(s)) \leq \sum_{\ell=1}^m \mathbb{E}(\text{ltrm}_\varphi(s^{(\ell)})) \leq \sum_{\ell=1}^m \left( r \cdot H_{n_\ell} + n_\ell \cdot \mathcal{Z} \right) \leq rm \cdot H_n + n \cdot \mathcal{Z}.$$

This proves the lemma. □

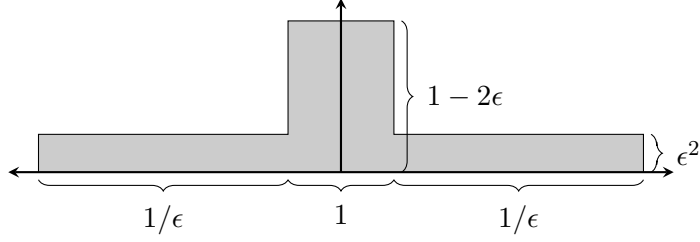


Figure 1: The distribution for Theorem 3.4.

Lemma 3.2 of Section 3.1 states that sorting a sequence can never decrease the expected number of left-to-right maxima – at least when the noise is drawn uniformly from a single interval. It is tempting to assume that a result similar to Lemma 3.2 will hold in the same way for every other noise distribution – at least if the noise distribution is reasonably well-behaved. This was also claimed by Damerow et al. [5] and, as we will see below, this claim is incorrect (fortunately, this claim does not affect the correctness of their other results). There exists a simple distribution and a sequence for which the sorted version has a lower expected number of left-to-right maxima than the original sequence (Theorem 3.4). The probability density is unimodal and symmetric, and it can be made smooth easily without changing the results. (For the sake of simplicity, we do not elaborate on this.)

**Theorem 3.4.** *There exist a sequence  $s$  and a unimodal, symmetric density function  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  such that sorting  $s$  to obtain  $\hat{s}$  decreases the expected number of left-to-right maxima after perturbation.*

*Proof.* The sequences are

$$s = \left( \underbrace{0, \dots, 0}_{K \text{ times}}, \underbrace{1 + \frac{1}{\epsilon}}_{s_1}, \underbrace{\frac{1}{\epsilon}}_{s_0} \right) \text{ and } \hat{s} = \left( \underbrace{0, \dots, 0}_{K \text{ times}}, \frac{1}{\epsilon}, 1 + \frac{1}{\epsilon} \right)$$

Here  $\epsilon > 0$  and  $K$  are numbers that will be fixed later. The distribution (see Figure 1) is defined as follows:  $\varphi(x) = 1 - 2\epsilon$  for  $|x| < 1/2$  and  $\varphi(x) = \epsilon^2$  for  $1/2 \leq |x| \leq 1/\epsilon$  and  $\varphi(x) = 0$  for  $|x| > 1/\epsilon$ . One can easily verify that  $\int_{-\infty}^{\infty} \varphi(x) dx = 1$ . By scaling the sequences and the distribution, one can achieve  $s, \hat{s} \in [0, 1]^{K+2}$  in compliance with our model.

The expected number of left-to-right maxima is the sum of the expected number of left-to-right maxima contributed by the first  $K$  elements plus the probability that at least one of  $\bar{s}_0$  and  $\bar{s}_1$  is a left-to-right maximum plus the probability that both  $\bar{s}_0$  and  $\bar{s}_1$  are left-to-right maxima. The contribution of the first elements is the same in  $\hat{s}$  and  $s$ . The probability that at least one of  $\bar{s}_0$  and  $\bar{s}_1$  contributes a left-to-right maximum is equal to the probability that one of them is greater than the maximum among the first elements, thus does not depend on the order of  $\bar{s}_0$  and  $\bar{s}_1$ . Only the probability that we have two left-to-right maxima differs in  $s$  and  $\hat{s}$ .

Let  $M$  be the maximum among the first  $K$  elements. Let us first assume that  $M = 1/2 + 1/\epsilon$ . After estimating the probabilities that both the last two elements are left-to-right maxima in  $s$  or  $\hat{s}$ , we show that we can get rid of this assumption by making  $K = K(\epsilon)$  sufficiently large.

If the maximum of the first  $K$  elements is  $M = 1/2 + 1/\epsilon$ , then  $\bar{s}_1 = s_1 + \nu_1 > M$  for  $\nu_1 > -1/2$ . Analogously,  $\bar{s}_0 = s_0 + \nu_0 > M$  for  $\nu_0 > 1/2$ . We have to estimate the probabilities  $\mathbb{P}(\bar{s}_0 > \bar{s}_1 > \frac{1}{2} + \frac{1}{\epsilon})$  and  $\mathbb{P}(\bar{s}_1 > \bar{s}_0 > \frac{1}{2} + \frac{1}{\epsilon})$ , which is done in Table 1.

Cases	$\mathbb{P}(\bar{s}_0 > \bar{s}_1 > \frac{1}{2} + \frac{1}{\epsilon})$	$\mathbb{P}(\bar{s}_1 > \bar{s}_0 > \frac{1}{2} + \frac{1}{\epsilon})$
$\nu_1 \in [-\frac{1}{2}, \frac{1}{2}], \nu_0 \in [\frac{1}{2}, \frac{3}{2}]$	$\epsilon^2 \cdot (1 - 2\epsilon) \cdot \frac{1}{2} =$	$\epsilon^2 \cdot (1 - 2\epsilon) \cdot \frac{1}{2}$
$\nu_1 \in [-\frac{1}{2}, \frac{1}{2}], \nu_0 > \frac{3}{2}$	$(\epsilon - \epsilon^2) \cdot (1 - 2\epsilon) \cdot \mathbf{1} \gg$	$\mathbf{0}$
$\nu_1 \in [\frac{1}{2}, \frac{1}{\epsilon} - \frac{1}{2}], \nu_0 \in [\frac{1}{2}, \frac{3}{2}]$	$\mathbf{0} <$	$\epsilon^2 \cdot (\epsilon - \epsilon^2) \cdot \mathbf{1}$
$\nu_1 \in [\frac{1}{2}, \frac{1}{\epsilon} - \frac{1}{2}], \nu_0 > \frac{3}{2}$	$(\epsilon - \epsilon^2) \cdot (\epsilon - \epsilon^2) \cdot \frac{1}{2} =$	$(\epsilon - \epsilon^2) \cdot (\epsilon - \epsilon^2) \cdot \frac{1}{2}$
$\nu_1 > \frac{1}{\epsilon} - \frac{1}{2}, \nu_0 \in \mathbb{R}$	$\mathbf{0} <$	$\mathbf{1} \cdot \epsilon^2 \cdot \mathbf{1}$

Table 1: The case distinction. The probabilities are always of the form  $p_0 \cdot p_1 \cdot p$ , where  $p_i$  is the probability that  $\nu_i$  is in the given interval and  $p$  (shown in bold face) is the probability that we have two left-to-right maxima, given that the noise values are in the respective intervals.

To compare the two probabilities, only the second, third, and fifth row play a role since the other two yield an equal contribution. The second row yields roughly  $\epsilon$  for the left probability and 0 for the right probability, while the third and fifth yield at most  $\epsilon^2 + o(\epsilon^2)$  for the right probability and 0 for the left. By making  $\epsilon$  sufficiently small, other terms can safely be ignored. Overall, the probability that both  $\bar{s}_0$  and  $\bar{s}_1$  yield a left-to-right maximum is larger in  $s$  than in  $\hat{s}$  by roughly  $\epsilon$ . By making  $K$  large enough, we can make sure that the maximum  $M$  among the first  $K$  elements is, with very high probability, at least  $1/2 + 1/\epsilon - \epsilon^c$  for some large constant  $c$ . The effect on the two probabilities is then negligible compared to  $\epsilon$ .  $\square$

### 3.3 Gaussian Noise

As a first application of Lemma 3.3 of the previous section, we prove an upper bound for the smoothed number of left-to-right maxima under Gaussian noise.

**Theorem 3.5.** *For all  $\sigma > 0$ , we have*

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{lrm}_{\mathbb{N}[\sigma]}(s)) \in O\left(\frac{\log^{3/2} n}{\sigma} + \log n\right).$$

*Proof.* In order to apply Lemma 3.3, we choose  $\delta = \sigma/\sqrt{\ln(n)}$ . For  $a \leq \sigma\sqrt{2\ln(n)}$ , we have

$$\frac{\varphi(a)}{\varphi(a + \delta)} = \exp\left(\frac{2\delta a}{2\sigma^2} + \frac{\delta^2}{2\sigma^2}\right) = \exp\left(\frac{a}{\sigma\sqrt{\ln n}} + \frac{1}{2\ln n}\right) \leq \exp\left(\sqrt{2} + \frac{1}{2\ln n}\right) \leq e^3.$$

Thus, for  $r = e^3$ , we have  $\mathcal{Z}_{\delta,r}^\varphi \subseteq [\sigma\sqrt{2\ln(n)}, \infty)$ . Now we need an upper bound for  $\mathcal{Z} = \int_{\mathcal{Z}_{\delta,r}^\varphi} \varphi(x) dx$ : The probability that a Gaussian random variable with standard deviation  $\sigma$  and mean 0 assumes a value of at least  $\sigma k$  for  $k \geq 0$  is

$$\begin{aligned} \int_{\sigma k}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) dx &= \int_k^{\infty} \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2}\right) dx \\ &\leq \frac{1}{k\sqrt{2\pi}} \cdot \exp\left(-\frac{k^2}{2}\right) \leq \exp\left(-\frac{k^2}{2}\right), \end{aligned}$$

where the equality follows from a substitution and the first inequality follows from Durrett [8, Theorem 1.3]. This bound yields

$$\mathcal{Z} = \int_{\mathcal{Z}_{\delta,r}^\varphi} \varphi(x) dx \leq \int_{\sigma\sqrt{2\ln(n)}}^{\infty} \varphi(x) dx \leq \frac{1}{n}.$$

Now we can apply Lemma 3.3 with parameters  $\delta = \frac{\sigma}{\sqrt{\ln n}}$ ,  $r = e^3$ , and  $\mathcal{Z} = 1/n$ . Then, for every input sequence  $s$ , the smoothed number of left-to-right maxima is at most

$$\mathbb{E}(\text{ltrm}_{N[\sigma]}(s)) \leq e^3 \cdot \left\lceil \sqrt{\ln(n)/\sigma} \right\rceil \cdot H_n + 1 \in O\left(\frac{\log^{3/2} n}{\sigma} + \log n\right)$$

as claimed.  $\square$

If  $\sigma \in \Omega(\sqrt{\log n})$ , then we obtain the average-case bound of  $O(\log n)$ . On the other hand, for  $\sigma \in O\left(\frac{\log^{3/2} n}{n}\right)$  we obtain the worst-case upper bound of  $O(n)$ . This means that for such small standard deviations, Gaussian perturbations of adversarial instances have no significant effects in our analysis.

### 3.4 Unimodal Noise

In this section we apply our findings for general noise distributions to obtain upper bounds for arbitrary unimodal distributions. Of course, the following analysis holds also for densities that have a single peak not at the origin but at any other point  $x \in \mathbb{R}$  and are monotonically decreasing to either side of  $x$ . For the sake of simplicity, we assume that the unimodal density  $\varphi$  assumes its maximum at 0.

The following theorem gives an upper bound on the number of left-to-right maxima for arbitrary unimodal noise distributions.

**Theorem 3.6.** *Let  $\varphi$  be a unimodal density function. Then*

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{ltrm}_\varphi(s)) \in O\left(\sqrt{n \log n \cdot \varphi(0)} + \log n\right).$$

*Proof.* Since  $\varphi$  is unimodal,  $\varphi(0)$  is its global maximum. We want to apply Lemma 3.3 with  $r = 2$ , and we postpone choosing  $\delta$ .

We need a bound for  $\mathcal{Z} = \int_{\mathcal{Z}_{\delta,r}^\varphi} \varphi(x) dx$ . To get such a bound, we aim at finding a covering  $\mathcal{Z}_1, \mathcal{Z}_2, \dots$  for the set  $\mathcal{Z}_{\delta,r}^\varphi$ , i. e.,  $\bigcup_i \mathcal{Z}_i \supseteq \mathcal{Z}_{\delta,r}^\varphi$ . Then  $\int_{\bigcup_i \mathcal{Z}_i} \varphi(x) dx$  is an upper bound for  $\mathcal{Z}$ .

For  $x + \delta \leq 0$ , we have  $\varphi(x) \leq \varphi(x + \delta)$  because of the monotonicity of  $\varphi$ . Hence,  $\mathcal{Z}_{\delta,r}^\varphi \subseteq [-\delta, \infty)$ . Partition  $[-\delta, \infty)$  into intervals of the form  $[(\ell - 1) \cdot \delta, \ell \delta]$  for  $\ell \in \mathbb{N}$ . Some of these intervals will have a non-empty intersection with  $\mathcal{Z}_{\delta,r}^\varphi$ . Let  $\mathcal{Z}_1$  be the first of them, let  $\mathcal{Z}_2$  be the second, and so on; and let  $\mathcal{Z}_k$  be the empty set if there are less than  $k$  intervals having a non-empty intersection with  $\mathcal{Z}_{\delta,r}^\varphi$ . This definition ensures that  $(\mathcal{Z}_i)_{i \in \mathbb{N}}$  is a covering of  $\mathcal{Z}_{\delta,r}^\varphi$ .

We can now bound  $\int_{\bigcup_i \mathcal{Z}_i} \varphi(x) dx$  as follows. For an interval  $\mathcal{Z}_i$ , let  $z_i \in \mathcal{Z}_i$  denote the point in  $\mathcal{Z}_i$  with maximum density. This means that for interval  $\mathcal{Z}_i \subseteq \mathbb{R}_{\geq 0}$ ,  $z_i$  is the left-endpoint. For the case that  $\mathcal{Z}_1 = [-\delta, 0]$ ,  $z_1 = 0$  is the right end-point of the interval. Then we have

$$\int_{\mathcal{Z}_i} \varphi(x) dx \leq \delta \cdot \varphi(z_i).$$

By our choice of  $r = 2$  and the definition of  $\mathcal{Z}_{\delta,r}^\varphi$  as well as  $\mathcal{Z}_1, \mathcal{Z}_2, \dots$ , we have  $\varphi(z_{i+2}) \leq \frac{1}{2} \varphi(z_i)$  for all  $i \in \mathbb{N}$ . To see this consider a point  $\hat{z}_i \in \mathcal{Z}_i \cap \mathcal{Z}_{\delta,r}^\varphi$ . Then

$$\varphi(z_i) \geq \varphi(\hat{z}_i) > 2 \cdot \varphi(\hat{z}_i + \delta) \geq 2 \cdot \varphi(z_{i+2}).$$

Putting everything together yields

$$\begin{aligned}
\int_{\bigcup_i \mathcal{Z}_i} \varphi(x) \, dx &\leq \sum_{i=1}^{\infty} \int_{\mathcal{Z}_{2^{i-1}}} \varphi(x) \, dx + \sum_{i=1}^{\infty} \int_{\mathcal{Z}_{2^i}} \varphi(x) \, dx \\
&\leq \sum_{i=1}^{\infty} \delta \cdot \varphi(z_{2^{i-1}}) + \sum_{i=1}^{\infty} \delta \cdot \varphi(z_{2^i}) \\
&\leq \sum_{i=1}^{\infty} \frac{1}{2^{i-1}} \cdot \delta \cdot \varphi(z_1) + \sum_{i=1}^{\infty} \frac{1}{2^{i-1}} \cdot \delta \cdot \varphi(z_2) \\
&\leq 2\delta \cdot \varphi(z_1) + 2\delta \cdot \varphi(z_2) \leq 4\delta \cdot \varphi(0) .
\end{aligned}$$

This allows us to bound  $\mathcal{Z} \leq 4\delta \cdot \varphi(0)$ . Lemma 3.3 yields that the smoothed number of left-to-right maxima is at most  $2 \cdot \lceil 1/\delta \rceil \cdot H_n + n4\delta \cdot \varphi(0)$ . Setting  $\delta = \sqrt{H_n/(n \cdot \varphi(0))}$  gives the upper bound that we aimed for:

$$\mathbb{E}(\text{ltrm}_{\varphi}(s)) \in O(\sqrt{n \log n \cdot \varphi(0)}) .$$

□

The Gaussian and the uniform distribution are unimodal distributions. For uniform noise in the interval  $[-\epsilon, \epsilon]$ , Theorem 3.6 yields an upper bound of  $O(\sqrt{n \log n / \epsilon})$ , which is only a factor of  $\sqrt{\log n}$  off the tight bound of Theorem 3.1 (Section 3.1).

If we consider Gaussian noise, we obtain a far less tight bound than the one shown in Theorem 3.5 since the maximal density is  $1/(\sqrt{2\pi}\sigma)$ .

### 3.5 Lower Bound

In this section we derive lower bounds for the smoothed number of left-to-right maxima. Our lower bound is tight for some noise distributions like, for example, the uniform distribution.

**Theorem 3.7.** *Let  $\varphi$  be a density function with support  $[-\epsilon, \epsilon]$  such that  $\varphi$  is smooth in this interval. Then*

$$\max_{x \in [0,1]^n} \mathbb{E}(\text{ltrm}_{\varphi}(x)) \in \Omega \left( \min \left\{ \sqrt{n/\epsilon} \cdot \left( 1 - \Phi(\epsilon - \sqrt{\epsilon/n})^{\sqrt{n\epsilon}} \right), n \right\} \right),$$

where  $\Phi$  denotes the distribution function of  $\varphi$ . If the probability distribution is unimodal and  $\varphi(\epsilon) \neq 0$ , then

$$\max_{x \in [0,1]^n} \mathbb{E}(\text{ltrm}_{\varphi}(x)) \in \Omega \left( \min \left\{ \sqrt{n\varphi(\epsilon)}, n \right\} \right).$$

*Proof.* Consider the following input sequence  $s = (s_1, \dots, s_n)$ : For some  $\ell \in \mathbb{N}$  that we will specify later on, we divide  $s$  into  $m = \lceil n/\ell \rceil$  subsequences  $s^{(1)}, \dots, s^{(m)}$  of length  $\ell$ : Let  $S_i = \{(i-1)\ell + 1, \dots, i\ell\} \cap \{1, \dots, n\}$  and  $s^{(i)} = s_{S_i}$ . We set  $s_{(i-1)\ell+1} = \dots = s_{i\ell} = i\ell/n$  for all  $i \in \{1, \dots, m\}$ . This means that all elements in the same subsequence  $s^{(i)}$  have the same value  $i\ell/n$ .

If element  $s_k$  is from subsequence  $s^{(i)}$ , then  $\bar{s}_k$  is distributed in the interval  $[i\ell/n - \epsilon, i\ell/n + \epsilon]$  after the perturbation. We say that  $\bar{s}_k$  is large if  $\bar{s}_k > (i-1) \cdot \ell/n + \epsilon$ . Then  $\bar{s}_k$  is larger than

all elements in the preceding subsequences  $s^{(1)}, \dots, s^{(i-1)}$ . Thus, if there is at least one element in subsequence  $s^{(i)}$  that assumes a value larger than  $(i-1) \cdot \ell/n + \epsilon$ , then at least one element of  $s^{(i)}$  is a left-to-right maximum.

For any  $k \in S_i$ , we have

$$\begin{aligned} \mathbb{P}(\bar{s}_k \leq (i-1) \cdot \ell/n + \epsilon) &= \mathbb{P}(i\ell/n + \nu_k \leq (i-1) \cdot \ell/n + \epsilon) \\ &= \mathbb{P}(\nu_k \leq \epsilon - \ell/n) \\ &= \Phi(\epsilon - \ell/n). \end{aligned}$$

By independence of the noise values, we thus obtain

$$\mathbb{P}(\text{no element of } s^{(i)} \text{ exceeds } (i-1) \cdot \ell/n + \epsilon) \leq \Phi(\epsilon - \ell/n)^\ell, \quad (8)$$

which yields

$$\mathbb{E}(\text{ltrm}_\varphi(s)) \geq m \cdot \left(1 - \Phi(\epsilon - \ell/n)^\ell\right).$$

Choosing  $\ell = \sqrt{n\epsilon}$ , we obtain  $m = n/\ell = \sqrt{n/\epsilon}$ . The first part of the theorem follows immediately by linearity of expectation and by observing that the number of left-to-right maxima never exceeds  $n$ .

For unimodal noise distributions with density function  $\varphi(\epsilon) \neq 0$ , we observe that  $1 - \Phi(\epsilon - \ell/n) \geq \varphi(\epsilon) \cdot \ell/n$ . Thus,  $\Phi(\epsilon - \ell/n) \leq 1 - \varphi(\epsilon) \cdot \ell/n$ . If we choose  $\ell = \sqrt{n/\varphi(\epsilon)}$ , we obtain an upper bound of  $1/e$  for (8). This yields

$$\mathbb{E}(\text{ltrm}_\varphi(x)) \geq \left(1 - \frac{1}{e}\right) \cdot \sqrt{n\varphi(\epsilon)}.$$

□

**Corollary 3.8.** *If the noise is uniformly distributed in  $[-\epsilon, \epsilon]$ , then*

$$\mathbb{E}(\text{ltrm}_\epsilon(s)) \in \Omega(\min\{\sqrt{\frac{n}{\epsilon}} + \log n, n\}).$$

*Proof.* We apply the second part of Theorem 3.7, and we plug in  $\varphi(\epsilon) = \frac{1}{2\epsilon}$ . If  $\epsilon \in \Omega(n/\log^2 n)$ , then we still have at least the average-case bound of  $\Theta(\log n)$ . □

## 4 Smoothed Height of Binary Search Trees

In this section we exploit our smoothed analysis of left-to-right maxima for uniform noise to prove an exact bound on the smoothed height of binary search trees under uniform noise. The bound is the same as for left-to-right maxima, as stated in the following theorem.

**Theorem 4.1.** *For  $\epsilon \geq 1/n$ , we have*

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{height}_{U[\epsilon]}(s)) \in \Theta(\sqrt{n/\epsilon} + \log n).$$

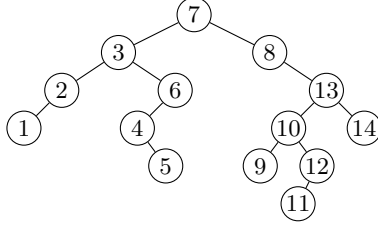


Figure 2: The tree  $T(s)$  obtained from the sequence  $s = (7, 8, 13, 3, 2, 10, 9, 6, 4, 12, 14, 1, 5, 11)$ . We have  $\text{height}(s) = 6$ . The root-to-leaf path ending at 11 yields the increasing run  $\alpha = (7, 8, 10, 11)$  and the decreasing run  $\beta = (13, 12, 11)$ .

In the rest of this section, we prove this theorem. We have to prove an upper and a lower bound, but the lower bound follows directly from the lower bound of  $\Omega(\sqrt{n/\epsilon} + \log n)$  for the smoothed number of left-to-right maxima (Theorem 3.1, the number of left-to-right maxima in a sequence is the length of the rightmost path of the sequence's search tree). Thus, we only need to focus on the upper bound. To prove the upper bound of  $O(\sqrt{n/\epsilon} + \log n)$  on the smoothed height of binary search trees, we need some preparations. In the next subsection we introduce the concept of *increasing and decreasing runs* and show how they are related to binary search tree height. As we will see, bounding the length of these runs implicitly bounds the height of binary search trees. This allows us to prove the upper bound on the smoothed height of binary search trees in the main part of this section.

#### 4.1 Increasing and Decreasing Runs

In order to analyze the smoothed height of binary search trees, we introduce a related measure for which an upper bound is easier to obtain. Given a sequence  $s$ , consider a root-to-leaf path of the tree  $T(s)$ . We extract two subsequences  $\alpha = (\alpha_1, \dots, \alpha_k)$  and  $\beta = (\beta_1, \dots, \beta_\ell)$  from this path according to the following algorithm: We start at the root. When we are at an element  $s_i$  of the path, we look at the direction in which the path continues from  $s_i$ . If it continues with the right child of  $s_i$ , we append  $s_i$  to  $\alpha$ ; if it continues with the left child, we append  $s_i$  to  $\beta$ ; and if  $s_i$  is a leaf (has no children), then we append  $s_i$  to both  $\alpha$  and  $\beta$ . This construction ensures  $\alpha_1 < \dots < \alpha_k = \beta_\ell < \dots < \beta_1$  and the length of  $s$  is  $k + \ell - 1$ . Figure 2 shows an example of how  $\alpha$  and  $\beta$  are constructed.

A crucial property of the sequence  $\alpha$  is the following: Let  $\alpha_i = s_{j_i}$  for all  $i \in \{1, \dots, k\}$  with  $j_1 < j_2 < \dots < j_k$ . Then none of  $s_1, \dots, s_{j_i-1}$  lies in the interval  $(\alpha_i, \alpha_{i+1})$ , for otherwise  $\alpha_i$  and  $\alpha_{i+1}$  cannot be on the same root-to-leaf path. A similar property holds for the sequence  $\beta$ : No element of  $s$  prior to  $\beta_i$  lies in the interval  $(\beta_{i+1}, \beta_i)$ . We introduce a special name for sequences with this property.

**Definition 4.2.** An increasing run of a sequence  $s$  is a subsequence  $(s_{i_1}, s_{i_2}, \dots, s_{i_k})$ ,  $s_{i_1} < \dots < s_{i_k}$  with the following property: For every  $j \in \{1, \dots, k-1\}$ , no element of  $s$  prior to  $s_{i_j}$  lies in the interval  $(s_{i_j}, s_{i_{j+1}})$ . Analogously, a decreasing run of  $s$  is a subsequence  $(s_{i_1}, \dots, s_{i_\ell})$  with  $s_{i_1} > \dots > s_{i_\ell}$  such that no element prior to  $s_{i_j}$  lies in the interval  $(s_{i_{j+1}}, s_{i_j})$ .

Let  $\text{inc}(s)$  and  $\text{dec}(s)$  denote the length of the longest increasing and decreasing run of  $s$ , respectively. Furthermore, let  $\text{dec}_{\text{U}[\epsilon]}(s)$  and  $\text{inc}_{\text{U}[\epsilon]}(s)$  denote the length of the longest runs under uniform noise  $\text{U}[\epsilon]$ . In Figure 2, we have  $\text{inc}(s) = 4$  because of  $(7, 8, 10, 12)$  or  $(7, 8, 13, 14)$  and  $\text{dec}(s) = 4$  because of  $(7, 3, 2, 1)$ .

Since every root-to-leaf path can be divided into an increasing and a decreasing run, we immediately obtain the following lemma.

**Lemma 4.3.** *For every sequence  $s$  and all  $\epsilon$ , we have*

$$\begin{aligned} \text{height}(s) &\leq \text{dec}(s) + \text{inc}(s), \\ \mathbb{E}(\text{height}_{\text{U}[\epsilon]}(s)) &\leq \mathbb{E}(\text{dec}_{\text{U}[\epsilon]}(s) + \text{inc}_{\text{U}[\epsilon]}(s)). \end{aligned}$$

In terms of upper bounds,  $\text{dec}(s)$  and  $\text{inc}(s)$  as well as  $\text{dec}_{\text{U}[\epsilon]}(s)$  and  $\text{inc}_{\text{U}[\epsilon]}(s)$  behave equally. The reason is that given a sequence  $s$ , the sequence  $s'$  with  $s'_i = 1 - s_i$  has the properties  $\text{dec}(s) = \text{inc}(s')$  and  $\mathbb{E}(\text{dec}_{\text{U}[\epsilon]}(s)) = \mathbb{E}(\text{inc}_{\text{U}[\epsilon]}(s'))$ . This observation together with Lemma 4.3 proves the next lemma.

**Lemma 4.4.** *For all  $\epsilon$ , we have*

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{height}_{\text{U}[\epsilon]}(s)) \leq 2 \cdot \max_{s \in [0,1]^n} \mathbb{E}(\text{inc}_{\text{U}[\epsilon]}(s)).$$

The lemma states that in order to bound the smoothed height of search trees from above we can instead bound the smoothed length of increasing or decreasing runs. To simplify the analysis even further, we show that we can once more restrict our attention to sorted sequences.

**Lemma 4.5.** *For every sequence  $s$  and its sorted version  $\hat{s}$ , we have*

$$\mathbb{E}(\text{inc}_{\text{U}[\epsilon]}(s)) \leq \mathbb{E}(\text{inc}_{\text{U}[\epsilon]}(\hat{s})).$$

*Proof.* We sort  $s$  successively as we already did to prove Lemma 3.2. Assume that  $s_i > s_{i+1}$  for some  $i$  and let  $\delta = s_i - s_{i+1} > 0$ . We show  $\mathbb{E}(\text{inc}_{\text{U}[\epsilon]}(s)) \leq \mathbb{E}(\text{inc}_{\text{U}[\epsilon]}(s'))$  where  $s'$  is obtained from  $s$  by swapping  $s_i$  and  $s_{i+1}$ . Let  $\nu$  denote the noise vector added to  $s$  and  $s'$ . Then  $\bar{s}'_i = s'_i + \nu_{i+1} = s_{i+1} + \nu_{i+1}$  and  $\bar{s}'_{i+1} = s'_{i+1} + \nu_i = s_i + \nu_i$ .

We distinguish two cases. First, we condition on  $\nu_i \in [-\epsilon, \epsilon - \delta]$  and  $\nu_{i+1} \in [\delta - \epsilon, \epsilon]$ . Similar to the argument in Lemma 3.2, both  $(\bar{s}_i, \bar{s}_{i+1})$  and  $(\bar{s}'_i, \bar{s}'_{i+1})$  are pairs of random numbers, all of which lie uniformly in the interval  $[s_i - \epsilon, s_{i+1} + \epsilon]$ , and the expected values of  $\text{inc}_{\text{U}[\epsilon]}(s)$  and  $\text{inc}_{\text{U}[\epsilon]}(s')$  are equal.

Second, we condition on the events that  $\nu_i \in (\epsilon - \delta, \epsilon]$  or  $\nu_{i+1} \in [-\epsilon, \delta - \epsilon)$ . In either case,  $\bar{s}_i > \bar{s}_{i+1}$  and  $\bar{s}'_i < \bar{s}'_{i+1}$ . Thus, every increasing run of  $\bar{s}$  corresponds to an increasing run of  $\bar{s}'$ : If the run of  $\bar{s}$  uses neither  $\bar{s}_i$  nor  $\bar{s}_{i+1}$ , this is obvious. If the run of  $\bar{s}$  uses  $\bar{s}_i$ , then we get the same run of  $\bar{s}'$ , where now  $\bar{s}'_{i+1}$  is used. The run cannot be interrupted by  $\bar{s}'_i$  because  $\bar{s}'_i < \bar{s}'_{i+1}$ . If the run of  $\bar{s}$  uses  $\bar{s}_{i+1}$ , then we obtain a run of the same length using  $\bar{s}'_i$ . This run is also an increasing run since the only difference of  $\bar{s}$  and  $\bar{s}'$  is that now the larger element  $\bar{s}'_{i+1}$  appears after  $\bar{s}'_i$ . Finally, the run of  $\bar{s}$  cannot use both  $\bar{s}_i$  and  $\bar{s}_{i+1}$  because  $\bar{s}_{i+1} < \bar{s}_i$ . Thus, we have  $\text{inc}(\bar{s}) \leq \text{inc}(\bar{s}')$ , which proves the lemma.  $\square$

## 4.2 Upper Bound on the Smoothed Height of Binary Search Trees

In this section we prove an upper bound on the height of binary search trees by proving an upper bound on the smoothed length of increasing runs. Throughout this section we will assume that the input sequence  $s$  is sorted. The following lemma gives an upper bound for the case that all input elements initially (i. e., before perturbation) lie within an interval of length  $\delta$  and the perturbation parameter  $\epsilon$  is rather low.



**Lemma 4.6.** *Given an input sequence  $s$  with start values from an interval of length  $\delta$  and a perturbation parameter  $\epsilon \leq \delta$  we have*

$$\mathbb{E}(\text{inc}_{\cup[\epsilon]}(s)) \in O(\sqrt{\delta n/\epsilon}).$$

*Proof.* Let us take a closer look at increasing runs. Assume that  $\alpha_1, \dots, \alpha_\ell$  is a longest increasing run of  $\bar{s}$  and let  $\alpha_k$  denote the first element in this sequence that is not a left-to-right maximum of  $\bar{s}$ . We can bound the length of the initial portion  $\alpha_1, \dots, \alpha_{k-1}$  of the run by using our results about left-to-right maxima. After scaling, Theorem 3.1 says that  $\mathbb{E}(\text{ltrm}_{\cup[\epsilon]}(s)) \in O(\sqrt{\delta n/\epsilon})$  which means that the length of the initial portion is bounded by the same quantity.

In the following we derive a bound on the length of the second part, the sequence  $\alpha_k, \dots, \alpha_\ell$ . We show that with high probability this length is at most  $O(\sqrt{\delta n/\epsilon})$ , as well, which gives the lemma.

For an element  $\alpha_i$  with  $k \leq i \leq \ell$  in this run, let  $\hat{\alpha}_i$  denote the smallest predecessor of  $\alpha_i$  that is larger than  $\alpha_i$ . This means  $\hat{\alpha}_i = \min\{\bar{s}_t \mid \bar{s}_t > \alpha_i, t \leq j_i\}$ , where  $j_i$  denotes the position of  $\alpha_i$  in the sequence  $\bar{s}$ .

Let  $i \geq k$ . Note that all elements  $\alpha_r$ ,  $r > i$  of the increasing run  $\alpha_k, \dots, \alpha_\ell$  must lie in the interval  $[\alpha_i, \hat{\alpha}_i]$ . We call this interval the *restriction interval* defined by  $\alpha_i$ . Observe that this interval only depends on random choices for elements  $s_t$ ,  $t \leq j_i$ . In the following we show that with high probability the length of this restriction interval shrinks a lot after seeing  $\sqrt{\delta n/\epsilon}$  more elements.

**Claim 4.7.** *Let  $[\alpha_i, \hat{\alpha}_i]$  denote a restriction interval, let  $\gamma_i = \hat{\alpha}_i - \alpha_i$  denote its length and let  $z = \sqrt{\delta n/\epsilon}$ . Then with high probability the length  $\eta$  of the restriction interval  $[\alpha_{i+z+1}, \hat{\alpha}_{i+z+1}]$  is at most  $\eta \leq \sqrt[4]{\epsilon/(\delta n)} \cdot \gamma_i$ .*

*Proof.* Note that this claim holds regardless of the choice of  $\alpha_{i+1}, \dots, \alpha_{i+z+1}$ .

Let  $j_i$  denote the index of  $\alpha_i$  in  $\bar{s}$ , and let  $j_{i+1} < j_{i+2} < \dots < j_{i+z}$  denote the next  $z$  positions in  $\bar{s}$  for which the corresponding element  $s_r$  has a perturbed value  $\bar{s}_r \in [\alpha_i, \hat{\alpha}_i]$ . Let  $j_{i+z+1}$  denote the position of the element  $\alpha_{i+z+1}$ . Note that  $j_{i+z+1} > j_{i+z}$  since the elements  $\alpha_{i+1}, \dots, \alpha_{i+z}$  must all fall into the interval  $[\alpha_i, \hat{\alpha}_i]$ .

Since the input sequence is sorted,  $\alpha_{i+z+1}$  must lie in the interval  $[\max\{\alpha_i, s_{i+z} - \epsilon\}, \hat{\alpha}_i]$  (regardless of the exact choice of  $\alpha_{i+z+1}$ ). We now show that with high probability the elements  $\bar{s}_{j_{i+1}}, \dots, \bar{s}_{j_{i+z}}$  split this interval into sub-intervals of length at most  $\eta$ . Then regardless of our choice of  $\alpha_{i+z+1}$ , we have  $\hat{\alpha}_{i+z+1} - \alpha_{i+z+1} \leq \eta$  and the claim follows.

Let  $\beta = \max\{\alpha_i, s_{i+z} - \epsilon\}$ , and let  $\gamma'_i = \hat{\alpha}_i - \beta < \gamma_i$ . We define  $\lfloor 2\gamma'_i/\eta \rfloor$  sub-intervals  $I_1, \dots, I_{\lfloor 2\gamma'_i/\eta \rfloor}$  of  $[\beta, \hat{\alpha}_i]$  by defining  $I_r := [\beta + (r-1)\frac{\eta}{2}, \beta + r\frac{\eta}{2}]$ . Each of the sub-intervals has length  $\eta/2$ . If every sub-interval contains at least one element from  $\bar{s}_{j_{i+1}}, \dots, \bar{s}_{j_{i+z}}$ , then  $[\beta, \hat{\alpha}_i]$  gets split into sub-intervals of length at most  $\eta$  and the claim follows.

Fix a sub-interval  $I_r$ . Note that each element  $s_{j_t}$  has a probability of at least  $\eta/(2\gamma_i)$  of falling into  $I_r$  (recall that we conditioned on the fact that  $\bar{s}_{j_t} \in [\alpha_i, \hat{\alpha}_i]$ ). Hence, the probability that  $I_r$  does not contain an element from  $\bar{s}_{j_{i+1}}, \dots, \bar{s}_{j_{i+z}}$  is at most

$$\left(1 - \frac{\eta}{\gamma_i}\right)^z \leq \exp\left(-\frac{\eta}{\gamma_i} \cdot z\right) \leq \exp\left(-\sqrt[4]{\delta n/\epsilon}\right).$$

Applying a union bound we obtain that the probability that there exists an empty set is at most  $\lfloor 2\gamma'_i/\eta \rfloor \exp(-\sqrt[4]{\delta n/\epsilon}) = \exp(\Theta(-\sqrt[4]{\delta n/\epsilon}))$ .  $\square$

Now, we use Claim 4.7 to finish the proof of Lemma 4.6. We assume that all high probability events actually occur. If this is not the case we can bound the number of elements in an increasing run by  $n$ . Because this case is so unlikely its effect on the expectation is negligible.

We want to estimate the length of the run  $\alpha_k, \dots, \alpha_\ell$ . Let  $z = \sqrt{\delta n/\epsilon}$  as in the above claim. The length of the restriction interval defined by  $\alpha_k$  is at most  $\delta + \epsilon \leq 2\delta$ . Assume that the run is longer than  $2z + 2$ . Then by applying Claim 4.7 twice we obtain that the restriction interval  $[\alpha_{k+z+2}, \hat{\alpha}_{k+z+2}]$  has length at most

$$\sqrt[4]{\epsilon/(\delta n)} \cdot \sqrt[4]{\epsilon/(\delta n)} \cdot 2\delta \leq 2\sqrt{\epsilon\delta/n}.$$

The probability for an element to fall into this interval is at most  $2\delta \leq 2\sqrt{\epsilon\delta/n} \cdot 2\frac{1}{\epsilon} = 4\sqrt{\delta/(\epsilon n)}$ . In expectation the total number of elements in this interval is only  $4\sqrt{\delta n/\epsilon}$ . Applying a Chernoff bound gives that with high probability the interval contains no more than  $8\sqrt{\delta n/\epsilon}$  elements.

Therefore, even if all elements that fall into the interval belong to the run  $\alpha_{2z+2}, \dots, \alpha_\ell$  the length of the run  $\alpha_k, \dots, \alpha_\ell$  is still at most  $O(\sqrt{\delta n \epsilon})$  with high probability. Together with the bound on the length of the initial sequence  $\alpha_1, \dots, \alpha_k$  this concludes the proof of the lemma.  $\square$

Before generalizing Lemma 4.6 to a wider range of  $\epsilon$  we require the following simple observation.

**Lemma 4.8.** *For every sequence  $s$ , all  $\epsilon$ , and every covering  $U_1, \dots, U_k$  of  $\{1, \dots, n\}$  (which means  $\bigcup_{i=1}^k U_i = \{1, \dots, n\}$ ), we have*

$$\begin{aligned} \text{height}(s) &\leq \sum_{i=1}^k \text{height}(s_{U_i}), \\ \mathbb{E}(\text{height}_{U[\epsilon]}(s)) &\leq \sum_{i=1}^k \mathbb{E}(\text{height}_{U[\epsilon]}(s_{U_i})). \end{aligned}$$

*Proof.* Let  $U_1, \dots, U_k$  cover  $\{1, \dots, n\}$ . For a fixed  $i$ , let  $a$  and  $b$  with  $a < b$  be two elements of  $s_{U_i}$  that do not lie on the same root-to-leaf path in  $T(s_{U_i})$ . Then there exists a  $c$  prior to  $a$  and  $b$  in  $s_{U_i}$  with  $a < c < b$ , which implies that  $a$  and  $b$  do not lie on the same root-to-leaf path in the tree  $T(s)$  either. Now consider a root-to-leaf path  $p$  of  $T(s)$  that has a length of  $\text{height}(s)$ . Let  $p_{U_i}$  be  $p$  restricted to elements of  $s_{U_i}$  and let  $\ell_{U_i}$  be its length. Then  $\sum_{i=1}^k \text{height}(s_{U_i}) \geq \sum_{i=1}^k \ell_{U_i} \geq \text{height}(s)$ , because the  $U_i$  cover  $\{1, \dots, n\}$ .

The second inequality follows directly from the first since taking expectations is a monotone operation.  $\square$

Now, we turn our attention to the case where  $\epsilon$  is relatively large, namely  $\epsilon = (n/\log^2 n) \cdot \delta$ . This case is critical since our overall goal is to show a bound of  $O(\sqrt{\delta n/\epsilon} + \log n)$ . For  $\epsilon = (n/\log^2 n) \cdot \delta$  this becomes  $O(\log n)$ . Therefore, high probability arguments of the form used to prove Lemma 4.6 are not sufficient.

**Lemma 4.9.** *Let  $s$  denote an input sequence of  $n$  elements with values from an interval of length  $\delta$  and let  $\epsilon \geq (n/\log^2 n) \cdot \delta$ . Then*

$$\mathbb{E}(\text{height}_{U[\epsilon]}(s)) \in O(\log n).$$

*Proof.* For simplicity we assume that all unperturbed values lie in the interval  $[0, \delta]$ . Note that the perturbed value  $\bar{s}_i$  of an element  $s_i$  from  $s$  can range over every value from the interval

$I := [-\epsilon + \delta, \epsilon] \subseteq [-\epsilon, \epsilon + \delta]$ . Furthermore, conditioned on the fact that  $\bar{s}_i \in I$ , it will be uniformly distributed within  $I$ .

We partition the elements of  $s$  into three not necessarily disjoint sets  $R$ ,  $F$ , and  $B$ . The set  $R$  of *regular positions* contains the positions of elements  $s_i$  that after perturbation lie in interval  $I$ . The set  $F$  contains positions of elements  $s_i$  for which the perturbation  $\nu_i$  is at most  $\delta - \epsilon$  and  $B$  contains positions of elements  $s_i$  with  $\nu_i \geq \epsilon - \delta$ . Note that every element of  $s_i$  contributes to at least one set because for an element not to end up in interval  $I$  (and hence its position *not* added to set  $R$ ) its perturbation must be either very large (at least  $\epsilon - \delta$ ) or very small (at most  $\delta - \epsilon$ ). Therefore we can obtain our result by bounding  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(\bar{s}_R))$ ,  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(\bar{s}_F))$  and  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(\bar{s}_B))$  by  $O(\log n)$  and applying Lemma 4.8.

Let us start with  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(\bar{s}_R))$ . Conditioned on the fact that  $\bar{s}_i \in I$  the element  $\bar{s}_i$  is uniformly distributed in this interval. Hence, every permutation of the elements of  $\bar{s}_R$  is equally likely which reduces  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(\bar{s}_R))$  to the average case. Hence,  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(\bar{s}_R)) \in O(\log n)$ .

It remains to deal with the sets  $F$  and  $B$ . We only consider  $F$  as the proof for  $B$  is analogous. Only conditioning on the fact that an element  $\bar{s}_i$  is in set  $F$  means that  $\nu_i \leq \delta - \epsilon$ . Within the interval  $[\delta - \epsilon, -\epsilon]$ ,  $\nu_i$  is distributed uniformly. One can view the perturbation of elements in  $F$  as being generated in a two step process. First the elements are moved down by  $\frac{\delta}{2} - \epsilon$  and then a uniform perturbation in the interval  $[-\delta/2, \delta/2]$  is added. Since after the first step all elements of  $F$  still lie in an interval of length  $\delta$  the second step fulfills the requirement for Lemma 4.6. Therefore,  $\mathbb{E}(\text{height}(\bar{s}_F)) \in O(\sqrt{2|F|})$ .

The probability for an element of  $s$  to end up in  $F$  is only  $\delta/\epsilon = \frac{\log^2 n}{n}$ . Using a Chernoff bound gives that with high probability (at least  $1 - n^{-(\log n)/3}$ )  $F$  contains at most  $2 \log^2 n$  elements, and hence  $\mathbb{E}(\text{height}(\bar{s}_F)) \in O(\log n)$ . In the case that  $F$  contains more elements, we can bound  $|F|$  by  $n$ . This case only contributes  $n \cdot n^{-(\log n)/3}$  to the expectation which is negligible.

This completes the proof of the lemma.  $\square$

**Lemma 4.10.** *For an input sequence  $s$  with values from the unit interval, we have*

$$\mathbb{E}(\text{height}_{\cup[\epsilon]}(s)) \in O(\sqrt{n/\epsilon} + \log n).$$

*Proof.* If  $\epsilon \geq n/\log^2 n$  then  $\mathbb{E}(\text{ltrim}_{\cup[\epsilon]}(s)) \in O(\log n)$  by Lemma 4.9. In order to prove the lemma for smaller values of  $\epsilon$ , we partition the sequence into subsequences of size at most  $N$ , where  $N$  will be chosen later. For each of the subsequences we apply Lemma 4.9 and sum the resulting expectations which, by Lemma 4.8, gives an upper bound on  $\mathbb{E}(\text{height}_{\cup[\epsilon]}(s))$ .

Suppose a subsequence contains only elements with values from an interval of length  $\delta$ . In order to be able to apply Lemma 4.9 we require  $\delta \leq (\frac{\log^2 N}{N}) \cdot \epsilon$ . In the following we choose  $\delta := (\frac{\log^2 N}{N}) \cdot \epsilon$  and partition the unit interval into  $\lceil 1/\delta \rceil$  sub-intervals of length at most  $\delta$ . Let  $S_j$  for  $j \in \{1, \dots, \lceil 1/\delta \rceil\}$  denote the positions in the input sequence  $s$  that lie in the  $j$ th interval. We partition  $S_j$  into  $\lceil |S_j|/N \rceil$  subsets of size at most  $N$  in an arbitrary manner and obtain  $\lceil |S_j|/N \rceil$  subsequences for which we can apply Lemma 4.9.

In total we constructed at most

$$\sum_{j=1}^{\lceil 1/\delta \rceil} \lceil |S_j|/N \rceil \leq \lceil 1/\delta \rceil + \sum_{j=1}^{\lceil 1/\delta \rceil} |S_j|/N = \left\lceil \frac{N}{\epsilon \cdot \log^2 N} \right\rceil + n/N$$

subsequences. Now, we choose  $N$  so as to approximately balance the two terms in the above equation. This means we choose  $N$  such that  $N^2/\log^2 N = \epsilon n$ . Then the total number of subsequences is  $O(n/N)$ .

Finally, we can apply Lemma 4.9 to each subsequence and sum the expectation of  $O(\log N)$  from each sequence to get a bound on  $\mathbb{E}(\text{height}_{U[\epsilon]}(s))$ . This gives

$$\mathbb{E}(\text{height}_{U[\epsilon]}(s)) \in O\left(\frac{n \log N}{N}\right) = O\left(\sqrt{n/\epsilon}\right).$$

□

## 5 Smoothed Number of Quicksort Comparisons

In this section we apply our results on binary search trees and left-to-right maxima under uniform noise to the performance of the quicksort algorithm under uniform noise. The following theorem summarizes the findings.

**Theorem 5.1.** *For  $\epsilon \geq 1/n$  we have*

$$\max_{s \in [0,1]^n} \mathbb{E}(\text{qs}_{U[\epsilon]}(s)) \in \Theta\left(\frac{n}{\epsilon+1} \sqrt{n/\epsilon} + n \log n\right).$$

In other words, for  $\epsilon \in O(1)$ , the number of comparisons is at most  $O(n\sqrt{n/\epsilon})$ , while for  $\epsilon \in \Omega(1)$ , it is at most  $O(\frac{n}{\epsilon} \sqrt{n/\epsilon})$ . This means that  $\epsilon$  has a stronger influence for  $\epsilon \in \Omega(1)$ .

### 5.1 Upper Bound on the Smoothed Number of Quicksort Comparisons

To prove the upper bound, we first need a lemma similar to Lemma 4.8 that allows us to estimate the number of comparisons of subsequences.

**Lemma 5.2.** *For every sequence  $s$ , all  $\epsilon$ , and every covering  $U_1, \dots, U_k$  of  $\{1, \dots, n\}$ , we have*

$$\begin{aligned} \text{qs}(s) &\leq \sum_{i=1}^k \text{qs}(s_{U_i}) + Q, \\ \mathbb{E}(\text{qs}_{U[\epsilon]}(s)) &= \mathbb{E}(\text{qs}(\bar{s})) \leq \sum_{i=1}^k \mathbb{E}(\text{qs}(\bar{s}_{U_i})) + \mathbb{E}(\bar{Q}), \end{aligned}$$

where  $Q$  is the number of comparisons of elements of  $s_{U_i}$  with elements of  $s_{\{1, \dots, n\} \setminus U_i}$  for any  $i$  and the random variable  $\bar{Q}$  is defined analogously for  $\bar{s}$ .

The proof goes along the same lines as the proof of Lemma 4.8 and is omitted.

**Lemma 5.3.** *For every sequence  $s$  and all  $\epsilon \geq 1/n$ , we have*

$$\mathbb{E}(\text{qs}_{U[\epsilon]}(s)) \in O\left(\frac{n}{\epsilon+1} \sqrt{n/\epsilon} + n \log n\right).$$

*Proof.* Given a sequence  $s$ , first observe that quicksort will make at most  $O(n\sqrt{n/\epsilon} + n \log n)$  comparisons, which follows directly from Lemma 4.10 and  $\mathbb{E}(\text{qs}(\bar{s})) \leq n \cdot \mathbb{E}(\text{height}(\bar{s}))$ : Every level of recursion of quicksort contributes at most  $n - 1$  comparisons, and we have  $\text{height}(\bar{s})$  levels of recursion. Thus, the claim of the theorem is correct for  $\epsilon \in O(1)$ .

Let us now consider the case  $\epsilon \in \omega(1)$ . Furthermore, assume that  $\epsilon \in O(\sqrt[3]{n/\log^2 n})$ . This is no restriction since we obtain the average-case bound of  $O(n \log n)$  already for  $\epsilon \in \Theta(\sqrt[3]{n/\log^2 n})$  and thus also for larger  $\epsilon$ .

Similar to the proof of Lemma 4.9, we divide the sequence  $\bar{s}$  into three parts. The set  $R = \{i \in \{1, \dots, n\} \mid \bar{s}_i \in [1 - \epsilon, \epsilon]\}$  of regular elements for the interval  $[1 - \epsilon, \epsilon]$  is defined as before. The set  $F$  is defined slightly differently, namely as  $F = \{i \in \{1, \dots, n\} \mid \nu_i \leq 4 - \epsilon\}$ . This means that  $F$  contains all  $i$  for which  $\nu_i$  is too small, plus some extra elements. Similarly  $B = \{i \in \{1, \dots, n\} \mid \nu_i \geq \epsilon - 4\}$ .

As in Lemma 4.9, the regular elements are easy to handle since they are uniformly distributed in  $[1 - \epsilon, \epsilon]$  and, thus,  $\mathbb{E}(\text{qs}_{U[\epsilon]}(s_R)) \in O(n \log n)$ .

We have  $\mathbb{E}(\text{height}_{U[\epsilon]}(s_F)) = \mathbb{E}(\text{height}_{U[\epsilon]}(s_B)) \in O(\sqrt{n/\epsilon})$ , which follows from the same scaling argument (see Lemma 2.1) that we used in Lemma 4.9:  $F$  contains  $2n/\epsilon$  elements in expectation. The probability that  $F$  contains more than  $4n/\epsilon$  elements is at most  $(e/4)^{2n/\epsilon} \in O((e/4)^{\sqrt{n}})$  due to the Chernoff bound and  $\epsilon \in O(\sqrt[3]{n/\log^2 n})$ . The same holds for  $B$ . If either contains more elements, we bound the height by  $n$ , which contributes at most  $o(1)$  to the expectation. Otherwise, we have sequences with  $O(n/\epsilon)$  elements that are perturbed with  $U[2]$ . We obtain

$$\mathbb{E}(\text{qs}(\bar{s}_F)) = \mathbb{E}(\text{qs}(\bar{s}_B)) \in O\left(\frac{n}{\epsilon} \sqrt{n/\epsilon}\right),$$

which is just the number of elements multiplied with the upper bound for the tree height.

By Lemma 5.2, what remains to be estimated is the number of comparisons of elements  $\bar{s}_i$  and  $\bar{s}_j$  where  $i$  and  $j$  are in two different sets of  $R$ ,  $F$ , and  $B$ .

Due to the symmetry between  $\bar{s}_F$  and  $\bar{s}_B$ , it suffices to restrict ourselves to estimating the number of comparisons of elements in  $\bar{s}_F$  with elements in  $\bar{s}_R$  and  $\bar{s}_B$ . This boils down to counting the number of comparisons of elements  $\bar{s}_i$  with  $\nu_i \leq \epsilon - 4$  to elements  $\bar{s}_j$  with  $\bar{s}_j \geq 1$ .

The number of comparisons between elements  $\bar{s}_i$  and  $\bar{s}_j$  with  $i \in F$  and  $j \in F \cap R$  can be bounded by the total number of comparisons between elements in  $F$ , but this number is  $\mathbb{E}(\text{qs}(\bar{s}_F)) \in O\left(\frac{n}{\epsilon} \sqrt{n/\epsilon}\right)$ . Similarly, since  $\mathbb{E}(\text{qs}(\bar{s}_R)) \in O(n \log n)$ , the expected number of comparisons between positions  $i \in F \cap R$  and  $j \in R$  is at most  $O(n \log n)$ .

Thus, we can concentrate on  $i \in F$  with  $\nu_i \leq 1 - \epsilon$  and  $j \in R$  with  $\bar{s}_j \geq \epsilon - 4$ , which includes all  $i \in F \setminus R$  and  $j \in R \setminus F$ .

We distinguish two cases: First, we estimate the expected number of such comparisons with  $\bar{s}_i$  being the pivot element. Second, we consider the case that  $\bar{s}_j$  is the pivot element.

The two elements  $\bar{s}_i \leq s_i - \epsilon + 1 \leq 2 - \epsilon$  and  $\bar{s}_j \geq 4 - \epsilon$  will be compared with  $\bar{s}_i$  being the pivot only if  $i < j$  and  $\bar{s}$  contains no element  $\bar{s}_k \in [\bar{s}_i, \bar{s}_j]$  for  $k < i$ . In particular,  $\bar{s}$  must not contain an element  $\bar{s}_k \in [2 - \epsilon, 4 - \epsilon]$  with  $k < i$ .

Since  $\epsilon \in \omega(1)$ , every element is eligible for the interval  $[2 - \epsilon, 4 - \epsilon]$ . Furthermore, for every  $i \in \{1, \dots, n\}$ , we have  $\mathbb{P}(\nu_i \leq 1) = \mathbb{P}(\bar{s}_k \in [2 - \epsilon, 4 - \epsilon]) = 1/\epsilon$  and these two events are disjoint. (If  $s_i = 1$ , then this is not true since  $\nu_i = 1$  is possible, but the probability of this is 0.) Thus, the probability that  $\bar{s}$  contains more than  $O(\log n)$  elements with  $\nu_i \leq 1$  prior to the first element  $\bar{s}_k \in [2 - \epsilon, 4 - \epsilon]$  is  $O(1/n)$ . If this happens nevertheless, we bound the number of comparisons by the trivial upper bound of  $n^2$ , which contributes only  $O(n^2 \cdot 1/n) = O(n)$  to the expected value.

Otherwise, at most  $O(\log n)$  elements  $\bar{s}_i$  with  $\nu_i \leq 1 - \epsilon$  are compared to elements  $\bar{s}_j$  with  $\bar{s}_j \geq 4 - \epsilon$  with  $\bar{s}_i$  being the pivot, which contributes  $O(n \log n)$  comparisons.

Now we consider the second case: How many comparisons of elements  $\bar{s}_j \geq 4 - \epsilon$  with elements  $\bar{s}_i \leq s_i + 1 - \epsilon$  with  $\bar{s}_j$  being the pivot element do we have to expect? The element  $\bar{s}_j$  is compared to  $\bar{s}_i$  only if  $j < i$  and there is no  $k < j$  with  $\bar{s}_k \in [\bar{s}_i, \bar{s}_j]$ . Thus, it is necessary that  $\bar{s}_j$  is minimal among all elements  $\bar{s}_k \geq 4 - \epsilon$  with  $k \leq j$ .

If we restrict ourselves to  $\bar{s}_k \in [4 - \epsilon, \epsilon]$ , then this corresponds just to the average number of left-to-right *minima*, which is  $O(\log n)$ . (The average number of left-to-right minima is equal to the average number of left-to-right maxima.) Thus, the expected number of elements  $\bar{s}_j \in [4 - \epsilon, \epsilon]$  that, when being the pivot element, are compared to any element  $\bar{s}_i \leq s_i + 1 - \epsilon$ , is  $O(\log n)$ . This contributes at most  $O(n \log n)$  to the expected number of comparisons.

Elements  $\bar{s}_k \geq \epsilon$  remain to be considered. Since  $\epsilon \in \omega(1)$ , there are at most  $O(\log n)$  such elements prior to the first element of the interval  $[4 - \epsilon, \epsilon]$  with high probability. Furthermore, there are at most  $O(\log n)$  elements of  $\bar{s}_F$  prior to the first element of  $[1 - \epsilon, \epsilon]$  with high probability. Thus, the contribution to the number of comparisons is only  $O(\log^2 n)$ .  $\square$

## 5.2 Lower Bound on the Smoothed Number of Quicksort Comparisons

The upper bound proved in the previous section is tight. The standard sorted sequence provides a worst case, but in the following lemma we use a sequence that is slightly easier to handle technically.

**Lemma 5.4.** *For the sequence  $s = (1/n, 2/n, 3/n, \dots, \frac{n}{2}/n, 1, 1, \dots, 1)$  and all  $\epsilon \geq 1/n$ , we have*

$$\mathbb{E}(\text{qs}_{U[\epsilon]}(s)) \in \Omega\left(\frac{n}{\epsilon+1}\sqrt{n/\epsilon} + n \log n\right).$$

*Proof.* In the perturbed sequence  $\bar{s}$  the first  $n/2$  elements contain an expected number of  $\Omega(\sqrt{n/\epsilon})$  left-to-right maxima according to Theorem 3.1. Every left-to-right maximum  $\bar{s}_i$  of  $\bar{s}$  has to be compared to all the elements that come later and are greater than  $\bar{s}_i$ .

If  $\epsilon \in o(1)$ , all  $n/2$  elements of the second half of  $\bar{s}$  are greater than any left-to-right maximum of the first half of  $\bar{s}$ . Thus, the expected number of comparisons is at least  $\Omega(n\sqrt{n/\epsilon}) = \Omega\left(\frac{n}{\epsilon+1}\sqrt{n/\epsilon} + n \log n\right)$ .

If  $\epsilon \in \Omega(1)$ , then the probability that an element  $\bar{s}_i$  of the second half of  $\bar{s}$  is greater than all left-to-right maxima of the first half of  $\bar{s}$  is

$$\mathbb{P}(\forall j \leq n/2: 1 + \nu_i \geq \bar{s}_j) \geq \mathbb{P}(1 + \nu_i \geq 1/2 + \epsilon) = \frac{1}{4\epsilon}.$$

Thus, the expected number of elements that are greater than all left-to-right maxima of the first half is at least the expected number of elements positions  $i > n/2$  with  $\nu_i \geq \epsilon + 1/2$ , which is  $\Omega(n/\epsilon)$ . The latter number is independent of the number of left-to-right maxima in the first half, so we can multiply the two expected values to get a lower bound of  $\Omega\left(\frac{n}{\epsilon}\sqrt{n/\epsilon}\right) \subseteq \Omega\left(\frac{n}{\epsilon+1}\sqrt{n/\epsilon}\right)$  comparisons. Since quicksort always needs at least  $\Omega(n \log n)$  comparisons, we get the claim.  $\square$

## 6 Smoothed Number of Points on a Convex Hull

In this section we study the smoothed number of points on a convex hull in two-dimensional space and apply our findings on left-to-right maxima. We model this problem as follows: The adversary chooses a set  $P = \{(x_1, y_1), \dots, (x_n, y_n)\} \subseteq [0, 1]^2$  in the plane. The coordinates of these points are then perturbed according to a certain probability distribution with density

function  $\phi$ , resulting in a set  $\bar{P} = \{(x_1 + \nu_1, y_1 + \nu'_1), \dots, (x_n + \nu_n, y_n + \nu'_n)\}$ . We are then interested in the expected number of points that lie on the convex hull of  $\bar{P}$  (the smallest convex polygon that includes all points of  $\bar{P}$ ). Letting  $\text{ch}(P) \subseteq P$  denote the set of points from  $P$  that lie on the convex hull of  $P$ , the *smoothed number of points on the convex hull of  $n$  points* is

$$\max_{P \subseteq [0,1]^2, |P|=n} \mathbb{E}(|\text{ch}_\phi(P)|).$$

## 6.1 Upper Bounds on the Smoothed Number of Points on a Convex Hull

We begin with upper bounds, which we derive from our bounds on the smoothed number of left-to-right maxima.

**Theorem 6.1.** *We have*

$$\begin{aligned} \max_{P \subseteq [0,1]^2, |P|=n} \mathbb{E}(|\text{ch}_{U[\epsilon]}(P)|) &\in O(\sqrt{n/\epsilon} + \log n), \\ \max_{P \subseteq [0,1]^2, |P|=n} \mathbb{E}(|\text{ch}_{N[\sigma]}(P)|) &\in O\left(\frac{\log^{3/2} n}{\sigma} + \log n\right). \end{aligned}$$

*Proof.* Given a set of points, its convex hull can be split into two parts: the upper part and the lower part, both of which start and end with the points having minimum and maximum  $x$ -coordinate, respectively. Both the upper and the lower part can be split into a left and a right part, where the split occurs at the points with the maximum or minimum  $y$ -coordinate, respectively. We focus on bounding the number of points in the upper left part of the convex hull, the total number of points on the convex hull is then at most four times this number. Consider the set of points in  $\bar{P}$ . If necessary, renumber the points so that  $x_1 + \nu_1 \leq x_2 + \nu_2 \leq \dots \leq x_n + \nu_n$ . Now, a *necessary* (though not sufficient) condition for a point  $(x_i, y_i)$  to be part of the upper left convex hull is that it is a left-to-right maximum of the sequence  $(y_1 + \nu'_1, \dots, y_n + \nu'_n)$ . But, then, the upper bounds on the number of left-to-right maxima on such a sequence from Theorems 3.1 and 3.5 yield the claim.  $\square$

## 6.2 Lower Bounds on the Smoothed Number of Points on a Convex Hull

Our lower bounds for the smoothed number of points on a two-dimensional convex hull turn out to be less sharp than the results we obtained in earlier sections for binary tree height and the quicksort algorithm.

**Theorem 6.2.** *We have*

$$\begin{aligned} \max_{P \subseteq [0,1]^2, |P|=n} \mathbb{E}(|\text{ch}_{U[\epsilon]}(P)|) &\in \Omega(\min\{\sqrt[5]{n}/\epsilon, n\}), \\ \max_{P \subseteq [0,1]^2, |P|=n} \mathbb{E}(|\text{ch}_{N[\sigma]}(P)|) &\in \Omega(\sqrt{\log n}). \end{aligned}$$

*Proof.* The second claim, for normally distribution noise, follows directly from the work of Rényi and Sulanke [20]. We now prove the lower bound for uniform noise.

Our aim is to construct an input of  $n$  points that has a large expected number of vertices on the convex hull after perturbation. Our model requires that the input points lie in  $[0, 1]^2$  and be perturbed in each direction by a value from the interval  $[-\epsilon, \epsilon]$ . However, for the purposes of the present proof it will be more natural to describe the construction for points from  $[-1, 1]^2$  perturbed in each direction by values from  $[-2\epsilon, 2\epsilon]$ , which is just a scaling by a factor of 2.

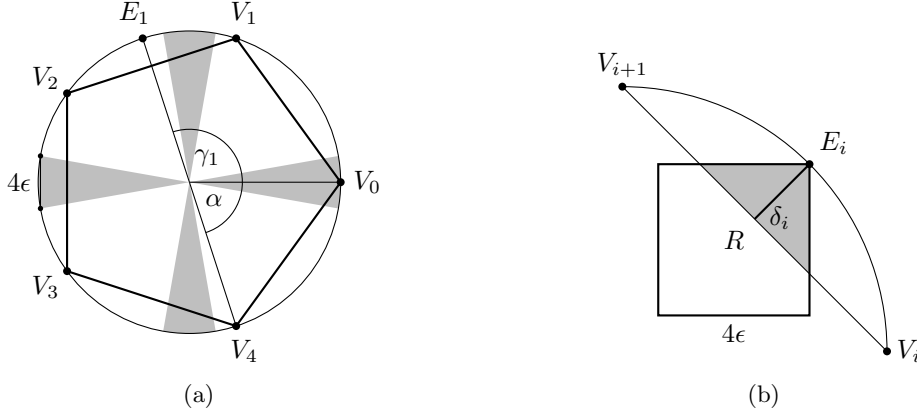


Figure 3: (a) The partitioning of the plane into different regions. If the extreme point  $E_i$  of a boundary region  $i$  falls into the shaded area the corresponding boundary region is not valid. (b) The situation where the intersection (in gray) between a boundary region and the corresponding range square  $R$  is minimal.

We inscribe an  $\ell$ -sided regular polygon into a unit circle centered at the origin. The number  $\ell$  will be fixed later. The interior of the polygon belongs to the *inner region* while everything outside the unit circle belongs to the *outer region*. Let  $V_0, \dots, V_{\ell-1}$  denote the vertices of the polygon. The  $i$ th *boundary region* is the segment of the unit circle defined by the chord  $V_iV_{i+1}$  where the indices are modulo  $\ell$ , see Figure 3(a). An important property of these regions is expressed in the following observation.

**Observation 6.3.** *If no point lies in the outer region, then every non-empty boundary region contains at least one point that is a vertex of the convex hull.*

*Proof.* If the outer region is empty, then the points in the  $i$ th boundary region can be separated from the remaining points by the straight line through  $V_i$  and  $V_{i+1}$ .  $\square$

In the following, we select the initial positions of the input points such that it is guaranteed that after the perturbation the outer region is empty and the expected number of non-empty boundary regions is large.

We need the following notations and definitions. For the  $j$ th input point we define the range square  $R$  to be the axis-parallel square with side length  $4\epsilon$  centered at position  $(x_j, y_j)$ . Note that for the uniform distribution by values from the interval  $[-2\epsilon, 2\epsilon]$  the perturbed position of  $(x_j, y_j)$  will lie in  $R$ . The intersection between the circle boundary and the perpendicular bisector of the chord  $V_iV_{i+1}$  is called the *extremal point* of boundary region  $i$  and is denoted with  $E_i$ . The line segment from the midpoint of the chord to  $E_i$  is denoted with  $\delta_i$ , see Figure 3(a).

The general outline for the proof is as follows. We try for a boundary region  $i$  to place a bunch of  $n/\ell$  input points in the plane such that a vertex of their common range square  $R$  lies in the extremal point  $E_i$  of the boundary region. Furthermore we require that no point of  $R$  lies in the outer region. If this is possible, it can be shown that the range square and the boundary region have a large intersection. Therefore it will be likely that one of the  $n/\ell$  input points corresponding to the square lies in the boundary region after perturbation. Then, we can derive a bound on the number of vertices in the convex hull by exploiting Observation 6.3, because we can guarantee that no perturbed point lies in the outer region.



Now, we formalize this idea. We call a boundary region  $i$  *valid* if we can place input points such that their range square  $R_i$  is contained in the unit circle and a vertex of it lies on  $E_i$ . Then  $R_i$  is called the *range square corresponding to boundary region  $i$* .

**Lemma 6.4.** *If  $\epsilon \leq 1/8$  and  $\ell \geq 23$ , then there are at least  $\ell/2$  valid boundary regions.*

*Proof.* Let  $\gamma_i$  denote the angle of vector  $E_i$  with respect to the positive  $x$ -axis. A boundary region is valid if, and only if,  $\sin \gamma_i \geq 2\epsilon$  and  $\cos \gamma_i \geq 2\epsilon$ . The invalid regions are depicted in Figure 3(a). If  $\epsilon \leq 1/8$ , these regions are small. To see this let  $\beta$  denote the central angle of each region. Then  $2 \sin(\beta/2) = 4\epsilon \leq 1/2$  and  $\beta \leq 2 \cdot \arcsin(1/4) \leq 0.51$ . At most  $\frac{\beta}{2\pi/\ell} + 1$  boundary regions can have their extreme point in a single invalid region. Hence the total number of invalid boundary regions is at most  $4\left(\frac{\beta}{2\pi/\ell} + 1\right) \leq \ell/2$ .  $\square$

The next lemma shows that a valid boundary region has a large intersection with the corresponding range square.

**Lemma 6.5.** *Let  $R_i$  denote the range square corresponding to boundary region  $i$ . Then the area of the intersection between  $R_i$  and the  $i$ th boundary region is at least  $\min\{(4/\ell)^4, 8\epsilon^2\}$  if  $\ell \geq 4$ .*

*Proof.* Let  $\alpha$  denote the central angle of the polygon. Then  $\alpha = \frac{2\pi}{\ell}$  and  $\delta_i \geq 1 - \cos(\frac{\alpha}{2})$ . By utilizing the inequality  $\cos(\psi) \leq 1 - \frac{1}{2}\psi^2 + \frac{1}{24}\psi^4$  we get  $\delta_i \geq \frac{11}{96}\alpha^2$  for  $\alpha \leq 2$ . Plugging in the value for  $\alpha$  this gives  $\delta_i \geq (4/\ell)^2$  for  $\ell \geq 4$ . The intersection between the range square and the boundary region is minimal when one diagonal of the square is parallel to  $\delta_i$ , see Figure 3(b). Therefore, the area of the intersection is at least  $\delta_i^2 \geq (4/\ell)^4$  if  $\delta_i \leq 4\sqrt{2}\epsilon$  and at least  $8\epsilon^2$  if  $\delta_i \geq 4\sqrt{2}\epsilon$ .  $\square$

**Lemma 6.6.** *If  $\ell \leq \min\{\sqrt[5]{n/(16\epsilon^2)}, n/2\}$ , then every valid boundary region is non-empty with probability at least  $1 - 1/e$ , after perturbation.*

*Proof.* We place  $n/\ell$  input points on the center of a valid range square. The probability that none of these points lies in the boundary region after perturbation is

$$\Pr[\text{boundary region is empty}] \leq \left(1 - \frac{\min\{\delta_i^2, 8\epsilon^2\}}{16\epsilon^2}\right)^{n/\ell},$$

because the area of the intersection is at least  $\min\{\delta_i^2, 8\epsilon^2\}$  and the whole area of the range square is  $16\epsilon^2$ . If  $\delta_i^2 = \min\{\delta_i^2, 8\epsilon^2\}$ , the result follows since

$$\frac{16\epsilon^2}{\min\{\delta_i^2, 8\epsilon^2\}} \leq \frac{16\epsilon^2}{\delta_i^2} \leq 16\epsilon^2 \cdot \ell^4 = 16\epsilon^2 \cdot \ell^5/\ell \leq n/\ell.$$

Here we utilized that  $\delta_i^2 \geq 1/\ell^4$ , which follows from the proof of Lemma 6.5. In the case that  $8\epsilon^2 = \min\{\delta_i^2, 8\epsilon^2\}$  the result follows since  $n/\ell \geq 2$ .  $\square$

Now Theorem 6.2 follows when we choose  $\ell = \Theta(\min\{\sqrt[5]{n}/\epsilon, n\})$ .  $\square$

## 7 Smoothed Motion Complexity

In this section we introduce the concept of *smoothed motion complexity* and give bounds on the smoothed motion complexity of a smallest axis-aligned box enclosing a moving set of points. The task to process a set of continuously moving objects arises in a broad variety of applications such as mobile ad-hoc networks, traffic control systems, and computer graphics (rendering moving objects). Therefore, researchers investigate data structures that can be efficiently maintained under continuous motion, e.g., to answer proximity queries [4], maintain a clustering [11], a convex hull [3], or some connectivity information of the moving point set [12]. In particular, in the framework of kinetic data structures [3] many interesting results on data structures for moving objects have been obtained. Within this framework, the efficiency of a data structure is analyzed with respect to the worst case number of combinatorial changes in the description of the maintained structure that occur during linear (or low degree algebraic) motion. These changes are called (*external*) *events*.

Let us consider the problem of maintaining the smallest axis-aligned bounding box of a moving point set in  $\mathbb{R}^d$  as an example. At any point of time this bounding box can be described by a set of at most  $2d$  points that attain the minimum and maximum value in each of the  $d$  coordinates. If any such minimum/maximum point changes, then an event occurs. We call the worst case number of events with respect to the maintenance of a certain structure under linear motion the *worst case motion complexity*. This is also used in kinetic data structures to measure the quality of the structure.

We introduce *smoothed motion complexity* as an alternative measure for the dynamics of moving data. Smoothed motion complexity asks for the worst case expected performance over all inputs where the expectation is taken with respect to small random noise added to the input. In the context of mobile data this means that both the speed value and the starting position of an input configuration are slightly perturbed by random noise. Thus the smoothed motion complexity is the worst case expected motion complexity over all inputs perturbed in such a way. We believe that smoothed motion complexity is a very natural measure for the dynamics of mobile data since in many applications the exact position of mobile data cannot be determined due to errors caused by physical measurements or fixed precision arithmetic. This is the case when, for instance, the positions of the moving objects are determined via GPS, sensors, and basically in any application involving “real life” data.

We illustrate our approach on the problem to maintain the smallest orthogonal bounding box of a point set moving in  $\mathbb{R}^d$  and show that an upper bound for this problem can be obtained via analyzing left-to-right maxima and the convex hull. The bounding box is a fundamental measure for the extend of a point set and it is useful in many applications, for example, in the construction of  $R$ -trees, for collision detection, and visibility culling.

We are given a set  $P$  of  $n$  points moving in  $\mathbb{R}^d$ . We use  $P(t) = \{p_1(t), \dots, p_n(t)\}$  to denote the set at time  $t$ , where  $p_i(t)$  denotes the position of point  $i$  at time  $t$ . We assume that the movement of each point is a linear function in  $t$ , i. e.,  $p_i(t) = s_i \cdot t + p_i(0)$ , where  $p_i(0)$  is the initial position (at time 0) and  $s_i$  is the speed at which the point moves. We normalize the speed vectors and initial positions such that  $p_i(0), s_i \in [0, 1]^d$ . We use  $p_i^{(j)}(t)$  and  $s_i^{(j)}$  to denote the  $j$ th coordinate of  $p_i(t)$  and  $s_i$ , respectively.

The motion complexity of the problem is the number of combinatorial changes to the set of  $2d$  extreme points defining the bounding box. Clearly, the motion complexity is  $O(d \cdot n)$  in the worst case, 0 in the best case, and  $O(d \cdot \log n)$  in the average case. When we consider smoothed motion complexity we add to each coordinate of the speed vector and each coordinate

of the initial position an i. i. d. random variable from a certain probability distribution with density function  $\varphi: \mathbb{R} \rightarrow \mathbb{R}^+$ . Let  $\bar{P}(t)$  be the perturbed point set at time  $t$ , i. e.,  $\bar{p}_i^{(j)}(t) = (s_i^{(j)} + \nu_{i,j}) \cdot t + p_i^{(j)}(0) + \kappa_{i,j}$ , where  $\nu_{i,j}$  and  $\kappa_{i,j}$  are drawn according to  $\varphi$ .

The smoothed motion complexity (with respect to  $\varphi$ ) is then defined as the maximum over all  $P$  of the expected number of combinatorial changes for  $\bar{P}$ .

We formalize this statement for the case of an axis-aligned bounding box. For any point of time  $t$  define  $\text{bb}(P(t))$  to be the set of points from  $P(t)$  that have minimum or maximum value in at least one coordinate. If for a coordinate the minimum or maximum is attained by more than one point, we take the lexicographically smallest point. We say that a combinatorial change of the bounding box takes place at a point of time  $t$ , if for every  $\epsilon > 0$  we have  $\text{bb}(P(t - \epsilon)) \neq \text{bb}(P(t + \epsilon))$ . By the linearity of motion we have that if for  $t' > t$  we have  $\text{bb}(P(t)) \neq \text{bb}(P(t'))$ , then for every  $t'' \geq t'$  we also have  $\text{bb}(P(t)) \neq \text{bb}(P(t''))$ , i. e., every set of points can define the bounding box only for a fixed interval of time. Hence, the motion complexity can be simply written as the number of distinct sets that define the bounding box over time:  $\text{mc}(P) := |\{\text{bb}(P(t)) \mid t \geq 0\}|$ . For this notation,  $\text{mc}_\varphi(P)$  is a random variable for the number of events that happen after perturbing the points according to  $\varphi$ . The smoothed motion complexity of an axis-aligned bounding box is now given by

$$\max_P \mathbb{E}(\text{mc}_\varphi(P)).$$

In order to obtain bounds on this value, we utilize two observations: First, an upper bound for the one-dimensional problem can be multiplied by  $d$  to yield a bound for the problem in  $d$  dimensions, which is why we can focus on the one-dimensional problem. Second, for the one-dimensional problem the number of external events is strongly related to number of points on the convex hull: If we map each point with initial position  $p_i$  and speed  $s_i$  to the point  $P_i = (p_i, s_i)$  in the two-dimensional plane, then the points on the upper right quarter and the lower left quarter of the convex hull correspond exactly to the external events.

Together with our results on the smoothed number of points on the convex hull we get the following theorem:

**Theorem 7.1.** *The smoothed motion complexity of maintaining an axis-aligned rectangle of a set of points moving in  $\mathbb{R}^d$  is*

$$\max_P \mathbb{E}(\text{mc}_{U[\epsilon]}(P)) \in O(d \cdot (\sqrt{n/\epsilon} + \log n)) \cap \Omega(\min\{\sqrt[5]{n/\epsilon}, n\})$$

and

$$\max_P \mathbb{E}(\text{mc}_{N[\sigma]}(P)) \in O\left(d \cdot \left(\frac{\log^{3/2} n}{\sigma} + \log n\right)\right) \cap \Omega(\sqrt{\log n}).$$

## 8 Conclusion

We have analyzed the smoothed number of left-to-right maxima for different noise distributions. It turns out that for uniform noise from the interval  $[-\epsilon, \epsilon]$ , the smoothed number of left-to-right maxima is  $O(\sqrt{n/\epsilon} + \log n)$  and a matching lower bound (for  $\epsilon > 1/n$ ) shows that it is also  $\Omega(\sqrt{n/\epsilon} + \log n)$ . In contrast to this, the smoothed number of left-to-right maxima under Gaussian noise with standard deviation  $\sigma$  is  $O\left(\frac{\log^{3/2} n}{\sigma} + \log n\right)$ .

We applied the result for uniform noise to the analysis of the smoothed height of binary search trees and the smoothed number of comparisons made by the quicksort algorithm under additive noise. The smoothed height of binary search trees and also the smoothed number of left-to-right maxima are  $\Theta(\sqrt{n/\epsilon} + \log n)$ ; the smoothed number of quicksort comparisons is  $\Theta(\frac{n}{\epsilon+1}\sqrt{n/\epsilon} + n \log n)$ .

While we obtain the average-case height of  $\Theta(\log n)$  for binary search trees only for  $\epsilon \in \Omega(n/\log^2 n)$  – which is large compared to the interval size  $[0, 1]$  from which the numbers are drawn –, for the quicksort algorithm  $\epsilon \in \Omega(\sqrt[3]{n/\log^2 n})$  suffices so that the expected number of comparisons equals the average-case number of  $\Theta(n \log n)$ . On the other hand, the recursion depth of quicksort, which is equal to the tree height, can be as large as  $\Omega(\sqrt{n/\epsilon})$ . Thus, although the average number of comparisons is already reached at  $\epsilon \in \Omega(\sqrt[3]{n/\log^2 n})$ , the recursion depth remains asymptotically larger than its average value for  $\epsilon \in o(n/\log^2 n)$ .

For the smoothed number of points on the convex hull of a set of points, our upper and lower bounds did not quite match – but they still clearly demonstrate the much stronger smoothing effect caused by the Gaussian noise model when compared to the uniform noise model. We also introduced the concept of smoothed motion complexity and showed that a smallest axis-aligned bounding box has small smoothed motion complexity.

A natural next step is the extension of the analysis for the height of binary search trees and the quicksort algorithm to other noise distributions. Also, studying the smoothed motion complexity of other basic structures is an interesting direction of future research.

## Acknowledgements

We thank Heiko Röglin for helpful discussions that led to Theorem 3.4.

## References

- [1] Pankaj K. Agarwal and Sarel Har-Peled. Maintaining approximate extent measures of moving points. In *Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 148–157, 2001.
- [2] Cyril Banderier, René Beier, and Kurt Mehlhorn. Smoothed analysis of three combinatorial problems. In *Proceedings of the 28th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, volume 2747 of *Lecture Notes in Computer Science*, pages 198–207. Springer, 2003.
- [3] Julien Basch, Leonidas J. Guibas, and John Hershberger. Data structures for mobile data. *Journal of Algorithms*, 31(1):1–28, 1999.
- [4] Julien Basch, Leonidas J. Guibas, and Li Zhang. Proximity problems on moving points. In *Proceedings of the 13th ACM Symposium on Computational Geometry (SocG)*, pages 344–351. ACM, 1997.
- [5] Valentina Damerow, Friedhelm Meyer auf der Heide, Harald Räcke, Christian Scheideler, and Christian Sohler. Smoothed motion complexity. In *Proceedings of the 11th European Symposium on Algorithms (ESA)*, volume 2832 of *Lecture Notes in Computer Science*, pages 161–171. Springer, 2003.

- [6] Michael Drmota. An analytic approach to the height of binary search trees II. *Journal of the ACM*, 50(3):333–374, 2003.
- [7] Michael Drmota. Profile and height of random binary search trees. *Journal of the Iranian Statistical Society*, 3(2):117–138, 2004.
- [8] Richard Durrett. *Probability: Theory and Examples*. Duxbury Press, 1991.
- [9] James Allen Fill and Svante Janson. Quicksort asymptotics. *Journal of Algorithms*, 44(1):4–28, 2002.
- [10] Leonidas J. Guibas, John Hershberger, Subhash Suri, and Li Zhang. Kinetic connectivity for unit disks. *Discrete & Computational Geometry*, 25(4):591–610, 2001.
- [11] Sariel Har-Peled. Clustering motion. *Discrete & Computational Geometry*, 31(4):545–565, 2004.
- [12] John Hershberger and Subhash Suri. Simplified kinetic connectivity for rectangles and hypercubes. In *Proceedings of the 12th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 158–167, 2001.
- [13] Rainer Kemp. *Fundamentals of Average Case Analysis of Particular Algorithms*. John Wiley & Sons and B. G. Teubner, 1984.
- [14] David G. Kirkpatrick, Jack Snoeyink, and Bettina Speckmann. Kinetic collision detection for simple polygons. *International Journal of Computational Geometry and Applications*, 12(1-2):3–27, 2002.
- [15] Donald E. Knuth. *Sorting and Searching*, volume 3 of *The Art of Computer Programming*. Addison–Wesley Longman Publishing Co., Redwood City, CA, USA, second edition, 1998.
- [16] Bodo Manthey and Rüdiger Reischuk. Smoothed analysis of binary search trees. *Theoretical Computer Science*, 378(3):292–315, 2007.
- [17] Bodo Manthey and Till Tantau. Smoothed analysis of binary search trees and quicksort under additive noise. In *Proceedings of the 33rd International Symposium on Mathematical Foundations of Computer Science (MFCS)*, volume 5162 of *Lecture Notes in Computer Science*, pages 467–478. Springer, 2008.
- [18] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.
- [19] Bruce Reed. The height of a random binary search tree. *Journal of the ACM*, 50(3):306–332, 2003.
- [20] Alfréd Rényi and Rolf Sulanke. Über die konvexe Hülle von  $n$  zufällig gewählten Punkten. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 2(1):75–84, 1963.
- [21] 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.

- [22] Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms and heuristics: Progress and open questions. In Luis M. Pardo, Allan Pinkus, Endre Süli, and Michael J. Todd, editors, *Foundations of Computational Mathematics, Santander 2005*, pages 274–342. Cambridge University Press, 2006.
- [23] Li Zhang, Harish Devarajan, Julien Basch, and Piotr Indyk. Probabilistic analysis for combinatorial functions of moving points. In *Proceedings of the 13th ACM Symposium on Computational Geometry (SocG)*, pages 442–444, 1997.